# - Quantitative Risk Management -

Prof. Dr. Dietmar Pfeifer

CARL VON OSSIETZKY universität OLDENBURG
FAKULTÄT V
INSTITUT FÜR MATHEMATIK

Last update: 03.07.2017

# Content

## Selected References

**N. Balakrishnan, C.R. Rao (Eds.):** *Order Statistics: Theory & Methods.* Handbook of Statistics 16, Elsevier, Amsterdam 1998.

**M. Bhatia:** *An Introduction to Economic Capital.* Risk Books, London 2009.

**U. Cherubini, E. Luciano, W. Vecchiato:** *Copula Methods in Finance*. Wiley, Chichester 2004.

**C. Cottin, S. Döhler:** *Risikoanalyse*. Modellierung, Beurteilung und Management von Risiken mit Praxisbeispielen. 2. Aufl., Vieweg+Teubner, Wiesbaden 2013.

**M. Cruz (Ed.):** *The Solvency II Handbook*. Developing ERM Frameworks in Insurance and Reinsurance Companies. Risk Books, London 2009.

**H.A. David, H.N. Nagaraja:** *Order Statistics*. 3rd ed., Wiley, N.Y. 2003.

**R. Doff:** *Risk Management for Insurers*. Risk Control, Economic Capital and Solvency II. 3rd ed., Risk Books, London 2015.

**P. Embrechts, C. Klüppelberg, T. Mikosch:** *Modelling Extremal Events*. Springer, Berlin 2001.

## Selected References

**S. Koryciorz:** *Sicherheitskapitalbestimmung und -allokation in der Schadenversicherung.* Eine risikotheoretische Analyse auf der Basis des Value-at-Risk und des Conditional Value-at-Risk. Verlag Versicherungswirtschaft, Karlsruhe 2004.

**G. Hofmann (Hrsg.):** *Basel III und MaRisk*. Regulatorische Vorgaben, bankinterne Verfahren, Risikomanagement. Frankfurt School Verlag, Frankfurt 2011.

**M. Kriele, J. Wolf:** *Wertorientiertes Risikomanagement von Versicherungsunternehmen*. Springer, Berlin 2012.

**A.M. Law, W.D. Kelton:** *Simulation Modeling & Analysis.* 2nd ed., McGraw-Hill, N.Y. 1991.

**J.F. Mai, M. Scherer:** *Simulating Copulas*. Stochastic Models, Sampling Algorithms, and Applications. Imperial College Press, London 2012.

**D.D. Mari, S. Kotz:** *Correlation and Dependence*. Imperial College Press, London 2001.

## Selected References

**A.J. McNeil, R. Frey, P. Embrechts:** *Quantitative Risk Management*. Concepts - Techniques - Tools. Revised ed., Princeton University Press, Princeton 2015.

**R.B. Nelsen:** *An Introduction to Copulas*. 2nd ed., Springer, N.Y. 2006.

**J. Rank (Ed.):** *Copulas*. From Theory to Application in Finance. Risk Books, London 2007.

**R.D. Reiss, M. Thomas:** *Statistical Analysis of Extreme Values*. With Applications to Insurance, Finance, Hydrology and Other Fields. Birkhäuser, Basel 2007.

**R.Y. Rubinstein, D.P. Kroese:** *Simulation and the Monte Carlo Method*. 2nd ed., Wiley, 2008.

**A. Sandstrøm:** *Solvency. Models, Assessment and Regulation*. Chapman & Hall /CRC, Boca Raton, 2006.

**A. Sandstrøm:** *Handbook of Solvency for Actuaries and Risk Managers.* Theory and Practice. Chapman & Hall /CRC, Boca Raton, 2011.

**F. de Weert:** *Bank and Insurance Capital Management*. Wiley, N.Y. 2011.

# I. Risk Measures

Let $\mathfrak{X}$ be a suitable set of non-negative random variables $X$ (risks) on a probability space $(\Omega, \mathcal{A}, P)$. A risk measure $R$ on $\mathfrak{X}$ is a mapping $\mathfrak{X} \to \mathbb{R}^+$ with the following properties:

$P^X = P^Y \Rightarrow R(X) = R(Y)$ for all $X, Y \in \mathfrak{X}$ [distribution invariance]

$R(cX) = c\,R(X)$ for all $X \in \mathfrak{X}$ and $c \geq 0$ [scale invariance]

$R(X + c) = R(X) + c$ for all $X \in \mathfrak{X}$ and $c \geq 0$ [translation invariance]

$R(X) \leq R(Y)$ for all $X, Y \in \mathfrak{X}$ with $X \leq Y$ [monotonicity]

The risk measure is called *coherent*, if it additionally has the property:

$R(X + Y) \leq R(X) + R(Y)$ for all $X, Y \in \mathfrak{X}$ [subadditivity]

This last property is the crucial point: it induces a *diversification effect* for *arbitrary* risks $X_1, \cdots, X_n$ (dependent or not) since it follows by induction that coherent risk measures have the property

$$R\left(\sum_{k=1}^{n} X_k\right) \leq \sum_{k=1}^{n} R(X_k) \qquad \text{for any } n \in \mathbb{N}.$$

The popular standard deviation principle *SDP* which is sometimes used for calculating premiums in insurance is defined as

$$SDP(X) = E(X) + \gamma\sqrt{Var(X)} \qquad \text{for a fixed } \gamma > 0 \text{ and } X \in \mathfrak{X} = \mathfrak{L}^2_+(\Omega, \mathcal{A}, P),$$

the set of non-negative square-integrable random variables on $(\Omega, \mathcal{A}, P)$.

*SDP* is coherent, but it is *not* a risk measure in the strict sense because it is *not monotone*.

The risk measure used in Basel II/III and Solvency II is the Value-at-Risk VaR, being defined as a (typically high) quantile of the risk distribution:

$$\mathrm{VaR}_\alpha(X) := Q_X(1-\alpha) \qquad \text{for } X \in \mathfrak{X} \text{ and } 0 < \alpha < 1,$$

where $Q_X$ denotes the quantile function

$$Q_X(u) := \inf\{x \in \mathbb{R} | P(X \le x) \ge u\} \qquad \text{for } 0 < u < 1.$$

VaR is a proper risk measure, but not coherent in general.

The "smallest" coherent risk measure above VaR is the expected shortfall ES (used in the Swiss Solvency Test - SST), which is in general defined as

$$\mathrm{ES}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_u(X)\,du \qquad (*)$$

for $0 < \alpha < 1$, where $\mathbb{1}_A$ denotes the indicator random variable of some event (measurable set) $A$. Equivalent representations are

$$\mathrm{ES}_\alpha(X) = \mathrm{VaR}_\alpha(X) + E\big(X - \mathrm{VaR}_\alpha(X) \,|\, X \geq \mathrm{VaR}_\alpha(X)\big) \cdot \frac{P\big(X \geq \mathrm{VaR}_\alpha(X)\big)}{\alpha}$$ and

$$\mathrm{ES}_\alpha(X) := \frac{1}{\alpha}\Big\{E\Big(X \cdot \mathbb{1}_{\{X \geq \mathrm{VaR}_\alpha(X)\}}\Big) - \mathrm{VaR}_\alpha(X)\big[P\big(X \geq \mathrm{VaR}_\alpha(X)\big) - \alpha\big]\Big\}. \qquad (**)$$

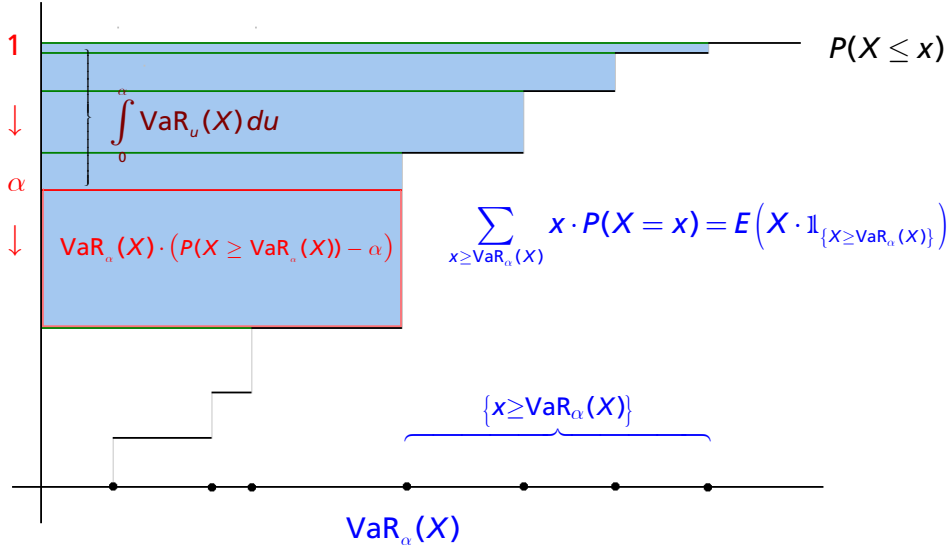
In case that $P\big(X \geq \mathrm{VaR}_\alpha(X)\big) = \alpha$, $\mathrm{ES}_\alpha(X)$ is also equivalent to

$$\mathrm{ES}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_u(X)\,du = E\big(X \,|\, X \geq \mathrm{VaR}_\alpha(X)\big)$$

Note that in general, we have

$$P\big(X \geq \mathrm{VaR}_\alpha(X)\big) \geq \alpha \geq P\big(X > \mathrm{VaR}_\alpha(X)\big).$$

Graphical sketch of proof for the equivalence of (*) and (**):



Example 1: Let the risks $X$ and $Y$ be independent and uniformly distributed over the interval $[0,1]$. Then, for $\alpha < \frac{1}{2}$,

$$\mathrm{VaR}_\alpha(X) = \mathrm{VaR}_\alpha(Y) = 1-\alpha \quad \text{and} \quad \mathrm{VaR}_\alpha(X+Y) = 2-\sqrt{2\alpha} < \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y),$$

$$\mathrm{ES}_\alpha(X) = \mathrm{ES}_\alpha(Y) = 1-\frac{\alpha}{2} \quad \text{and} \quad \mathrm{ES}_\alpha(X+Y) = 2-\frac{2}{3}\sqrt{2\alpha} < \mathrm{ES}_\alpha(X) + \mathrm{ES}_\alpha(Y),$$

hence there is a strict diversification effect for both risk measures.

For $Y = X$, however, we obtain

$$\mathrm{VaR}_\alpha(X) = \mathrm{VaR}_\alpha(Y) = 1-\alpha \quad \text{and} \quad \mathrm{VaR}_\alpha(X+Y) = 2-2\alpha = \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y),$$

$$\mathrm{ES}_\alpha(X) = \mathrm{ES}_\alpha(Y) = 1-\frac{\alpha}{2} \quad \text{and} \quad \mathrm{ES}_\alpha(X+Y) = 2-\alpha = \mathrm{ES}_\alpha(X) + \mathrm{ES}_\alpha(Y),$$

hence there is no strict diversification effect for both risk measures.

Some formulas for normal and lognormal risk distributions:

1. Case $P^X = \mathcal{N}(\mu, \sigma^2)$:

$$\mathrm{VaR}_\alpha(X) = \mu + u_{1-\alpha}\sigma, \quad \mathrm{ES}_\alpha(X) = \mu + \frac{\varphi(u_{1-\alpha})}{\alpha}\sigma$$

with $\varphi$ denoting the density, $\Phi$ the cumulative distribution function and $u_\beta = \Phi^{-1}(\beta)$ the $\beta$-quantile of $\mathcal{N}(0,1)$.

Note that here $\mathrm{VaR}_\alpha(X)$ and $\mathrm{ES}_\alpha(X)$ are exactly *SDP*'s for $\alpha < \frac{1}{2}$, hence both risk measures are coherent in this case!

2. Case $P^X = \mathcal{LN}(\mu, \sigma^2)$:

$$\mathrm{VaR}_\alpha(X) = \exp(\mu + u_{1-\alpha}\sigma), \quad ES_\alpha(X) = \frac{1}{\alpha}\exp\left(\mu + \frac{\sigma^2}{2}\right)\left(1 - \Phi(u_{1-\alpha} - \sigma)\right).$$

Some remarks on the interplay between correlation and diversification under the risk measure $\mathrm{VaR}_\alpha(X)$:

" 'Diversification effects' means the reduction in the risk exposure of insurance and reinsurance undertakings and groups related to the diversification of their business, resulting from the fact that the adverse outcome from one risk can be offset by a more favourable outcome from another risk, where those risks are not fully correlated. The Basic Solvency Capital Requirement shall comprise individual risk modules, which are aggregated […] The correlation coefficients for the aggregation of the risk modules […], shall result in an overall Solvency Capital Requirement […] Where appropriate, diversification effects shall be taken into account in the design of each risk module."

[Directive 2009/138/EC, (64) p. 7; (37) p. 24; Article 104, p. 52]

Example 2: Let the joint distribution of the risks $X$ and $Y$ be given by the following table (where $\alpha = 0.005$ (Solvency II standard)):

| $P(X=x, Y=y)$ | | $x$ | | | | |
|---|---|---|---|---|---|---|
| | | 0 | 50 | 100 | $P(Y=y)$ | $P(Y \le y)$ |
| $y$ | 0 | $\beta$ | $0.440-\beta$ | 0.000 | 0.440 | 0.440 |
| | 40 | $0.554-\beta$ | $\beta$ | 0.001 | 0.555 | **0.995** |
| | 50 | 0.000 | 0.001 | 0.004 | 0.005 | 1.000 |
| | $P(X=x)$ | 0.554 | 0.441 | 0.005 | | |
| | $P(X \le x)$ | 0.554 | **0.995** | 1.000 | | |

with $0 \le \beta \le 0.440$, giving $\mathrm{VaR}_\alpha(X) = 50$, $\mathrm{VaR}_\alpha(Y) = 40$.

For the moments of $X$ and $Y$, we obtain, with $\sigma$ denoting the standard deviation:

| $E(X)$ | $E(Y)$ | $\sigma(X)$ | $\sigma(Y)$ | $\rho(\beta) = \rho(X,Y)$ |
|---|---|---|---|---|
| 22.550 | 22.450 | 25.377 | 19.912 | $-0.9494 + 3.9579\,\beta$ |

which shows that the range of possible risk correlations is the interval $[-0.9494;\ 0.7921]$, with a zero correlation being attained for $\beta = 0.2399$.

The following table shows the distribution of the aggregated risk $S = X + Y$:

| $s$ | 0 | 40 | 50 | 90 | 100 | 140 | 150 |
|---|---|---|---|---|---|---|---|
| $P(S = s)$ | $\beta$ | $0.554 - \beta$ | $0.440 - \beta$ | $\beta$ | 0.001 | 0.001 | 0.004 |
| $P(S \leq s)$ | $\beta$ | 0.554 | $0.994 - \beta$ | 0.994 | **0.995** | 0.996 | 1.000 |

giving a *risk concentration* (as opposite to *risk diversification*) due to

$$\mathrm{VaR}_\alpha(S) = 100 > 90 = \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y),$$

independent of the parameter $\beta$ and hence also independent of the possible correlations between $X$ and $Y$!

# II. Copulas

A copula (in $d$ dimensions) is a function $C$ defined on the unit cube $[0,1]^d$ with the following properties:

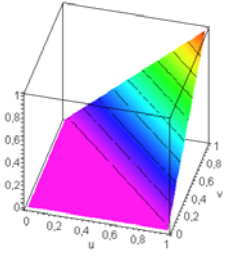
- the range of $C$ is the unit interval $[0,1]$;
- $C(\mathbf{u})$ is zero for all $\mathbf{u} = (u_1, \cdots, u_d)$ in $[0,1]^d$ for which at least one coordinate is zero;
- $C(\mathbf{u}) = u_k$ if all coordinates of $\mathbf{u}$ are 1 except the $k$-th one;
- $C$ is $d$-increasing in the sense that for every $\mathbf{a} \leq \mathbf{b}$ in $[0,1]^d$ the volume assigned by $C$ to the subinterval $[\mathbf{a}, \mathbf{b}] = [a_1, b_1] \times \cdots \times [a_d, b_d]$ is non-negative.

A copula can alternatively be characterized as a multivariate distribution function with univariate marginal distribution functions that belong to a continuous uniform distribution over the unit interval $[0,1]$.

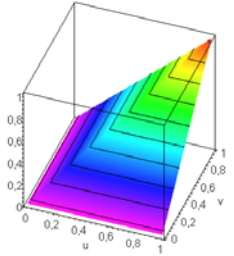
Every copula is bounded by the so-called Fréchet-Hoeffding bounds, i.e.

$$C_*(\mathbf{u}) := \max(u_1 + \cdots + u_d - d + 1, 0) \leq C(u_1, \cdots, u_d) \leq C^*(\mathbf{u}) := \min(u_1, \cdots, u_d).$$

The upper Fréchet-Hoeffding bound $C^*$ is a copula itself for *any* dimension; the lower Fréchet-Hoeffding bound $C_*$ is a copula in *two* dimensions only.




Representations: $(U, 1-U)$ or $(1-U, U)$




$(U, U)$ [with uniform $U$]

**Sklar's Theorem:** Let $H$ denote some $d$-dimensional distribution function with marginal distribution functions $F_1, \cdots, F_d$. Then there exists a copula $C$ such that for all real $(x_1, \cdots, x_d)$,

$$H(x_1, \cdots, x_d) = C\big(F_1(x_1), \cdots, F_d(x_d)\big).$$

If all the marginal distribution functions are continuous, then the copula is unique. Moreover, the converse of the above statement is also true in case of continuity. If we denote by $F_1^{-1}, \cdots, F_d^{-1}$ the generalized inverses of the marginal distribution functions (or quantile functions), then for every $(u_1, \cdots, u_d)$ in the unit cube,

$$C(u_1, \cdots, u_d) = H\big(F_1^{-1}(u_1), \cdots, F_d^{-1}(u_d)\big).$$

Example 3: Consider a random vector $(X,Y)$ with the density

$$f(x,y)=\begin{cases} x+y & \text{if } 0\le x,y\le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then by integration, we obtain for the cdf $F$

$$F(x,y)=\frac{1}{2}xy(x+y) \text{ for } 0\le x,y\le 1.$$

The quantile functions are given by

$$Q_X(u)=Q_Y(u)=-\frac{1}{2}+\frac{1}{2}\sqrt{1+8u} \text{ for } 0\le u\le 1.$$

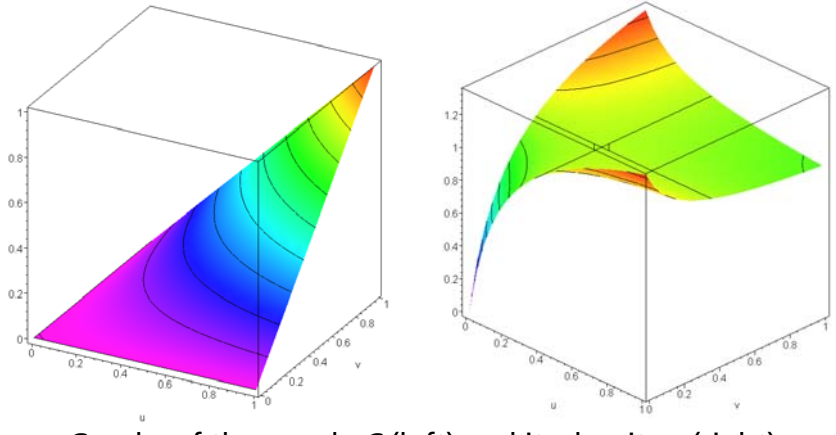
Hence by Sklar's Theorem, the corresponding copula is given by

$$C(u,v)=F\big(Q_X(u),Q_Y(v)\big)=\frac{1}{16}\Big(-1+\sqrt{1+8u}\Big)\Big(-1+\sqrt{1+8v}\Big)\Big(-2+\sqrt{1+8u}+\sqrt{1+8v}\Big)$$

for $0\le u,v\le 1$.

For the copula density we obtain

$$c(u,v) = \frac{\partial^2}{\partial u\,\partial v} C(u,v) = 2\frac{\sqrt{1+8u}+\sqrt{1+8v}-2}{\sqrt{1+8u}\sqrt{1+8v}} \quad \text{for } 0 \le u,v \le 1.$$




Graphs of the copula $C$ (left) and its density $c$ (right)

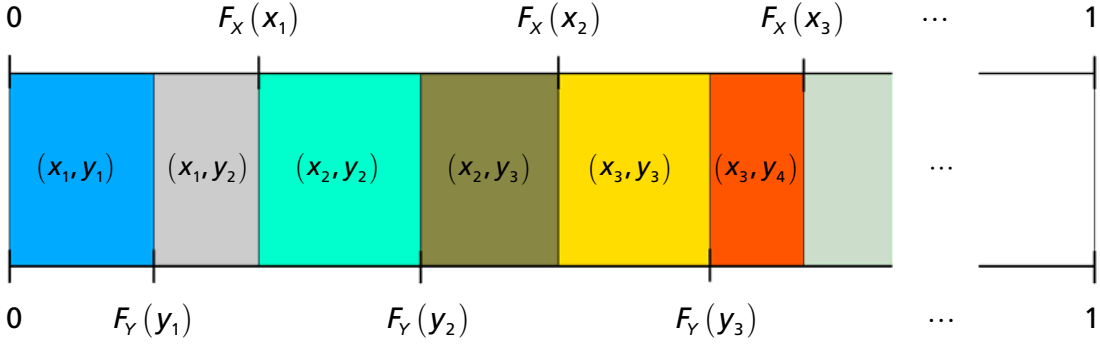
If $X$ is any real random variable, then

- the random vector $\mathbf{X} = (X, X, \cdots, X)$ with $d$ components possesses the upper Fréchet-Hoeffding bound $C^*$ as copula
- the random vector $\mathbf{X} = (X, -X)$ with two components possesses the lower Fréchet-Hoeffding bound $C_*$ as copula.

Random variables who have $C^*$ or $C_*$ as copula are called *comonotone* or *countermonotone*, resp.

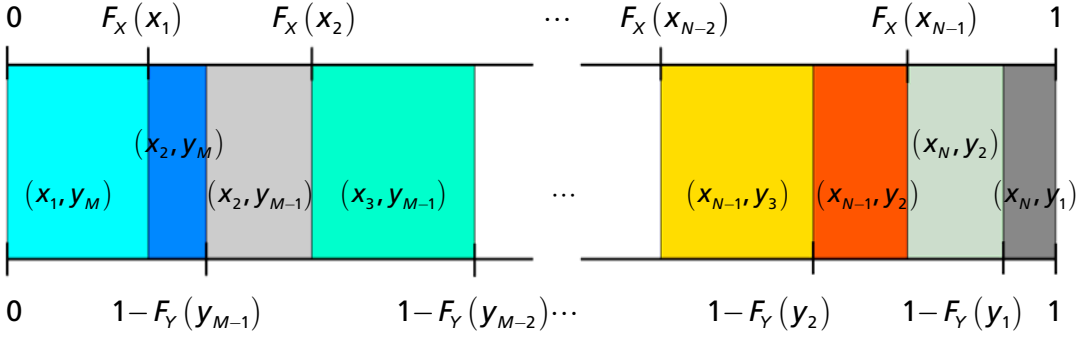
Any copula is invariant against (the same type of) monotone transformations of the marginal random variables / distributions.

Here is a canonical construction of comonotone random variables $X$ and $Y$ with a discrete distribution over the set $\{x_1, x_2, \cdots\}$ with $x_1 < x_2 < \cdots$ for $X$ and the set $\{y_1, y_2, \cdots\}$ with $y_1 < y_2 < \cdots$ for $Y$:




The occurrence probability of a pair $(x_i, y_j)$ corresponds to the width of the corresponding coloured rectangle.

Here is a canonical construction of countermonotone random variables $X$ and $Y$ with a discrete distribution over the set $\{x_1, x_2, \cdots, x_N\}$ with $x_1 < x_2 < \cdots < x_N$ for $X$ and the set $\{y_1, y_2, \cdots, y_M\}$ with $y_1 < y_2 < \cdots < y_M$ for $Y$:




(similarly for unbounded random variables)

The occurrence probability of a pair $(x_i, y_j)$ corresponds to the width of the corresponding coloured rectangle.

Example 4: Consider random variables X and Y with the following marginal distributions:

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X = x)$ | 0.1 | 0.3 | 0.6 |
| $F_X(x)$ | 0.1 | 0.4 | 1.0 |

| $y$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(Y = y)$ | 0.2 | 0.3 | 0.1 | 0.4 |
| $F_Y(y)$ | 0.2 | 0.5 | 0.6 | 1.0 |
| $1-F_Y(y)$ | 0.8 | 0.5 | 0.4 | 0.0 |

Comonotone case:

| $P(X = x, Y = y)$ | | $x$ | | |
|---|---|---|---|---|
| | | 0 | 1 | 2 |
| $y$ | 0 | 0.1 | 0.1 | 0.0 |
| | 1 | 0.0 | 0.2 | 0.1 |
| | 2 | 0.0 | 0.0 | 0.1 |
| | 3 | 0.0 | 0.0 | 0.4 |

Example 4: Consider random variables X and Y with the following marginal distributions:

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X=x)$ | 0.1 | 0.3 | 0.6 |
| $F_X(x)$ | 0.1 | 0.4 | 1.0 |

| $y$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(Y=y)$ | 0.2 | 0.3 | 0.1 | 0.4 |
| $F_Y(y)$ | 0.2 | 0.5 | 0.6 | 1.0 |
| $1-F_Y(y)$ | 0.8 | 0.5 | 0.4 | 0.0 |

Countermonotone case:

| $P(X=x, Y=y)$ | | $x$ | | |
|---|---|---|---|---|
| | | 0 | 1 | 2 |
| $y$ | 0 | 0.0 | 0.0 | 0.2 |
| | 1 | 0.0 | 0.0 | 0.3 |
| | 2 | 0.0 | 0.0 | 0.1 |
| | 3 | 0.1 | 0.3 | 0.0 |

Correlations can be expressed through the copula, but not conversely:

**Hoeffding's Lemma:** Let $(X,Y)$ be a random vector with a copula $C$ and marginal cdf's $F_X$ and $F_Y$ such that $E(|X|)<\infty$, $E(|Y|)<\infty$ and $E(|XY|)<\infty$. Then the covariance between $X$ and $Y$ can be expressed in the following way through the copula $C$:

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[C(F_X(x),F_Y(y))-F_X(x)F_Y(y)\right]dx\,dy.$$

**Attainable Correlations Theorem:** Let $(X,Y)$ be a random vector with a copula $C$ and marginal cdf's $F_X$ and $F_Y$ such that $E(|X|)<\infty$, $E(|Y|)<\infty$, $E(|XY|)<\infty$ and the variances $\sigma_X^2$ and $\sigma_Y^2$ are strictly positive.

- The attainable correlations $\rho(X,Y)=\dfrac{Cov(X,Y)}{\sigma_X\sigma_Y}$ form a closed interval $[\rho_{\min},\rho_{\max}]\subseteq[-1,1]$ with $\rho_{\min}<0<\rho_{\max}$.
- The minimum correlation $\rho(X,Y)=\rho_{\min}$ is attained iff $X$ and Y are counter-monotonic. The maximum correlation $\rho(X,Y)=\rho_{\max}$ is attained iff $X$ and Y are comonotonic.

Note that the extreme cases of $\rho_{\min}=-1$ and $\rho_{\max}=1$ are attained iff $X$ and $Y$ are almost surely *linearly* dependent.

**No Diversification Theorem (Pfeifer 2012):** Let $X$ and $Y$ be risks with cdf's $F_X$ and $F_Y$, resp. which are continuous and strictly increasing on their support. Denote, for a fixed $\alpha \in (0,1)$,

$$Q^*(\alpha,\delta) := \min\{Q_X(u) + Q_Y(2-\alpha-\delta-u) | 1-\alpha-\delta \le u \le 1\} \quad \text{for } 0 \le \delta < 1-\alpha.$$

Then there exists a sufficiently small $\varepsilon \in (0, 1-\alpha)$ with the property

$$Q^*(\alpha,\varepsilon) > Q_X(1-\alpha) + Q_Y(1-\alpha) = \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y).$$

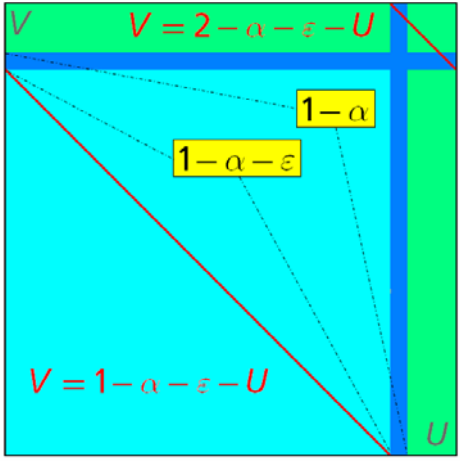
Assume further that the random vector $(U,V)$ has a copula $C$ as joint distribution function with the properties

$$V < 1-\alpha-\varepsilon \text{ iff } U < 1-\alpha-\varepsilon \text{ and } V = 2-\alpha-\varepsilon-U \text{ iff } U \ge 1-\alpha-\varepsilon.$$

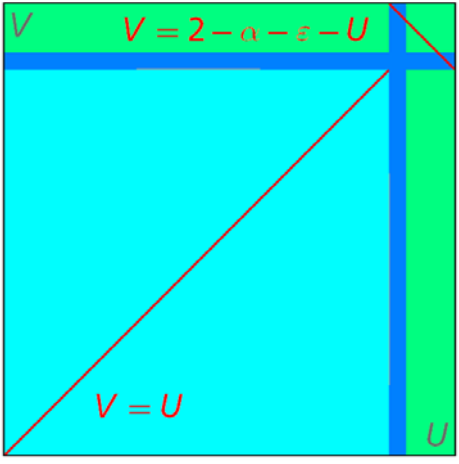
If we define $X^* := Q_X(U)$, $Y^* := Q_Y(V)$, $S^* := X^* + Y^*$, then $(X^*, Y^*)$ has the same marginal distributions as $(X, Y)$, and it holds

$$\mathrm{VaR}_\alpha\left(X^* + Y^*\right) \geq Q^*(\alpha, \varepsilon) > \mathrm{VaR}_\alpha\left(X^*\right) + \mathrm{VaR}_\alpha\left(Y^*\right) = \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y),$$

i.e. there is a risk concentration effect. Moreover, the correlation $\rho\left(X^*, Y^*\right)$ is minimal if $V = 1 - \alpha - \varepsilon - U$ for $U < 1 - \alpha - \varepsilon$ (lower extremal copula $\underline{C}$) and maximal if $V = U$ for $U < 1 - \alpha - \varepsilon$ (upper extremal copula $\overline{C}$).




lower extremal copula $\underline{C}$



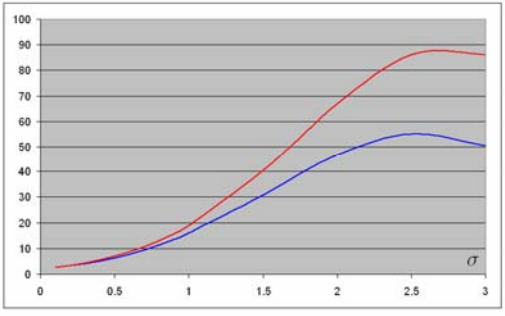

upper extremal copula $\overline{C}$

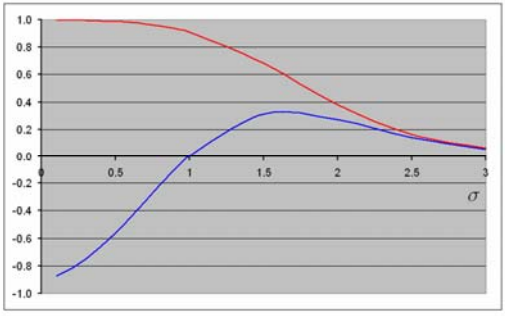
Example 5: Assume that the risks $X$ and $Y$ follow the same lognormal distribution $\mathcal{LN}\left(-\frac{\sigma^2}{2}, \sigma^2\right)$ with $\sigma > 0$ which corresponds to $E(X) = E(Y) = 1$.

The following table shows some numerical results for the extreme copulas $\underline{C}$ and $\overline{C}$ in the last Theorem, especially the maximal range of correlations induced by them. According to the Solvency II standard, we choose $\alpha = 0.005$ (and $\varepsilon = 0.001$, which will be sufficient here).

| $\sigma$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 |
|---|---|---|---|---|---|---|---|
| $\mathrm{VaR}_\alpha(X) = \mathrm{VaR}_\alpha(Y)$ | 1.2873 | 1.6408 | 2.0704 | 2.5866 | 3.1992 | 3.9177 | 4.7497 |
| $\mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y)$ | 2.5746 | 3.2816 | 4.1408 | 5.1732 | 6.3984 | 7.8354 | 9.4994 |
| $\mathrm{VaR}_\alpha\left(X^* + Y^*\right)$ | 2.6205 | 3.3994 | 4.3661 | 5.5520 | 6.9901 | 8.7134 | 10.7537 |
| $\rho_{\min}\left(X^*, Y^*\right)$ | –0.8719 | –0.8212 | –0.7503 | –0.6620 | –0.5598 | –0.4480 | –0.3310 |
| $\rho_{\max}\left(X^*, Y^*\right)$ | 0.9976 | 0.9969 | 0.9951 | 0.9920 | 0.9873 | 0.9802 | 0.9700 |

$$P^X = P^Y = \mathcal{LN}\left(-\frac{\sigma^2}{2}, \sigma^2\right)$$




Left: graph of $\mathrm{VaR}_\alpha\left(X^*+Y^*\right)$ and $\mathrm{VaR}_\alpha(X)+\mathrm{VaR}_\alpha(Y)$ as functions of $\sigma$

Right: graph of $\rho_{\max}\left(X^*,Y^*\right)$ and $\rho_{\min}\left(X^*,Y^*\right)$ as functions of $\sigma$

In general, three types of copulas can be distinguished:

1. Copulas which can be described only implicitly, but can be constructed explicitly (e.g. Gauß-, *t*- and, more generally, elliptic copulas)

2. Copulas which can be described explicitly, but which are difficult to construct (e.g. Archimedean copulas)

3. Copulas which can be described **and** constructed explicitly (e.g. vine copulas, checkerboard copulas, Bernstein copulas).

Examples for type 1 copulas ($\Sigma =$ positive definite correlation matrix):

*Gauß copula* $C_\Sigma^G$ :

$$C_\Sigma^G(u_1,\cdots,u_d) = \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} \exp\left(-\frac{1}{2}\mathbf{v}^T\Sigma^{-1}\mathbf{v}\right) dv_1 \cdots dv_d$$

*t-Copula* $C_\Sigma^{t_\nu}$ with $\nu \in \mathbb{N}$ degrees of freedom:

$$C_\Sigma^{t_\nu}(u_1,\cdots,u_d) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \cdots \int_{-\infty}^{t_\nu^{-1}(u_d)} \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^d \det(\Sigma)}} \left(1+\frac{1}{\nu}\mathbf{v}^T\Sigma^{-1}\mathbf{v}\right)^{\left(-\frac{\nu+d}{2}\right)} dv_1 \cdots dv_d$$

**Gauß copula**: Suppose that the random vector $\mathbf{X} = (X_1, \cdots, X_d)^T$ follows a multivariate normal distribution $\mathcal{N}(\mathbf{0}, \Sigma)$ with zero expectation and some positive definite correlation matrix $\Sigma$. Then the joint cumulative distribution function of the random vector $\mathbf{Z} = (Z_1, \cdots, Z_d)^T$ with

$$Z_k = \Phi(X_k), \ k = 1, \cdots, d$$

is the Gaussian copula $C_\Sigma^G$, i.e. $C_\Sigma^G$ is also the underlying copula for $\mathbf{X}$.

Using the linear structure of jointly normally distributed random variables a possible construction for a Gauß copula can be obtained as follows:

- Decompose the given matrix $\Sigma$ as a product $\Sigma = AA^T$, e.g. using a *spectral representation* [via eigenvalues and eigenvectors] or using a *Cholesky-decomposition.*
- If $\mathbf{Z}$ is a random vector of $d$ independent standard normally distributed components $Z_1, \cdots, Z_d$, then $\mathbf{X}$ can be stochastically represented as $\mathbf{X} = A\mathbf{Z}$.

The *spectral decomposition* of an arbitrary positive definite symmetric matrix $M$ is characterized by the following matrix product:

$$M = T\Delta T^{-1} = T\Delta T^{T}$$

where $\Delta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{bmatrix}$ is the diagonal matrix of the (positive) eigenvalues of $M$ and $T$ is an orthonormal matrix consisting of the corresponding eigenvectors. The required transformation matrix $A$ is then given by

$$A = T\Delta^{1/2} \quad \text{with} \quad \Delta^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_d} \end{bmatrix}.$$

Example 6: Let $\mathbf{I}$ denote the unit matrix. The matrix $M = \begin{bmatrix} 5 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 5 \end{bmatrix}$ involves the characterististic polynomial

$$\varphi(\lambda) = \det(M - \lambda \mathbf{I}) = -\lambda^3 + 11\lambda^2 - 11\lambda + 1$$

with the three zeros

$$\lambda_1 = 1, \quad \lambda_{2,3} = 5 \pm 2\sqrt{6}$$

and a possible orthonormal matrix

$$T = \begin{bmatrix} -0.7071 & 0.6739 & -0.2142 \\ 0.0000 & 0.3029 & 0.9530 \\ 0.7071 & 0.6739 & -0.2142 \end{bmatrix}.$$

This implies

$$A = T\Delta^{1/2} = \begin{bmatrix} -0.7071 & 2.1202 & -0.0681 \\ 0.0000 & 0.9530 & 0.3029 \\ 0.7071 & 2.1202 & -0.0681 \end{bmatrix}.$$

Note that the spectral decomposition method is quite tedious especially in higher dimensions.

In comparison, the *Cholesky decomposition* is more efficient in general. W.l.o.g., we can assume that $A$ is a *lower triangular matrix*:

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{bmatrix}$$

This implies

$$M = \left[m_{ij}\right] = AA^T = \begin{bmatrix} a_{11}^2 & a_{11}a_{21} & \cdots & a_{11}a_{d1} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 & \cdots & a_{21}a_{d1} + a_{22}a_{d2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1}a_{11} & a_{d1}a_{21} + a_{d2}a_{22} & \cdots & \sum_{k=1}^{d} a_{dk}^2 \end{bmatrix}.$$

This equation can be solved recursively, giving

$$a_{11} = \sqrt{m_{11}},\ a_{kk} = \sqrt{m_{kk} - \sum_{i=1}^{k-1} a_{ki}^2},\ a_{k1} = \frac{m_{k1}}{a_{11}},\ a_{kj} = \frac{m_{kj} - \sum_{i=1}^{j-1} a_{ki}a_{ji}}{a_{jj}},\quad 1 \le j \le k \le d.$$

In the example above, $M = \begin{bmatrix} 5 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 5 \end{bmatrix}$ giving

$$A = \begin{bmatrix} \sqrt{5} & 0 & 0 \\ \frac{2}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} & 0 \\ \frac{4}{5}\sqrt{5} & \frac{2}{5}\sqrt{5} & 1 \end{bmatrix} = \begin{bmatrix} 2.2361 & 0.0000 & 0.000 \\ 0.8944 & 0.4472 & 0.000 \\ 1.7889 & 0.8944 & 1.000 \end{bmatrix}.$$

**Remark:** The equation $M = AA^T$ can in general have infinitely many different solutions, for instance here also

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

In the 2-dimensional case dependent standard-normally distributed random vectors $\mathbf{X} = (X_1, X_2)^T$ can also be represented by the Box-Muller-transformation

$$X_1 := \sqrt{-2\ln(U)} \cdot \cos(2\pi V)$$
$$X_2 := \sqrt{-2\ln(U)} \cdot \cos(2\pi V + \alpha),$$

where U and $V$ are continuous uniformly distributed random variables over $[0,1]$ and $\alpha \in [0,\pi]$ is an arbitrary angle. Note that, denoting $R := \sqrt{-2\ln(U)}$, it follows that $R^2$ has a $\mathcal{E}(1/2)$-exponential distribution, hence

$$Cov(X_1, X_2) = E(R^2) \int_0^1 \cos(2\pi v) \cdot \cos(2\pi v + \alpha)\, dv = \frac{E(R^2)}{2} \cos(\alpha) = \cos(\alpha).$$

For the correlation this means

$$\rho(X_1, X_2) = \cos(\alpha) \in [-1,1].$$

For $\alpha = \frac{\pi}{2}$ we have $\cos(2\pi V + \alpha) = -\sin(2\pi V)$, in which case $X_1$ and $X_2$ are uncorrelated. Due to symmetry, the negative sign can also be dropped here.

***t*-Copula:** This type of a copula is derived from a multivariate *t*-distribution $t_\nu(\boldsymbol{\mu}, \Sigma)$ with $\nu$ degrees of freedom ($\nu \in \mathbb{N}$) which is obtained from a multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ via a variance mixture. To be more precise, a random vector $\mathbf{X}$ with such a distribution can be represented by a $\mathcal{N}(\mathbf{0}, \Sigma)$-distributed random vector $\mathbf{Z}$ and a $\chi^2_\nu$-distributed random variable $W$, independent from $\mathbf{Z}$, by

$$\mathbf{X} := \boldsymbol{\mu} + \sqrt{\frac{\nu}{W}}\mathbf{Z}.$$

Note that the variance-covariance matrix of $\mathbf{X}$ only exists if $\nu \geq 3$, and is then given by $\frac{\nu}{\nu - 2}\Sigma$.

In a similar way we can represent random vectors $\mathbf{X}$ that follow so called elliptic distributions as

$$\mathbf{X} := \mu + ARS$$

where $A$ is again a suitable linear transformation matrix. Here $R$ is a non-negative random variable, independent of the random vector $\mathbf{S}$ which is continuous uniformly distributed over the $d$-dimensional unit sphere $\mathcal{S} := \left\{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{x}^T\mathbf{x} = 1\right\}$.

Such a vector $\mathbf{S}$ can generally be represented as

$$\mathbf{S} = \frac{1}{\|\mathbf{Z}\|}\mathbf{Z}$$

where e.g. $\mathbf{Z}$ is standard normally $\mathcal{N}(\mathbf{0}, \mathbf{I})$-distributed.

## Archimedean Copulas:

These are characterised by their so called *generator* $\varphi$ via

$$C_d(u_1,\cdots,u_d) = \varphi^{-1}\left(\sum_{i=1}^{d}\varphi(u_i)\right) \quad \text{for} \quad u_1,\cdots,u_d \in [0,1].$$

Special case:

$$\varphi(x) = -\ln x$$

with

$$C_d(u_1,\cdots,u_n) = \varphi^{-1}\left(\sum_{i=1}^{d}\varphi(u_i)\right) = \exp\left(\sum_{i=1}^{d}\ln u_i\right) = \prod_{i=1}^{d} u_i$$

(independence copula)

*Strict* generators are characterized by the following result:

**Generator Theorem:** Let $\varphi : (0,1] \to \mathbb{R}$ be continuous, strictly decreasing and convex with $\varphi(1) = 0$ and $\lim_{z \downarrow 0} \varphi(z) = \infty$. If $\varphi^{-1}$ denotes the inverse mapping on the interval $[0, \infty)$, then

$$C_d(u_1, \cdots, u_d) = \varphi^{-1}\left( \sum_{i=1}^{d} \varphi(u_i) \right) \quad \text{for} \quad u_1, \cdots, u_d \in [0,1]$$

is a copula for $d = 2$. It is a copula **for all** $d \geq 2$ iff $\varphi^{-1}$ is *completely monotone,* i.e. iff

$$(-1)^k \frac{d^k}{ds^k} \varphi^{-1}(s) \geq 0 \quad \text{for all } k \in \mathbb{N} \text{ and } s > 0.$$

**Bernstein's Theorem:** Each inverse $\varphi^{-1}$ of a strict generator can be represented as the Laplace transform of a suitable non-negative random variable $Z$ via

$$\varphi^{-1}(s) = E\left(e^{-sZ}\right), \ s \geq 0.$$

Note that in the special case $Z \equiv 1$ we have $\varphi^{-1}(s) = e^{-s}, \ s \geq 0,$ which leads to the independence copula (in any dimension).

Examples of Archimedean copulas with strict generators:

*Clayton-Copula*:

$$C_\theta(u_1,\cdots,u_d)=\left[\sum_{i=1}^{d}u_i^{-\theta}-d+1\right]^{-1/\theta},\ \mathbf{u}\in(0,1]^d,\ \theta>0$$

with $P^Z=\Gamma(\alpha,\alpha)$ for $\alpha=\frac{1}{\theta}>0$ and density $f_Z(z)=\frac{z^{\alpha-1}}{\Gamma(\alpha)}\alpha^\alpha e^{-\alpha z},\ z>0,$

i.e. $\varphi^{-1}(s)=E\left(e^{-sZ}\right)=\left(\frac{\alpha}{\alpha+s}\right)^\alpha,\ s\geq 0,$ with $\varphi(t)=\frac{t^{-\theta}-1}{\theta},\ t\in(0,1].$

Examples of Archimedean copulas with strict generators:

*Frank-Copula*:

$$C_\theta(u_1,\cdots,u_d) = -\frac{1}{\theta}\ln\left(1+\left(e^{-\theta}-1\right)\prod_{i=1}^{d}\left\{\frac{e^{-\theta u_i}-1}{e^{-\theta}-1}\right\}\right),\ \mathbf{u}\in(0,1]^d,\ \theta>0$$

with $P^Z = \mathcal{LS}\left(e^{-\theta}\right)$ (log-series distribution over $\mathbb{N}$) for $\theta>0$, i.e.

$$\varphi^{-1}(s) = E\left(e^{-sZ}\right) = -\frac{1}{\theta}\sum_{k=1}^{\infty}\frac{\left(\left(1-e^{-\theta}\right)e^{-s}\right)^k}{k} = -\frac{\ln\left(1-\left(1-e^{-\theta}\right)e^{-s}\right)}{\theta},\ s\geq 0,$$ with

$$\varphi(t) = -\ln\frac{1-e^{-\theta t}}{1-e^{-\theta}},\ t\in(0,1].$$

Examples of Archimedean copulas with strict generators:

*Gumbel-Copula*:

$$C_\theta(u_1,\cdots,u_d) = \exp\left(-\left\{\sum_{i=1}^{d}\left(-\ln(u_i)\right)^\theta\right\}^{1/\theta}\right),\ \mathbf{u}\in(0,1]^d,\ \theta\geq 1$$

The mixing random variable $Z$ here follows a particular *positively stable* distribution with Laplace transform

$$\varphi^{-1}(s) = E\left(e^{-sZ}\right) = e^{-s^{1/\theta}},\ s\geq 0.$$

**Summary:**

| Name | Copula $C_\theta$ | Generator $\varphi_\theta$ | Mixing distribution |
|---|---|---|---|
| Clayton | $\left[\sum_{i=1}^{d} u_i^{-\theta} - d + 1\right]^{-1/\theta}, \ \theta > 0$ | $\frac{1}{\theta}\left(t^{-\theta} - 1\right)$ | $\Gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$ |
| Gumbel | $\exp\left(-\left\{\sum_{i=1}^{d}\left(-\ln(u_i)\right)^{\theta}\right\}^{1/\theta}\right), \ \theta \geq 1$ | $(-\ln t)^{\theta}$ | positively stable |
| Frank | $-\frac{1}{\theta}\ln\left(1-\left(1-e^{-\theta}\right)\prod_{i=1}^{d}\left\{\frac{1-e^{-\theta u_i}}{1-e^{-\theta}}\right\}\right), \theta > 0$ | $-\ln\frac{1-e^{-\theta t}}{1-e^{-\theta}}$ | $\mathcal{LS}\left(e^{-\theta}\right)$ over $\mathbb{N}$ |

**Checkerboard copulas:**

For $d \in \mathbb{N}$ let $\mathbf{U} = (U_1, \cdots, U_d)$ be a random vector whose components $U_i$ follow a discrete uniform distribution over the set $T := \{0, 1, \cdots, m-1\}$ with $m \in \mathbb{N}$ for $i = 1, \cdots, d$. Let further denote

$$p_m(k_1, \cdots, k_d) := P\left(\bigcap_{i=1}^{d} \{U_i = k_i\}\right) \text{ for all } (k_1, \cdots, k_d) \in T^d$$

the joint probabilities of $\mathbf{U}$ (forming a $d$-dimensional contingency table) and

$$I_{k_1, \cdots, k_d} := \bigtimes_{j=1}^{d} \left[\frac{k_j}{m}, \frac{k_j + 1}{m}\right] \text{ for } (k_1, \cdots, k_d) \in T^d$$

giving all possible subcubes of $(0,1]^d$ with edge length $1/m$. The checkerboard copula density $c_{m\square}$ of the checkerboard copula $C_{m\square}$ is then defined by

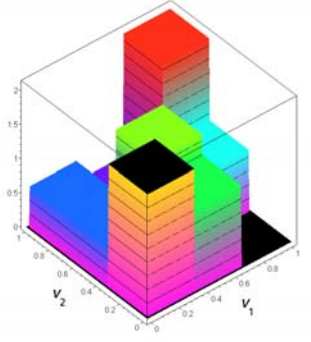
$$c_{m\square} := m^d \sum_{k_1=0}^{m-1} \cdots \sum_{k_d=0}^{m-1} p_m(k_1, \cdots, k_d)\, \mathbb{1}_{I_{k_1, \cdots, k_d}}.$$

Interpretation:

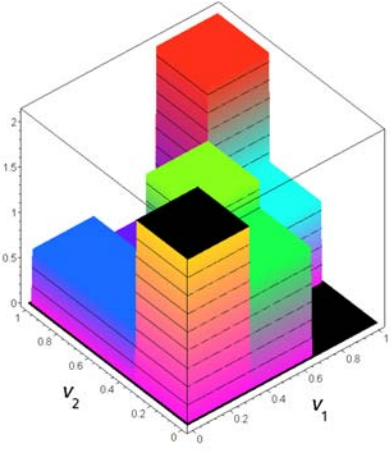
A random vector $\mathbf{V} = (V_1, \cdots, V_d)$ with a copula as cdf has a checkerboard copula iff the conditional distribution of $\mathbf{V}$ given the event $\left\{\mathbf{V} \in I_{k_1,\cdots,k_d}\right\}$ is continuous uniform over $I_{k_1,\cdots,k_d}$ with $p_m(k_1,\cdots,k_d) = P\left(\mathbf{V} \in I_{k_1,\cdots,k_d}\right)$ for all $(k_1,\cdots,k_d) \in T^d$.

Example 7: Assume $d = 2$ and $m = 3$ and consider the contingency table

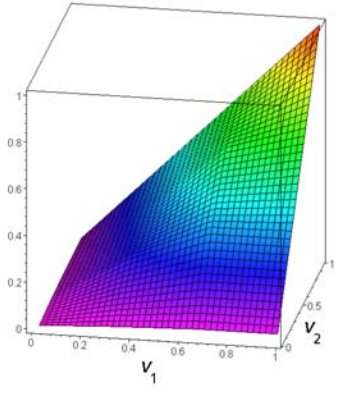
| $P(U_1 = i, U_2 = j)$ | | $i$ | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | $P(U_2 = j)$ |
| $j$ | 0 | 6/30 | 4/30 | 0 | 1/3 |
| | 1 | 2/30 | 5/30 | 3/30 | 1/3 |
| | 2 | 2/30 | 1/30 | 7/30 | 1/3 |
| | $P(U_1 = i)$ | 1/3 | 1/3 | 1/3 | |




density of $\mathbf{V}$




copula density $c_{3\square}(v_1, v_2)$




copula $C_{3\square}(v_1, v_2)$

**Approximation Theorem:** Every copula $C$ in $d$ dimensions can be uniformly approximated by a sequence $\{C_{m\square}\}_{m\in\mathbb{N}}$ of checkerboard copulas. A particular choice of admissible parameters is given by

$$p_m(k_1,\cdots,k_d) = P\left(\mathbf{Z}\in I_{k_1,\cdots,k_d}\right) = P\left(\bigcap_{j=1}^{d}\left\{\frac{k_j}{m} < Z_j \le \frac{k_j+1}{m}\right\}\right) \text{ for all } (k_1,\cdots,k_d)\in T^d$$

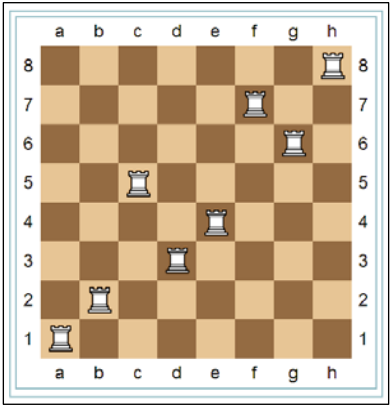
where $\mathbf{Z}=(Z_1,\cdots,Z_d)$ denotes a random vector that has $C$ as joint cdf.

Interpretation: If $\mathbf{V}_m$ is a random vector that has the approximating checkerboard copula $C_{m\square}$ as cdf, then

$$P\left(\mathbf{V}_m\in I_{k_1,\cdots,k_d}\right) = p_m(k_1,\cdots,k_d) = P\left(\mathbf{Z}\in I_{k_1,\cdots,k_d}\right) \text{ for all } (k_1,\cdots,k_d)\in T^d.$$

The difference is that the conditional distribution of $\mathbf{Z}$ given $\left\{\mathbf{Z}\in I_{k_1,\cdots,k_d}\right\}$ is "smoothed" by a local continuous uniform distribution.

**Rook copulas:** Particular checkerboard copulas that distribute probability mass according to the placement of rooks on a checkerboard without mutual threatening. It can in general be constructed in $d$ dimensions as follows: Let



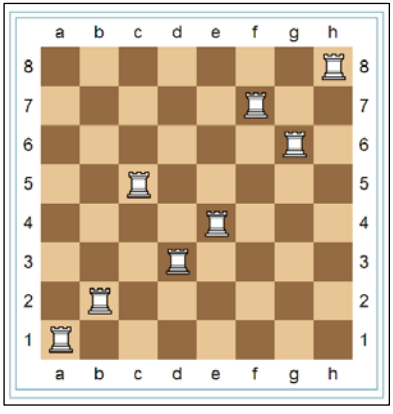

$$M := \begin{bmatrix} \sigma_{01} & \sigma_{02} & \cdots & \sigma_{0,d-1} & \sigma_{0d} \\ \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,d-1} & \sigma_{1d} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{m-2,1} & \sigma_{m-2,2} & \cdots & \sigma_{m-2,d-1} & \sigma_{m-2,d} \\ \sigma_{m-1,1} & \sigma_{m-1,2} & \cdots & \sigma_{m-1,d-1} & \sigma_{m-1,d} \end{bmatrix}$$

denote a matrix of transposed permutations , i.e. each $(\sigma_{0k}, \sigma_{1k}, \cdots, \sigma_{m-1,k})$ is a permutation of the set $T := \{0, 1, \cdots, m-1\}$ for $k = 1, \cdots, d$. A checkerboard copula $C$ is a rook copula iff for the non-zero probabilities there holds

$$p_m(k_1, \cdots, k_d) = P\left(\bigcap_{i=1}^{d} \{U_i = k_i\}\right) = \frac{1}{m} \Leftrightarrow (k_1, \cdots, k_d) = (\sigma_{t1}, \sigma_{t2}, \cdots, \sigma_{t,d}) \text{ for some } t \in T.$$

Example 8: The rook copula corresponding to the picture above is given by the matrix

$$M = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 2 & 4 \\ 3 & 2 \\ 4 & 3 \\ 5 & 6 \\ 6 & 5 \\ 7 & 7 \end{bmatrix}.$$




$(k_1, k_2) = (7,7)$

$(k_1, k_2) = (4,3)$

$(k_1, k_2) = (0,0)$

Rook copulas are of special importance for Monte Carlo simulations on the basis of empirical copulas which will be discussed in a later section.

**Bernstein copulas:** Bernstein-polynomial of degree $m$:

$$B(m,k,z)=\binom{m}{k}z^k(1-z)^{m-k},\ 0\le z\le 1,\ k=0,\cdots,m\in\mathbb{N}$$

For $d\in\mathbb{N}$ let $\mathbf{U}=(U_1,\cdots,U_d)$ again be a random vector whose components $U_i$ follow a discrete uniform distribution over the set $T:=\{0,1,\cdots,m-1\}$ with $m\in\mathbb{N}$ for $i=1,\cdots,d$. Let further denote again

$$p_m(k_1,\cdots,k_d):=P\left(\bigcap_{i=1}^{d}\{U_i=k_i\}\right)\text{ for all }(k_1,\cdots,k_d)\in T^d$$

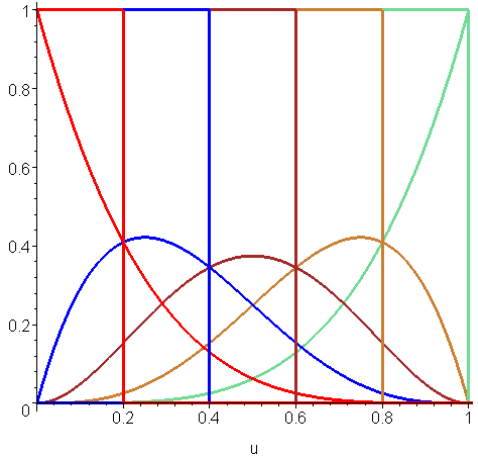
the joint probabilities of $\mathbf{U}$ (forming a $d$-dimensional contingency table). Then

$$c_B(u_1,\cdots,u_d):=m^d\sum_{k_1=0}^{m-1}\cdots\sum_{k_d=0}^{m-1}p(k_1,\cdots,k_d)\prod_{i=1}^{d}B(m-1,k_i,u_i),\ (u_1,\cdots,u_d)\in[0,1]^d$$

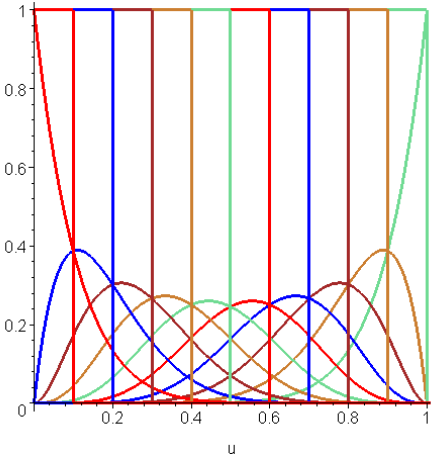
defines the density $c_B$ of a *Bernstein-Copula* $C_B$ induced by $\mathbf{U}$.

**Remark:** A Bernstein copula can be considered as a smoothed version of a checkerboard copula, distributing the mass of the local continuous uniform distribution over the whole unit cube in an appropriate way.

Visualization of the smoothing effect for $d = 1$:




$m = 5$

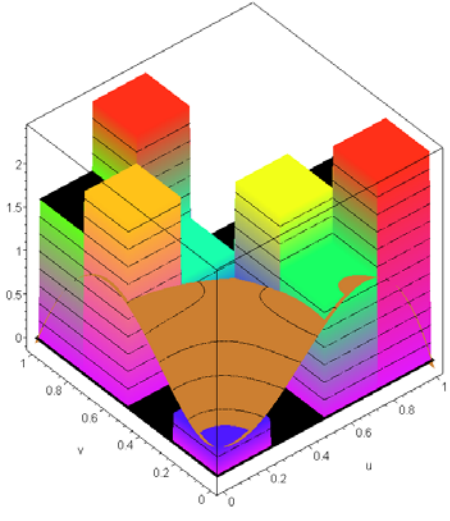


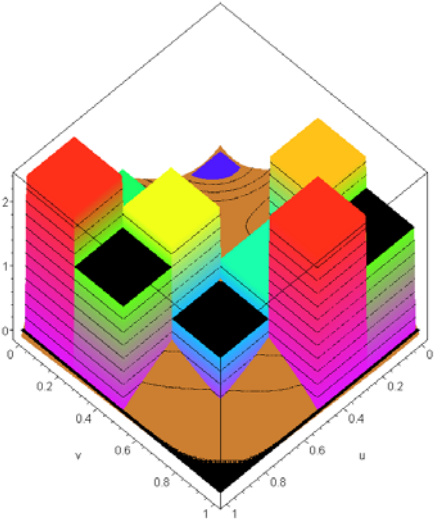

$m = 10$

Example 9: Smoothing effect for $d = 2$ and $m = 4$; the distribution of $\mathbf{U} = (U_1, U_2)$ is given in the following table:

| $P(\mathbf{U} = (i,j))$ | | $i$ | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| $j$ | 0 | 0.02 | 0.00 | 0.08 | 0.15 |
| | 1 | 0.00 | 0.03 | 0.12 | 0.10 |
| | 2 | 0.13 | 0.07 | 0.05 | 0.00 |
| | 3 | 0.10 | 0.15 | 0.00 | 0.00 |

The following pictures show the smoothing effect of the corresponding Bernstein copula in comparison with the raw checkerboard copula:


2
1.5
1
0.5
0
1
0.8
0.6
0.4
0.2
0
v
0
0.2
0.4
0.6
0.8
1
u


2
1
0
0
0.2
0.4
0.6
0.8
1
v
0
0.2
0.4
0.6
0.8
1
u

**Partition of unity copulas:** Generalizations of checkerboard and Bernstein copulas, based on a family of functions $\{\phi(m,k,\bullet)|0\le k\le m-1,\, m\in\mathbb{N}\}$ (so called partition of unity) with the following properties:

- $\int_0^1 \phi(m,k,u)\,du = \frac{1}{m}$ for $k=0,\cdots,m-1$
- $\sum_{k=0}^{m-1}\phi(m,k,\bullet)=1$ for $m\in\mathbb{N}$.

Under the conditions of checkerboard and Bernstein copulas,

$$c_\phi(u_1,\cdots,u_d) := m^d \sum_{k_1=0}^{m-1}\cdots\sum_{k_d=0}^{m-1} P\left(\bigcap_{i=1}^d \{U_i=k_i\}\right)\prod_{i=1}^d \phi(m,k_i,u_i),\ (u_1,\cdots,u_d)\in[0,1]^d$$

defines the density $c_\phi$ of a partition of unity copula $C_\phi$ induced by $\mathbf{U}$.

**Remark 1:** Checkerboard and Bernstein copulas are special cases of a partition of unity copula:

Checkerboard copula: $\phi(m,k,u) = \mathbb{1}_{\left(\frac{k}{m},\frac{k+1}{m}\right]}(u)$

Bernstein copula: $\phi(m,k,u) = \binom{m-1}{k} u^k (1-u)^{m-1-k}$

for $0 \leq u \leq 1$ and $0 \leq k \leq m-1$, $m \in \mathbb{N}$.

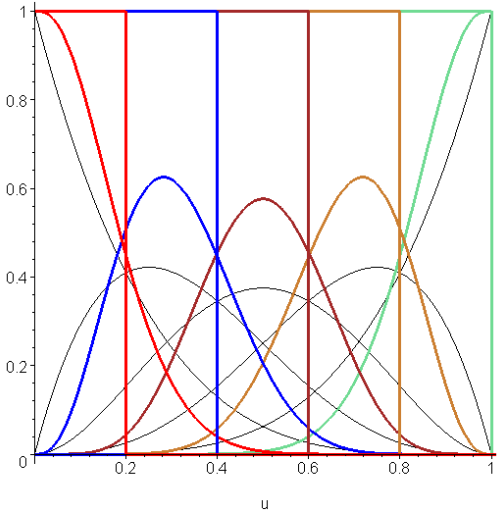
**Remark 2:** Every partition of unity $\{\phi(m,k,\bullet)|0\le k\le m-1,\, m\in\mathbb{N}\}$ generates a new partition of unity $\{\phi_K(m,k,\bullet)|0\le k\le m-1,\, m\in\mathbb{N}\}$ for any fixed $K\in\mathbb{N}$ via

$$\phi_K(m,k,\bullet) := \sum_{j=0}^{K-1}\phi(K\cdot m, K\cdot k+j,\bullet) \quad \text{for } k=0,\cdots,m-1$$

since

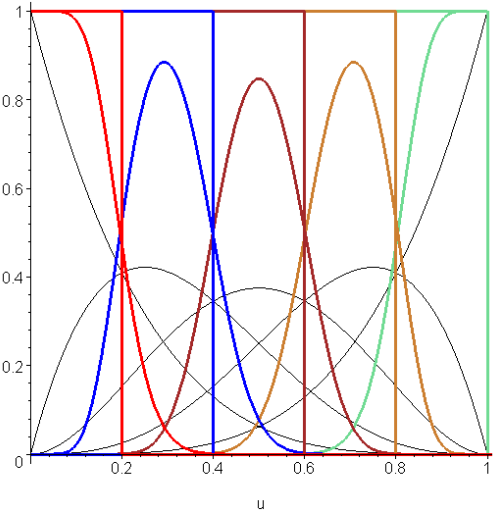
- $$\int_0^1 \phi_K(m,k,u)\,du = \sum_{j=0}^{K-1}\int_0^1 \phi(K\cdot m,K\cdot k+j,u)\,du = \sum_{j=0}^{K-1}\frac{1}{K\cdot m} = \frac{1}{m},\ k=0,\cdots,m-1$$
- $$\sum_{k=0}^{m-1}\phi_K(m,k,\bullet) = \sum_{j=0}^{K-1}\sum_{k=0}^{m-1}\phi(K\cdot m,K\cdot k+j,\bullet) = \sum_{i=0}^{K\cdot m}\phi(K\cdot m,i,\bullet) = 1,\ m\in\mathbb{N}.$$

Visualization of the resulting smoothing effect for Bernstein copulas:




$K = 3$

$m = 5$




$K = 10$

**Other extensions of (known) copulas:**

- For every $d$-dimensional copula $C$ and arbitrary $0 \leq \alpha_k \leq 1,\ k = 1, \cdots, d$

$$C^*(u_1, \cdots, u_d) = \prod_{k=1}^{d} u_k^{(1-\alpha_k)} \cdot C\left(u_1^{\alpha_1}, \cdots, u_d^{\alpha_d}\right),\ 0 \leq u_1, \cdots, u_d \leq 1$$

defines a new copula which is not symmetric if some of the $\alpha_k$ are not identical.

- For every $d$-dimensional copula $C$ and arbitrary $K \in \mathbb{N}$

$$C_{(K)}(u_1, \cdots, u_d) = C^K\left(u_1^{1/K}, \cdots, u_d^{1/K}\right),\ 0 \leq u_1, \cdots, u_d \leq 1$$

defines a new copula. In case that the limit

$$C_{(\infty)}(u_1, \cdots, u_d) = \lim_{K \to \infty} C_{(K)}\left(u_1, \cdots, u_d\right),\ 0 \leq u_1, \cdots, u_d \leq 1$$

exists as a copula, then $C_{(\infty)}$ is called *extreme value copula*.

Proof of the first case:

Let $\mathbf{V} = (V_1, \cdots, V_d)$ be a random vector with copula $C$ as cdf and $\mathbf{U} = (U_1, \cdots, U_d)$ a random vector with the independence copula as cdf, also independent of $\mathbf{V}$. Define

$$W_k := \max\left\{V_k^{1/\alpha_k}, U_k^{1/(1-\alpha_k)}\right\} \quad \text{for } k = 1, \cdots, d.$$

Then $\mathbf{W} = (W_1, \cdots, W_d)$ has the copula $C^*$ since

$$P\left(\bigcap_{k=1}^{d}\{W_k \le u_k\}\right) = P\left(\bigcap_{k=1}^{d}\left(\left\{V_k \le u_k^{\alpha_k}\right\} \cap \left\{U_k \le u_k^{1-\alpha_k}\right\}\right)\right)$$
$$= P\left(\bigcap_{k=1}^{d}\left\{V_k \le u_k^{\alpha_k}\right\}\right) \cdot P\left(\bigcap_{k=1}^{d}\left\{U_k \le u_k^{1-\alpha_k}\right\}\right) = C\left(u_1^{\alpha_1}, \cdots, u_d^{\alpha_d}\right) \cdot \prod_{k=1}^{d} u_k^{(1-\alpha_k)}$$

with $P(W_k \le u_k) = P\left(V_k \le u_k^{\alpha_k}\right) \cdot P\left(U_k \le u_k^{1-\alpha_k}\right) = u_k^{\alpha_k} \cdot u_k^{1-\alpha_k} = u_k, \quad 0 \le u_1, \cdots, u_d \le 1.$

Proof of the second case:

Let $\mathbf{V}_i = (V_{i1}, \cdots, V_{id})$, $i = 1, \cdots, K$ be independent random vectors with copula $C$ as cdf's each. Define

$$W_m := \left(\max\{V_{1m}, \cdots, V_{Km}\}\right)^K \quad \text{for } m = 1, \cdots, d.$$

Then $\mathbf{W} = (W_1, \cdots, W_d)$ has the copula $C^*$ since

$$P\left(\bigcap_{m=1}^{d}\{W_m \le u_m\}\right) = P\left(\bigcap_{m=1}^{d}\bigcap_{i=1}^{K}\{V_{im} \le u_m^{1/K}\}\right) = P\left(\bigcap_{i=1}^{K}\bigcap_{m=1}^{d}\{V_{im} \le u_m^{1/K}\}\right)$$
$$= \prod_{i=1}^{K} C\left(u_1^{1/K}, \cdots, u_d^{1/K}\right) = C^K\left(u_1^{1/K}, \cdots, u_d^{1/K}\right)$$

with $P(W_m \le u_m) = P\left(\bigcap_{i=1}^{K}\{V_{im} \le u_m^{1/K}\}\right) = \left(u_m^{1/K}\right)^K = u_m, \quad 0 \le u_1, \cdots, u_d \le 1.$

**Remark:** The Gumbel copula is a particular extreme value copula since from

$$C(u_1,\cdots,u_n)=\exp\left(-\left\{\sum_{i=1}^{n}\left(-\ln(u_i)\right)^{\theta}\right\}^{1/\theta}\right),\quad \theta\geq 1$$

it follows for every $K\in\mathbb{N}$ that

$$C^K\left(u_1^{1/K},\cdots,u_d^{1/K}\right)=\exp\left(-K\left\{\sum_{i=1}^{d}\left(-\ln\left(u_i^{1/K}\right)\right)^{\theta}\right\}^{1/\theta}\right)=\exp\left(-K\left\{\sum_{i=1}^{d}\left(-\frac{1}{K}\ln(u_i)\right)^{\theta}\right\}^{1/\theta}\right)$$

$$=\exp\left(-K\left\{\frac{1}{K^{\theta}}\sum_{i=1}^{d}\left(-\ln\left(u_i\right)\right)^{\theta}\right\}^{1/\theta}\right)=C(u_1,\cdots,u_d),\quad \theta\geq 1$$

and hence also

$C=C_{(K)}=C_{(\infty)}$ for all $K\in\mathbb{N}$.

**Nesting of copulas** ($\rightarrow$ compatibility problem):

**Question:** If $C_1, C_2$ are bivariate copulas, are then also

$$C_3(u,v,w) = C_2\big(C_1(u,v),w\big) \quad \text{and / or} \quad C_2\big(u,C_1(v,w)\big) \text{ with } 0 \le u,v,w \le 1$$

copulas (likewise for larger dimensions)?

**Answer:** Generally **not**! Counterexample: Choose $C_1 = C_2 = C_*$ (lower Fréchet-Hoeffding bound), then

$$C_3(u,v,w) = C_2\big(C_1(u,v),w\big) = C_*(u,v,w),$$

but this is *not* a copula.

**Nesting of Archimedean copulas:**

If $\varphi_1$ and $\varphi_2$ are two strict generators of Archimedean copulas $C_1$ and $C_2$ whose inverses $\varphi_1^{-1}$ und $\varphi_2^{-1}$ are completely monotone and for which also the composition $-\varphi_2\left(\varphi_1^{-1}\right)$ is completely monotone, i.e.

$$(-1)^{k-1}\frac{d^k}{dt^k}\varphi_2\left(\varphi_1^{-1}(t)\right)\geq 0 \text{ for all } k\in\mathbb{N},$$

then

$$C(u,v,w)=C_2\left(C_1(u,v),w\right)=\varphi_2^{-1}\left(\varphi_2\circ\varphi_1^{-1}\left(\varphi_1(u)+\varphi_1(v)\right)+\varphi_2(w)\right), 0\leq u,v,w\leq 1$$

also is a copula. This construction principle can similarly be extended to higher dimensions (→ hierarchical Archimedean copulas).

Example 10: Nesting of Gumbel copulas: consider

$\varphi_1(t) = (-\ln t)^a$, $\varphi_2(t) = (-\ln t)^b$ with $a \geq b \geq 1$, then

$$\varphi_2\left(\varphi_1^{-1}(t)\right) = \left(-\ln\left(\varphi_1^{-1}(t)\right)\right)^b = \left(-\ln\left(\exp\left(-t^{1/a}\right)\right)\right)^b = t^{b/a} \text{ and}$$

$$(-1)^{k-1}\frac{d^k}{dt^k}\varphi_2\left(\varphi_1^{-1}(t)\right) = (-1)^{k-1}c_k t^{b/a-k} \geq 0 \text{ with } c_k = \prod_{j=0}^{k-1}\left(\frac{b}{a} - j\right).$$

Hierarchical Gumbel copula:

$$C(u,v,w) = C_2(C_1(u,v),w) = \exp\left(-\left(\left\{-\ln(C_1(u,v))\right\}^b + (-\ln w)^b\right)^{1/b}\right)$$

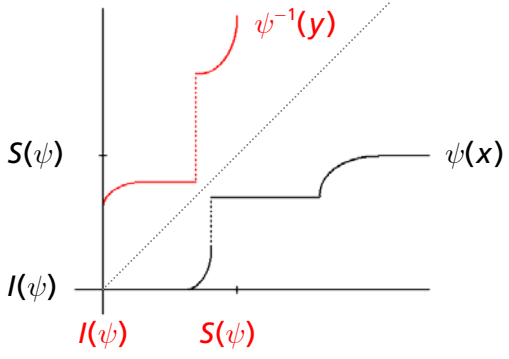
$$= \exp\left(-\left(\left\{(-\ln u)^a + (-\ln v)^a\right\}^{b/a} + (-\ln w)^b\right)^{1/b}\right), \quad 0 \leq u,v,w \leq 1.$$

# III. Order Statistics and Extreme Value Theory

A short review on quantile and related functions:

Let $\psi : \mathbb{R} \to \mathbb{R}$ be a weakly increasing right-continuous function. Denote $I(\psi) := \inf\{\psi(x) \mid x \in \mathbb{R}\}$, $S(\psi) := \sup\{\psi(x) \mid x \in \mathbb{R}\}$. Then the *pseudo-inverse* $\psi^{-1}$ of $\psi$ is defined by

$$\psi^{-1}(y) := \inf\{x \in \mathbb{R} \mid \psi(x) \geq y\}, \quad y \in (I(\psi), S(\psi)).$$



**Properties of a pseudo-inverse:**

- $\psi(x) < y$ iff $x < \psi^{-1}(y)$ for $x \in \mathbb{R}$ and $y \in (I(\psi), S(\psi))$, or, alternatively,
- $\psi(x) \geq y$ iff $x \geq \psi^{-1}(y)$ for $x \in \mathbb{R}$ and $y \in (I(\psi), S(\psi))$.
- $\psi^{-1}$ is weakly increasing over the interval $(I(\psi), S(\psi))$ and left-continuous.
- $\psi\left(\psi^{-1}(y)\right) \geq y$ for all $y \in (I(\psi), S(\psi))$.
- If $\psi$ is continuous in $\psi^{-1}(y)$ for some $y \in (I(\psi), S(\psi))$, then $\psi\left(\psi^{-1}(y)\right) = y$.
- $\psi^{-1}\left(\psi(x)\right) \leq x$ for all $x \in \mathbb{R}$ with $\psi(x) \in (I(\psi), S(\psi))$.
- If $\psi^{-1}$ is continuous in $\psi(x)$ for some $x \in \mathbb{R}$ with $\psi(x) \in (I(\psi), S(\psi))$, then $\psi^{-1}\left(\psi(x)\right) = x$.

Remember that the pseudo-inverse $F^{-1}$ of a univariate cdf $F$ is also called *quantile function*.

**Empirical distribution function:**

Let $X_1, \cdots, X_n$ $(n \in \mathbb{N})$ be independent identically distributed (i.i.d.) random variables with cdf $F$. Then the empirical distribution function $\hat{F}_n$ is defined as

$$\hat{F}_n(x) := \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{(-\infty, x]}(X_k) \quad \text{for all } x \in \mathbb{R}.$$

Interpretation:

The empirical distribution function $\hat{F}_n$ is the cdf of a discrete random probability measure $\xi$ (random Laplace distribution) which gives the mass $\frac{1}{n}$ to each of the random variables $X_1, \cdots, X_n$, i.e.

$$\xi(A) = \frac{\#\left(k \mid X_k \in A\right)}{n} \quad \text{for all Borel sets } A \in \mathcal{B}^1.$$

**Glivenko-Cantelli Theorem:**

The empirical distribution function $\hat{F}_n$ converges uniformly to the true distribution function $F$ for $n \to \infty$, i.e.

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \left| \hat{F}_n(x) - F(x) \right| = 0.$$

Note that the weaker proposition

$$\lim_{n\to\infty} \left| \hat{F}_n(x) - F(x) \right| = 0 \text{ for all } x \in \mathbb{R}$$

is a direct consequence of the Law of Large Numbers (LLN) applied to the i.i.d. sequence $\{Y_n(x)\}_{n\in\mathbb{N}}$ with $Y_n(x) := \mathbb{1}_{(-\infty,x]}(X_n)$ for $n \in \mathbb{N}$, $x \in \mathbb{R}$ with a binomial distribution $P^{Y_n(x)} = B(1, F(x))$ each.

**Order statistics:**

Let $X_1, \cdots, X_n$ be independent identically distributed random variables with cdf $F$. Then the $k$-th order statistic $X_{k:n}$ is defined as a quantile of the empirical distribution function:

$$X_{k:n} := \hat{F}_n^{-1}\left(\frac{k}{n}\right) \text{ for } k = 1, \cdots, n.$$

**Interpretation:**

The $k$-th order statistic corresponds (up to uniqueness) to the $k$-th largest value among $X_1, \cdots, X_n$, i.e.

$$\min\{X_1, \cdots, X_n\} = X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n} = \max\{X_1, \cdots, X_n\}.$$

Note that ties can occur among order statistics with positive probability if the underlying distribution has atoms.

**Marginal distribution of order statistics:**

$$P(X_{k:n} \le x) = \sum_{i=k}^{n} \binom{n}{i} F^i(x)[1-F(x)]^{n-i} \text{ for } x \in \mathbb{R},\ k = 1, \cdots, n.$$

Proof: By the second property of the pseudo-inverse, we have, using the notation above:

$$P(X_{k:n} \le x) = P\left(\hat{F}_n^{-1}\left(\frac{k}{n}\right) \le x\right) = P\left(\frac{k}{n} \le \hat{F}_n(x)\right) = P\left(\sum_{i=1}^{n} Y_i(x) \ge k\right)$$
$$= \sum_{i=k}^{n} \binom{n}{i} F^i(x)[1-F(x)]^{n-i}.$$

**Joint distribution of two order statistics:**

$$P(X_{k:n} \le x,\ X_{r:n} \le y) = \sum_{j=r}^{n} \sum_{i=k}^{j} \binom{n}{j} \binom{j}{i} F^i(x)[F(y)-F(x)]^{j-i}[1-F(y)]^{n-j}$$

for real $x < y$, $1 \le k < r \le n$.

Proof: The random vectors

$\mathbf{Y}_i(x,y) = (Y_{1i}(x,y), Y_{2i}(x,y), Y_{3i}(x,y))$ with

$Y_{1i}(x,y) = \mathbb{1}_{(-\infty,x]}(X_i), Y_{2i}(x,y) = \mathbb{1}_{(x,y]}(X_n), Y_{3i}(x,y) = \mathbb{1}_{[y,\infty)}(X_n)$ for $i = 1,\cdots,n$

are independent and *multinomially distributed*, i.e.

$$P(\mathbf{Y}_i(x,y) = (k_1,k_2,k_3)) = F^{k_1}(x)[F(y)-F(x)]^{k_2}[1-F(y)]^{k_3}$$

for $k_1,k_2,k_3 \in \{0,1\}$ with $k_1+k_2+k_3 = 1$.

**Joint distribution of two order statistics:**

$$P(X_{k:n} \le x,\ X_{r:n} \le y) = \sum_{j=r}^{n} \sum_{i=k}^{j} \binom{n}{j} \binom{j}{i} F^i(x) [F(y) - F(x)]^{j-i} [1 - F(y)]^{n-j}$$

for real $x < y$, $1 \le k < r \le n$.

This implies that $\mathbf{Y}(x,y) := \sum_{i=1}^{n} \mathbf{Y}_i(x,y)$ is also multinomially distributed, with

$$P(\mathbf{Y}(x,y) = (k_1,k_2,k_3)) = \binom{n}{k_1,k_2,k_3} F^{k_1}(x) [F(y) - F(x)]^{k_2} [1 - F(y)]^{k_3}$$

for $k_1,k_2,k_3 \in \mathbb{Z}^+$ with $k_1 + k_2 + k_3 = n$.

**Joint distribution of two order statistics:**

$$P(X_{k:n} \le x,\, X_{r:n} \le y) = \sum_{j=r}^{n} \sum_{i=k}^{j} \binom{n}{j} \binom{j}{i} F^i(x) [F(y) - F(x)]^{j-i} [1 - F(y)]^{n-j}$$

for real $x < y$, $1 \le k < r \le n$.

Now

$$\begin{aligned} P(X_{k:n} \le x,\, X_{r:n} \le y) &= P\left(\frac{k}{n} \le \hat{F}_n(x),\, \frac{r}{n} \le \hat{F}_n(y)\right) \\ &= P(\mathbf{Y}_1(x,y) \ge k,\, \mathbf{Y}_1(x,y) + \mathbf{Y}_2(x,y) \ge r) = \sum_{j=r}^{n} \sum_{i=k}^{j} P(\mathbf{Y}(x,y) = (i, j-i, n-j)) \\ &= \sum_{j=r}^{n} \sum_{i=k}^{j} \binom{n}{i, j-i, n-j} F^i(x) [F(y) - F(x)]^{j-i} [1 - F(y)]^{n-j} \\ &= \sum_{j=r}^{n} \sum_{i=k}^{j} \binom{n}{j} \binom{j}{i} F^i(x) [F(y) - F(x)]^{j-i} [1 - F(y)]^{n-j}. \end{aligned}$$

**Univariate density for order statistics:** In case that $F$ has a density $f$ we have

$$f_{X_{k:n}}(x) = \frac{f(x)}{B(k, n-k+1)} F^{k-1}(x)\left[1-F(x)\right]^{n-k} \quad \text{for } x \in \mathbb{R}.$$

Proof: From $F_{X_{k:n}}(x) = \sum_{i=k}^{n} \binom{n}{i} F^i(x)\left[1-F(x)\right]^{n-i}$ we obtain, by differentiation:

$$\begin{aligned}
f_{X_{k:n}}(x) &= \frac{d}{dx} F_{X_{k:n}}(x) = \sum_{i=k}^{n} \binom{n}{i} \frac{d}{dx}\left\{F^i(x)\left[1-F(x)\right]^{n-i}\right\} \\
&= f(x) \sum_{i=k}^{n} \binom{n}{i} \left\{i F^{i-1}(x)\left[1-F(x)\right]^{n-i} - (n-i) F^i(x)\left[1-F(x)\right]^{n-i-1}\right\} \\
&= f(x)\left\{n \sum_{i=k-1}^{n-1} \binom{n-1}{i} F^i(x)\left[1-F(x)\right]^{n-1-i} - n \sum_{i=k}^{n-1} \binom{n-1}{i} F^i(x)\left[1-F(x)\right]^{n-1-i}\right\} \\
&= n \binom{n-1}{k-1} f(x) F^{k-1}(x)\left[1-F(x)\right]^{n-k} = \frac{f(x)}{B(k, n-k+1)} F^{k-1}(x)\left[1-F(x)\right]^{n-k}.
\end{aligned}$$

**Bivariate density for order statistics:** In case that $F$ has a density $f$ we have

$$f_{(X_{k:n},X_{r:n})}(x,y) = \begin{cases} n(n-1)\binom{n-2}{k-1,r-k-1,n-r} f(x)f(y)F^{k-1}(x)\left[F(y)-F(x)\right]^{r-k-1}\left[1-F(y)\right]^{n-r} & \text{if } x \le y \\ 0 & \text{otherwise} \end{cases}$$

for $x, y \in \mathbb{R}$ and $1 \le k < r \le n$.

The proof is similar to the one for univariate densities by calculating the partial derivatives

$$f_{(X_{k:n},X_{r:n})}(x,y) = \frac{\partial^2}{\partial x\,\partial y}\left\{\sum_{j=r}^{n}\sum_{i=k}^{j}\binom{n}{j}\binom{j}{i}F^{i}(x)\left[F(y)-F(x)\right]^{j-i}\left[1-F(y)\right]^{n-j}\right\}.$$

**Joint density of all order statistics:** In case that $F$ has a density $f$ we have

$$f_{(X_{1:n},\cdots,X_{n:n})}(x_1,\cdots,x_n) = \begin{cases} n!\prod_{k=1}^{n} f(x_k) & \text{if } x_1 \le x_2 \le \cdots \le x_n \\ 0 & \text{otherwise.} \end{cases}$$

Proof: Let $\Sigma$ denote the set of all permutations $\sigma = (\sigma_1,\cdots,\sigma_n)$ of $(1,2,\cdots,n)$ and $B := \{(x_1,\cdots,x_n) \in \mathbb{R}^n \mid x_1 \le x_2 \le \cdots \le x_n\}$. Then, by the i.i.d. property,

$$\begin{aligned} P(X_{1:n} \le x_1,\cdots,X_{n:n} \le x_n) &= \sum_{\sigma\in\Sigma} P\left(\{X_{\sigma_1} \le X_{\sigma_2} \le \cdots \le X_{\sigma_n}\} \cap \{X_{\sigma_1} \le x_1,\cdots,X_{\sigma_n} \le x_n\}\right) \\ &= n!\sum_{\sigma\in\Sigma} P\left(\{X_1 \le X_2 \le \cdots \le X_n\} \cap \{X_1 \le x_1,\cdots,X_1 \le x_n\}\right) \\ &= n!\int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_1} 1_B(u_1,\cdots,u_n)\prod_{k=1}^{n} f(u_k)\,du_1\cdots du_n \end{aligned}$$

from which the statement follows immediately.

**Order statistics under special distributions:**

- If the random variables are continuous uniformly distributed over $[0,1]$, then $X_{k:n}$ is $\mathcal{B}(k, n-k+1)$ Beta-distributed.
- If the random variables are $\mathcal{E}(\lambda)$-exponentially distributed, then the increments $X_{1:n}, X_{2:n} - X_{1:n}, X_{3:n} - X_{2:n}, \cdots, X_{n:n} - X_{n-1:n}$ are also independent and exponentially distributed with

  $P^{X_{k:n} - X_{k-1:n}} = \mathcal{E}\big((n-k+1)\lambda\big)$ for $1 \le k \le n$ (with the convention $X_{0:n} \equiv 0$).
- In general, if $F$ is continuous, the (finite) sequence of order statistics forms a (finite) Markov chain with transition probabilities

$$P\big(X_{k:n} \le x \,|\, X_{k-1:n} = y\big) = F_{X_{k:n}}(x \,|\, X_{k-1:n} = y) = 1 - \left(\frac{1-F(x)}{1-F(y)}\right)^{n-k+1}, \quad x > y,\ 2 \le k \le n.$$

(This property does not hold in the discrete case.)

**Remark:** Order statistics can similarly be defined also for arbitrary random variables $X_1, \cdots, X_n$ $(n \in \mathbb{N})$, with the same formal definition of $\hat{F}_n$. In general, however, neither the Glivenko-Cantelli Theorem nor the distribution formulas remain valid.

Note that a (general) order statistic $X_{k:n}$ can usually not be identified with some element $X_{R_k}$ of the sequence with a non-random index $R_k$ for $k = 1, \cdots, n$.

In the i.i.d. case, the order statistics are almost surely pairwise different. Here a **rank vector** $R(\mathbf{X}) = (R(X_1), \cdots, R(X_n))$ can be properly defined for $\mathbf{X} = (X_1, \cdots, X_n)$ a.s. via the relation $R(X_i) = k$ iff $X_i = X_{k:n}$, i.e. the rank vector indicates the position of the components of $\mathbf{X}$ w.r.t. a sequential ordering.

More generally, a rank vector $r(\mathbf{x}) = (r(x_1), \cdots, r(x_n))$ can be defined uniquely for every vector $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ in the following way:

**Proposition:** For every $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ there exists a unique permutation $\sigma = (\sigma_1, \cdots, \sigma_n)$ of $(1, \cdots, n)$ with the following properties:

- $x_{\sigma_1} \leq x_{\sigma_2} \leq \cdots \leq x_{\sigma_n}$
- $\sigma_k < \sigma_{k+1}$ if $x_{\sigma_k} = x_{\sigma_{k+1}}$ for $k = 1, \cdots, n-1$.

Under these conditions, the rank vector $r(\mathbf{x}) = (r(x_1), \cdots, r(x_n))$ can be uniquely defined as the inverse permutation $r(\mathbf{x}) := \sigma^{-1}$.

Example 11: Let $x = (2,1,4,3,4,4,1,2,1,3)$. In tabular form, we obtain

| index | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| ordered data | $x_2$ | $x_7$ | $x_9$ | $x_1$ | $x_8$ | $x_4$ | $x_{10}$ | $x_3$ | $x_5$ | $x_6$ |
| | 1 | 1 | 1 | 2 | 2 | 3 | 3 | 4 | 4 | 4 |
| $\sigma$ | 2 | 7 | 9 | 1 | 8 | 4 | 10 | 3 | 5 | 6 |
| $r(\mathbf{x}) = \sigma^{-1}$ | 4 | 1 | 8 | 6 | 9 | 10 | 2 | 5 | 3 | 7 |

For example, we get $r(x_1) = 4$, i.e. the first data value $x_1 = 2$ has rank 4 (it is the 4th largest in the ordered sequence); $r(x_5) = 9$, i.e. the fifth data value $x_5 = 4$ has rank 9 (it is the 9th largest in the ordered sequence, behind the 8th largest data value $x_3 = 4$ because the index of 5 is larger than the index of 3).

Now if $\mathbf{X} = (X_1, \cdots, X_n)$ is any random vector then the (random) rank vector $R(\mathbf{X}) = (R(X_1), \cdots, R(X_n))$ can be defined pointwise as

$R(\mathbf{X})(\omega) = (R(X_1(\omega)), \cdots, R(X_n(\omega)))$ for all $\omega \in \Omega$

(the basic set on which all random variables are defined).

**Theorem:** If $\mathbf{X} = (X_1, \cdots, X_n)$ is a random vector with independent and identically distributed components then the rank vector $R(\mathbf{X}) = (R(X_1), \cdots, R(X_n))$ has a Laplace distribution over the set $\Sigma$ of all permutations of $(1, \cdots, n)$, $n \in \mathbb{N}$.

**Corollary:** In the i.i.d. case, the (individual) rank $R(X_k)$ has a Laplace distribution over the set $\{1, 2, \cdots, n\}$ for every $k \in \{1, 2, \cdots, n\}$.

**Max-stability and domain of attraction:**

A non-degenerate cdf $G$ is called max-stable, if there exist constants $A_n > 0$, $B_n \in \mathbb{R}$ such that

$$G^n\left(\frac{x}{A_n} + B_n\right) = G(x) \text{ for all } x \in \mathbb{R},\ n \in \mathbb{N}.$$

A probability distribution is called max-stable, if the corresponding cdf is max-stable.

If $F$ and $G$ are non-degenerate cdf's and if for suitable constants $A_n > 0$, $B_n \in \mathbb{R}$ there holds

$$\lim_{n\to\infty} F^n\left(\frac{x}{A_n} + B_n\right) = G(x) \text{ for all continuity points } x \text{ of } G \text{ (in symbols: } x \in \mathcal{C}(G)),$$

then we say that $F$ is in the *domain of attraction of* $G$, in symbols: $F \in \mathcal{D}(G)$.

**Skorokhod's Theorem:** Let $F_n$ and $G$ be cdf's with

$$\lim_{n\to\infty} F_n(x) = G(x) \text{ for all } x \in \mathcal{C}(G).$$

Then there exist random variables $X_n$ and $Y$ such that

$$F_{X_n} = F_n, \; F_Y = G \text{ and } Y = \lim_{n\to\infty} X_n \text{ a.s.}$$

Sketch of proof: Let $U$ be continuous uniformly distributed over $[0,1]$. Then

$X_n := F_n^{-1}(U)$ and $Y := G^{-1}(U)$ are appropriate choices because

$$G^{-1}(u) = \lim_{n\to\infty} F_n^{-1}(u) \text{ for } u \in (0,1)$$ almost everywhere.

**Corollary:** If - under the conditions of Skorokhod's Theorem - there are $\alpha_n, \alpha > 0$ and $\beta_n, \beta \in \mathbb{R}$ with

$$\alpha = \lim_{n\to\infty} \alpha_n \text{ and } \beta = \lim_{n\to\infty} \beta_n,$$

then also

$$\lim_{n\to\infty} F_n(\alpha_n x + \beta_n) = G(\alpha x + \beta) \text{ for all } x \in \mathbb{R} \text{ with } \alpha x + \beta \in \mathcal{C}(G).$$

Proof: Take $X_n$ and $Y$ as in Skorokhod's Theorem. Then also

$$\lim_{n\to\infty} \frac{X_n - \beta_n}{\alpha_n} = \frac{Y - \beta}{\alpha} \text{ a.s.}$$ from which we obtain the weak convergence result

$$\lim_{n\to\infty} F_n(\alpha_n x + \beta_n) = \lim_{n\to\infty} P(X_n \le \alpha_n x + \beta_n) = P(Y \le \alpha x + \beta) = G(\alpha x + \beta),\ \alpha x + \beta \in \mathcal{C}(G).$$

**Chintchin's Theorem:** Let $F_n$ and $G$ be cdf's, $G$ non-degenerate, and assume that there are $\alpha_n > 0$ and $\beta_n \in \mathbb{R}$ with

$$\lim_{n\to\infty} F_n(\alpha_n x + \beta_n) = G(x) \text{ for all } x \in \mathcal{C}(G).$$

Then necessary and sufficient conditions for the existence of a non-degenerate cdf $G^*$ and $\alpha_n^* > 0$ and $\beta_n^* \in \mathbb{R}$ with

$$\lim_{n\to\infty} F_n(\alpha_n^* x + \beta_n^*) = G^*(x) \text{ for all } x \in \mathcal{C}(G^*)$$

are: $\lim_{n\to\infty} \frac{\alpha_n^*}{\alpha_n} = \alpha$ and $\lim_{n\to\infty} \frac{\beta_n^* - \beta_n}{\alpha_n} = \beta$ for suitable $\alpha > 0$ and $\beta \in \mathbb{R}$.

In this case, we necessarily have $G^*(x) = G(\alpha x + \beta)$ for all $x \in \mathbb{R}$. (We say that $G$ and $G^*$ are *of the same type*.)

**Characterization Theorem:** Let $G$ be a non-degenerate cdf. Then there holds:

- $G$ is max-stable iff there exists non-degenerate cdf's $F_n$ and $A_n > 0, B_n \in \mathbb{R}$ with $n \in \mathbb{N}$ such that

$$\lim_{n\to\infty} F_n\left(\frac{x}{A_{n\cdot k}} + B_{n\cdot k}\right) = G^{1/k}(x) \text{ for all } x \in \mathcal{C}(G) \text{ and all } k \in \mathbb{N}.$$

- $\mathcal{D}(G) \neq \varnothing$ iff $G$ is max-stable. In this case, $G \in \mathcal{D}(G)$, in particular.

**Extremal Types Theorem:** Any (non-degenerate) max-stable cdf $G$ is necessarily of one of the following three types (up to positive-linear transformations):

$G_1(x) = e^{-e^{-x}}$, $x \in \mathbb{R}$ [Gumbel type]

$G_{2,\alpha}(x) = e^{-x^{-\alpha}}$, $x > 0$ $(\alpha > 0)$ [Fréchet type]

$G_{3,\alpha}(x) = e^{-(-x)^{\alpha}}$, $x < 0$ $(\alpha > 0)$ [Weibull type max]

**Corollary.** Let $G$ be a non-degenerate cdf and $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables. If there exists $A_n > 0$, $B_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} P\big(A_n \cdot (X_{n:n} - B_n) \le x\big) = G(x) \text{ for all } x \in \mathcal{C}(G),$$

Then $G$ is max-stable, i.e. the normalized partial maxima converge weakly to one of the three types of limit distributions from the Extremal Types Theorem.

Proof: Clear since $P(X_{n:n} \le x) = F_{X_{n:n}}(x) = F^n(x)$ for all $x \in \mathbb{R}$.

**Remarks.** The limit distributions in the Extremal Type Theorem are usually called *extreme value distributions*.

It is possible to formulate a corresponding limit theorem also for *normalized partial minima* of an i.i.d. sequence of random variables, e.g. by changing the sign.

**Characterization Theorem of domains of attraction (I):**

For any cdf $F$, let $x_{\sup}(F) := \sup\{x \in \mathbb{R} | F(x) < 1\}$ and $\gamma_n(F) := F^{-1}\left(1 - \frac{1}{n}\right)$, $n \in \mathbb{N}$.

Necessary and sufficient conditions for a cdf $F$ to be a member of a domain of attraction of an extreme value cdf are:

1. for $F \in \mathcal{D}(G_1)$: There exists a positive measurable function $g$ on $\mathbb{R}$ such that

$$\lim_{t \uparrow x_{\sup}(F)} \frac{1 - F(t + x\, g(t))}{1 - F(t)} = e^{-x} \text{ for all } x \in \mathbb{R}.$$

A possible choice for $g$ is $g(t) = \int_t^{x_{\sup}(F)} \frac{1 - F(u)}{1 - F(t)} du$ for $t < x_{\sup}(F)$.

A possible choice for the normalizing constants is

$$A_n = \frac{1}{g(\gamma_n(F))},\ B_n = \gamma_n(F) \text{ for } n \in \mathbb{N}.$$

**Characterization Theorem of domains of attraction (I):**

For any cdf $F$, let $x_{\sup}(F) := \sup\{x \in \mathbb{R} | F(x) < 1\}$ and $\gamma_n(F) := F^{-1}\left(1 - \frac{1}{n}\right)$, $n \in \mathbb{N}$.

Necessary and sufficient conditions for a cdf $F$ to be a member of a domain of attraction of an extreme value cdf are:

2. for $F \in \mathcal{D}(G_{2,\alpha})$: $x_{\sup}(F) = \infty$ and

$$\lim_{t \uparrow \infty} \frac{1 - F(t \cdot x)}{1 - F(t)} = x^{-\alpha} \text{ for all } x > 0.$$

A possible choice for the normalizing constants is

$$A_n = \frac{1}{\gamma_n(F)}, \; B_n = 0 \text{ for } n \in \mathbb{N}.$$

**Characterization Theorem of domains of attraction (I):**

For any cdf $F$, let $x_{\sup}(F) := \sup\{x \in \mathbb{R} | F(x) < 1\}$ and $\gamma_n(F) := F^{-1}\left(1 - \frac{1}{n}\right)$, $n \in \mathbb{N}$.

Necessary and sufficient conditions for a cdf $F$ to be a member of a domain of attraction of an extreme value cdf are:

3. for $F \in \mathcal{D}(G_{3,\alpha})$: $x_{\sup}(F) < \infty$ and

$$\lim_{h \downarrow 0} \frac{1 - F\left(x_{\sup}(F) - x \cdot h\right)}{1 - F\left(x_{\sup}(F) - h\right)} = x^{\alpha} \text{ for all } x > 0.$$

A possible choice for the normalizing constants is

$$A_n = \frac{1}{x_{\sup}(F) - \gamma_n(F)}, \; B_n = x_{\sup}(F) \text{ for } n \in \mathbb{N}.$$

**Remark:**

The sequence $\{\gamma_n(F)\}_{n\in\mathbb{N}}$ from above fulfils in particular the relationship

$$\lim_{n\to\infty} n\cdot\left(1-F\left(\gamma_n(F)\right)\right)=1.$$

More generally, any sequence $\{\gamma_n(F)\}_{n\in\mathbb{N}}$ with this property can be used to choose appropriate normalizing constants $A_n$ and $B_n$ as above.

**Characterization Theorem of domains of attraction (II):**

Sufficient conditions for a cdf $F$ with density $f$ to be a member of a domain of attraction of an extreme value cdf are:

1. for $F \in \mathcal{D}(G_1)$: $f'(x) < 0$ for sufficiently large $x < x_{\sup}(F)$, $f\left(x_{\sup}(F)\right) = 0$ and

$$-\lim_{x \uparrow x_{\sup}(F)} \frac{f'(x)(1-F(x))}{f^2(x)} = \lim_{x \uparrow x_{\sup}(F)} \left(\frac{1}{f}\right)'(x)(1-F(x)) = 1 \text{ for all } x \in \mathbb{R}$$

2. for $F \in \mathcal{D}(G_{2,\alpha})$: $f(x) > 0$ for sufficiently large $x$ and

$$\lim_{x \uparrow \infty} \frac{x \cdot f(x)}{1-F(x)} = \alpha$$

3. for $F \in \mathcal{D}(G_{3,\alpha})$: $f(x) > 0$ for sufficiently large $x < x_{\sup}(F)$ and

$$\lim_{x \uparrow x_{\sup}(F)} \frac{\left(x_{\sup}(F) - x\right) f(x)}{1-F(x)} = \alpha.$$

**Tail equivalence:**

Two cdf's $F$ and $G$ are tail equivalent, in symbols: $F \overset{tail}{\sim} G$, if $x_{\sup}(F) = x_{\sup}(G)$ and there holds

$$\lim_{x \uparrow x_{\sup}(F)} \frac{1-F(x)}{1-G(x)} = \beta \text{ for some } \beta \in (0, \infty).$$

Two distributions are called tail equivalent (with the same symbolism), if their corresponding cdf's are tail equivalent.

**Corollary:**

- If $F_1 \in \mathcal{D}(G)$ for an extreme value cdf $G$ and $F_1 \overset{tail}{\sim} F_2$, then also $F_2 \in \mathcal{D}(G)$.
- If $G$ is an extreme value cdf and $F \overset{tail}{\sim} G$, then $F \in \mathcal{D}(G)$.

This means that the domain of attraction $\mathcal{D}(G)$ of an extreme value cdf $G$ consists exactly of all cdf's $F$ that are tail equivalent to $G$ (cf. the Characterization Theorem of domains of attraction).

Example 12: The following distributions are tail equivalent (with $\beta = 1$):

- $G_1 \overset{tail}{\sim} \mathcal{E}(1)$, i.e. the Gumbel distribution has an exponential tail
- $G_{2,\alpha} \overset{tail}{\sim} \mathcal{P}a(\alpha)$ for $\alpha > 0$, i.e. the Fréchet distribution has a Pareto tail
- $G_{3,\alpha} \overset{tail}{\sim} P^{(X-1)}$ for $P^X = \mathcal{B}(1,\alpha)$ and $\alpha > 0$, i.e. the Weibull type max distribution has a shifted Beta tail.

Sketch of proof:

- $$\lim_{x\uparrow\infty} \frac{1-e^{-e^{-x}}}{e^{-x}} = \lim_{h\downarrow 0} \frac{1-e^{-h}}{h} = -\frac{d}{dz} e^{-z}\Big|_{z=0} = 1$$
- $$\lim_{x\uparrow\infty} \frac{1-e^{-x^{-\alpha}}}{\frac{1}{(1+x)^{\alpha}}} = \lim_{x\uparrow\infty} \frac{1-e^{-x^{-\alpha}}}{x^{-\alpha}} = \lim_{h\downarrow 0} \frac{1-e^{-h}}{h} = -\frac{d}{dz} e^{-z}\Big|_{z=0} = 1$$

Example 12: The following distributions are tail equivalent (with $\beta = 1$):

- $G_1 \overset{tail}{\sim} \mathcal{E}(1)$, i.e. the Gumbel distribution has an exponential tail
- $G_{2,\alpha} \overset{tail}{\sim} \mathcal{P}a(\alpha)$ for $\alpha > 0$, i.e. the Fréchet distribution has a Pareto tail
- $G_{3,\alpha} \overset{tail}{\sim} P^{(X-1)}$ for $P^X = \mathcal{B}(1,\alpha)$ and $\alpha > 0$, i.e. the Weibull type max distribution has a shifted Beta tail.

Sketch of proof:

- For $P^X = \mathcal{B}(1,\alpha)$ we have $F_X(x) = 1-(1-x)^\alpha$ for $0 \le x \le 1$, hence $F_{X-1}(x) = F_X(x+1) = 1-(-x)^\alpha$ for $-1 \le x \le 0$, which implies

$$\lim_{x\uparrow 0} \frac{1-e^{-(-x)^\alpha}}{1-F_{X-1}(x)} = \lim_{x\uparrow 0} \frac{1-e^{-(-x)^\alpha}}{(-x)^\alpha} = \lim_{h\downarrow 0} \frac{1-e^{-h}}{h} = -\frac{d}{dz} e^{-z}\Big|_{z=0} = 1.$$

Example 13: Let, as usual, $\Phi$ denote the cdf of the standard normal distribution $\mathcal{N}(0,1)$ and $\varphi$ its density. Then $\Phi \in \mathcal{D}(G_1)$. A possible choice of normalizing constants is

$$A_n = \sqrt{2\ln n} \text{ and } B_n = \sqrt{2\ln n} - \frac{1}{2} \cdot \frac{\ln(\ln n) + \ln(4\pi)}{\sqrt{2\ln n}} \text{ for } n \geq 2.$$

Sketch of proof: We have

$$\left(\frac{1}{\varphi}\right)'(x) = \sqrt{2\pi}\, x \exp\left(\frac{x^2}{2}\right) = \frac{x}{\varphi(x)} \text{ for } x \in \mathbb{R}, \text{ hence, by L'Hospital's rule,}$$

$$\lim_{x\uparrow\infty} \left(\frac{1}{\varphi}\right)'(x)(1-\Phi(x)) = \lim_{x\uparrow\infty} \frac{x \cdot (1-\Phi(x))}{\varphi(x)} = \lim_{x\uparrow\infty} \frac{(1-\Phi(x)) - x \cdot \varphi(x)}{\varphi'(x)}$$

$$= -\lim_{x\uparrow\infty} \frac{x \cdot \varphi(x)}{\varphi'(x)} = \lim_{x\uparrow\infty} \frac{x \cdot \varphi(x)}{x \cdot \varphi(x)} = 1.$$

Example 13: Let, as usual, $\Phi$ denote the cdf of the standard normal distribution $\mathcal{N}(0,1)$ and $\varphi$ its density. Then $\Phi \in \mathcal{D}(G_1)$. A possible choice of normalizing constants is

$$A_n = \sqrt{2\ln n} \text{ and } B_n = \sqrt{2\ln n} - \frac{1}{2} \cdot \frac{\ln(\ln n) + \ln(4\pi)}{\sqrt{2\ln n}} \text{ for } n \geq 2.$$

Sketch of proof: Alternatively, with $g(t) = \frac{1}{t}$, $t > 0$, we obtain, by L'Hospital's rule,

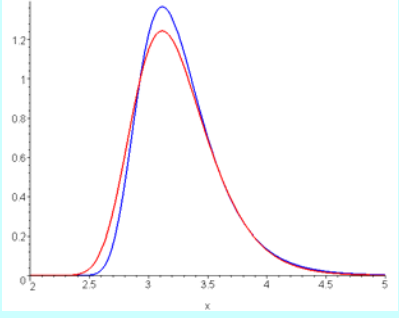
$$\lim_{t \uparrow \infty} \frac{1 - F(t + x\, g(t))}{1 - F(t)} = \lim_{t \uparrow \infty} \frac{1 - \Phi\left(t + \frac{x}{t}\right)}{1 - \Phi(t)} = \lim_{t \uparrow \infty} \frac{\varphi\left(t + \frac{x}{t}\right)}{\varphi(t)} \cdot \left(1 - \frac{x}{t^2}\right)$$

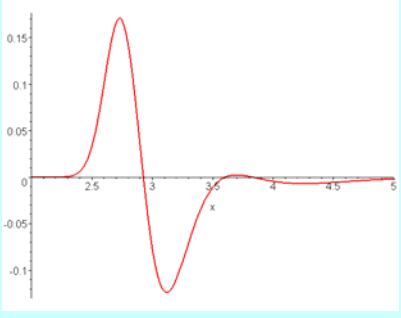
$$= \lim_{t \uparrow \infty} \exp\left(-\frac{1}{2}\left[\left(t + \frac{x}{t}\right)^2 - t^2\right]\right) = e^{-x}$$

for all $x \in \mathbb{R}$.

Example 13: Let, as usual, $\Phi$ denote the cdf of the standard normal distribution $\mathcal{N}(0,1)$ and $\varphi$ its density. Then $\Phi \in \mathcal{D}(G_1)$. A possible choice of normalizing constants is

$$A_n = \sqrt{2\ln n} \text{ and } B_n = \sqrt{2\ln n} - \frac{1}{2} \cdot \frac{\ln(\ln n) + \ln(4\pi)}{\sqrt{2\ln n}} \text{ for } n \geq 2.$$




Left: plot of the density of $X_{n:n}$ for $n = 1000$ [red] and the scaled Gumbel density $A_n \cdot g_1(A_n \cdot (\bullet - B_n))$ [blue]; right: plot of deviation of densities

Example 14: Let $F$ be the cdf of a standard $\mathcal{LN}(0,1)$-lognormal distribution. Then $F \in \mathcal{D}(G_1)$. A possible choice of normalizing constants is

$\tilde{A}_n = A_n \cdot \exp(-B_n)$ and $\tilde{B}_n = \exp(B_n)$ for $n \geq 2$ with the normalizing constants $A_n$ and $B_n$ for the standard normal distribution.

Sketch of Proof: We have

$F(x) = \Phi(\ln x)$, $f(x) = \dfrac{\varphi(\ln x)}{x}$, thus $\left(\dfrac{1}{f}\right)'(x) = \dfrac{\varphi(\ln x) - \varphi'(\ln x)}{\varphi^2(\ln x)}$ for $x > 0$, hence, by L'Hospital's rule,

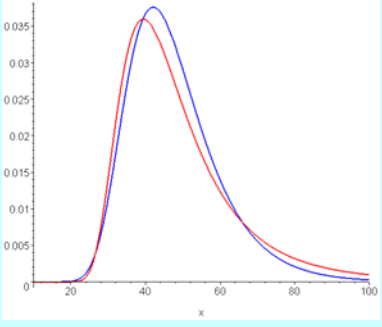
$$\lim_{x \uparrow \infty} \left(\frac{1}{f}\right)'(x)(1-F(x)) = \lim_{x \uparrow \infty} \frac{\varphi(\ln x) - \varphi'(\ln x)}{\varphi^2(\ln x)}(1-\Phi(\ln x)) = \lim_{z \uparrow \infty} \frac{\varphi(z) - \varphi'(z)}{\varphi^2(z)}(1-\Phi(z))$$

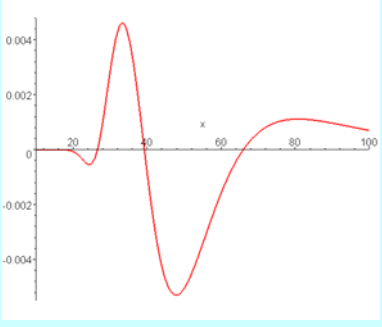
$$= \lim_{z \uparrow \infty} \frac{\varphi'(z)}{\varphi^2(z)}(1-\Phi(z)) = \lim_{z \uparrow \infty} \frac{z \cdot (1-\Phi(z))}{\varphi(z)} = 1$$

as before.

Example 14: Let $F$ be the cdf of a standard $\mathcal{LN}(0,1)$-lognormal distribution. Then $F \in \mathcal{D}(G_1)$. A possible choice of normalizing constants is

$\tilde{A}_n = A_n \cdot \exp(-B_n)$ and $\tilde{B}_n = \exp(B_n)$ for $n \geq 2$ with the normalizing constants $A_n$ and $B_n$ for the standard normal distribution.

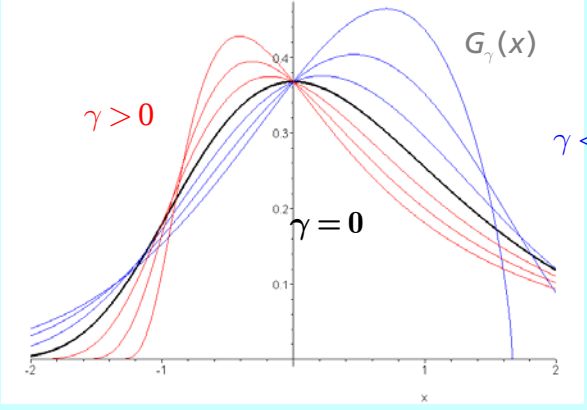



Left: plot of the density of $X_{n:n}$ for $n = 10000$ [red] and the scaled Gumbel density $\tilde{A}_n \cdot g_1\left(\tilde{A}_n \cdot (\bullet - \tilde{B}_n)\right)$ [blue]; right: plot of deviation of densities

**One-parameter representation of extreme value cdf's:**

$$G_\gamma(x) = \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right) & \text{for } 1+\gamma x > 0 \text{ and } \gamma \neq 0 \\ \lim_{\delta \to 0} \exp\left(-(1+\delta x)^{-1/\delta}\right) = e^{-e^{-x}} & \text{for } x \in \mathbb{R} \text{ and } \gamma = 0 \end{cases} \qquad (\gamma \in \mathbb{R})$$




Plot of densities of $G_\gamma(x)$ for various values of $\gamma$

**Smirnov's Theorem I:** Let $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables with a cdf $F$ that belongs to the domain of attraction of an extreme value cdf $G$ with normalizing constants $A_n > 0$, $B_n \in \mathbb{R}$, i.e.

$$\lim_{n\to\infty} P\left(A_n \cdot (X_{n:n} - B_n) \le x\right) = \lim_{n\to\infty} F^n\left(\frac{x}{A_n} + B_n\right) = G(x) \text{ for all } x \in \mathcal{C}(G).$$

Then for every fixed $k \in \{1,\cdots,n\}$, it holds

$$\lim_{n\to\infty} P\left(A_n \cdot (X_{n-k:n} - B_n) \le x\right) = \lim_{n\to\infty} \sum_{i=n-k}^{n} \binom{n}{i} F^i\left(\frac{x}{A_n} + B_n\right)\left[1 - F\left(\frac{x}{A_n} + B_n\right)\right]^{n-i}$$

$$= G(x;k) := G(x)\sum_{j=0}^{k} \frac{\left[-\ln G(x)\right]^j}{j!}$$

for all $x \in \mathcal{C}(G)$.

**Smirnov's Theorem II:** Let $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables and let $\{k_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers with $\lim_{n\to\infty}\sqrt{n}\cdot\left(\frac{k_n}{n}-\beta\right)=0$ for some $\beta>0$. If there exist normalizing constants $A_n(\beta)>0$, $B_n(\beta)\in\mathbb{R}$ such that

$$\lim_{n\to\infty} P\left(A_n(\beta)\cdot\left(X_{n-k_n:n}-B_n(\beta)\right)\le x\right)=G^*(x) \quad \text{for all } x\in\mathcal{C}(G^*),$$

then $G^*$ corresponds exactly to one of the following four (!) types (with $\Phi$ denoting the cdf of the standard normal distribution, as usual):

$$G_{1,\alpha,c}(x)=\begin{cases}0, & \text{if } x<0\\ \Phi\left(c\,x^{\alpha}\right) & \text{if } x\ge 0\end{cases} \qquad G_{2,\alpha,c}(x)=\begin{cases}\Phi\left(-c\,(-x)^{\alpha}\right), & \text{if } x<0\\ 1, & \text{if } x\ge 0\end{cases}$$

$$G_{3,\alpha,c,d}(x)=\begin{cases}\Phi\left(-c\,(-x)^{\alpha}\right), & \text{if } x<0\\ \Phi\left(c\,x^{\alpha}\right) & \text{if } x\ge 0\end{cases} \qquad G_4(x)=\begin{cases}0, & \text{if } x<-1\\ 1/2, & \text{if } -1\le x\le 1\\ 1, & \text{if } x>1\end{cases} \qquad (\alpha,c,d>0)$$

**Remark:** For a calculation of the normalizing constants $A_n(\beta) > 0$, $B_n(\beta) \in \mathbb{R}$ in Smirnov's Theorem II we can potentially make use of the following result:

Suppose that with $p_n(x) := 1 - F\left(\frac{x}{A_n(\beta)} + B_n(\beta)\right)$, $n \in \mathbb{N}, x \in \mathbb{R}$ there holds

$\lim_{n\to\infty} p_n(x) = 0$ and $\lim_{n\to\infty} \sqrt{n}\,\frac{\beta - p_n(x)}{\sqrt{\beta(1-\beta)}} = \tau(x)$ for $x \in \mathbb{R}$, then we have

$$\lim_{n\to\infty} P\left(A_n(\beta)\cdot\left(X_{n-k_n:n} - B_n(\beta)\right) \le x\right) = \Phi(\tau(x)) \text{ for all } x \in \mathbb{R}.$$

If $F$ has a density $f$ with $f\left(F^{-1}(1-\beta)\right) > 0$, then $\tau(x) = x$, and a possible choice for the normalizing constants is

$$A_n(\beta) = \sqrt{n}\,\frac{f\left(F^{-1}(1-\beta)\right)}{\sqrt{\beta(1-\beta)}} \text{ and } B_n(\beta) = -F^{-1}(1-\beta) \text{ for } 0 < \beta < 1 \text{ and } n \in \mathbb{N}.$$

Example 15: Let $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables with an $\mathcal{E}(1)$-exponential distribution. Then with

$k_n := \lfloor \beta \cdot n \rfloor$, $A_n(\beta) := \sqrt{\frac{n\beta}{1-\beta}}$, $B_n(\beta) := -\ln \beta$ for $\beta > 0$, $n \in \mathbb{N}$ we obtain

$$\lim_{n\to\infty} P\left(A_n(\beta)\cdot\left(X_{n-k_n:n} - B_n(\beta)\right) \le x\right) = \Phi(x) \quad \text{for all } x \in \mathbb{R}.$$

Proof: We have $\left|\frac{k_n}{n} - \beta\right| \le \frac{1}{n}$ and hence $\lim_{n\to\infty} \sqrt{n}\cdot\left(\frac{k_n}{n} - \beta\right) = 0$. Further,

$$p_n(x) = \exp\left(-\frac{x}{A_n(\beta)} - B_n(\beta)\right) = \beta \exp\left(-x\sqrt{\frac{1-\beta}{n\beta}}\right) \text{ with}$$

$$\lim_{n\to\infty} p_n(x) = 0 \text{ and } \lim_{n\to\infty} \sqrt{n}\,\frac{\beta - p_n(x)}{\sqrt{\beta(1-\beta)}} = \sqrt{\frac{\beta}{1-\beta}} \lim_{n\to\infty} \sqrt{n}\left(1 - \exp\left(-x\sqrt{\frac{1-\beta}{n\beta}}\right)\right) = x.$$

Example 15: Let $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables with an $\mathcal{E}(1)$-exponential distribution. Then with

$k_n := \lfloor \beta \cdot n \rfloor$, $A_n(\beta) := \sqrt{\frac{n\beta}{1-\beta}}$, $B_n(\beta) := -\ln\beta$ for $\beta > 0$, $n \in \mathbb{N}$ we obtain

$$\lim_{n\to\infty} P\Big(A_n(\beta)\cdot\big(X_{n-k_n:n} - B_n(\beta)\big) \le x\Big) = \Phi(x) \quad \text{for all } x \in \mathbb{R}.$$

Alternatively,

$$A_n(\beta) = \sqrt{n}\,\frac{f\big(F^{-1}(1-\beta)\big)}{\sqrt{\beta(1-\beta)}} = \sqrt{\frac{n}{\beta(1-\beta)}}\exp\big(-(-\ln(\beta))\big) = \beta\sqrt{\frac{n}{\beta(1-\beta)}} = \sqrt{\frac{n\beta}{1-\beta}}$$

and $B_n(\beta) = -F^{-1}(1-\beta) = -\big(-\ln(\beta)\big) = \ln\beta$ for $0 < \beta < 1$ and $n \in \mathbb{N}$,

as before.

**Remark:** Smirnov's Theorem II is related to a statistical estimate of the Value at Risk on the basis of i.i.d. risks $X_1, \cdots, X_n$ distributed as $X$ by

$$\widehat{\mathrm{VaR}_\alpha(X)} := X_{n-k_n:n} \text{ with } k_n := \lfloor \alpha \cdot n \rfloor.$$

Also, in this case,

$$P\text{-}\lim_{n\to\infty} X_{n-k_n:n} = \mathrm{VaR}_\alpha(X).$$

**Multivariate extremes:**

If for $d \geq 2$ $\{\mathbf{X}_n\}_{n\in\mathbb{N}}$ is an i.i.d. sequence of $d$-dimensional random vectors $\mathbf{X}_n = (X_{1n}, \cdots, X_{dn})$ with marginal cdf's $F_1, \cdots, F_d$ and Copula C, then the copula $C_{(n)}$ with

$$C_{(n)}(u_1, \cdots, u_d) = C^n\left(u_1^{1/n}, \cdots, u_d^{1/n}\right) \text{ for } 0 \leq u_1, \cdots, u_d \leq 1$$

is the copula of the random vector

$$\mathbf{X}_{n:n} := \left(\max\{X_{11}, \cdots, X_{1n}\}, \cdots, \max\{X_{d1}, \cdots, X_{dn}\}\right) \text{ for } n \in \mathbb{N}.$$

Proof: For $x_1, \cdots, x_d \in \mathbb{R}$ we have

$$P\left(\max\{X_{11}, \cdots, X_{1n}\} \leq x_1, \cdots, \max\{X_{d1}, \cdots, X_{dn}\} \leq x_d\right) = P\left(\bigcap_{i=1}^{d}\bigcap_{j=1}^{n}\{X_{ij} \leq x_i\}\right)$$

$$= P\left(\bigcap_{j=1}^{n}\bigcap_{i=1}^{d}\{X_{ij} \leq x_i\}\right) = C^n\left(F_1(x_1), \cdots, F_d(x_d)\right) = C^n\left(\left\{F_1^n(x_1)\right\}^{1/n}, \cdots, \left\{F_d^n(x_d)\right\}^{1/n}\right).$$

**Corollary:** If an extreme value limit copula $C_{(\infty)}(u_1,\cdots,u_d) = \lim_{n\to\infty} C^n\left(u_1^{1/n},\cdots,u_d^{1/n}\right)$ exists and each $F_k$ is in the domain of attraction of some extreme value cdf $G_k$ for $k = 1,\cdots,d$, then there exist normalizing vector constants $\mathbf{A}_n > \mathbf{0}$ and $\mathbf{B}_n \in \mathbb{R}^d$ such that

$$\lim_{n\to\infty} P\left(\mathbf{A}_n \cdot (\mathbf{X}_{n:n} - \mathbf{B}_n) \le \mathbf{x}\right) = C_{(\infty)}(G_1(x_1),\cdots,G_k(x_d)) \text{ for all } \mathbf{x} = (x_1,\cdots,x_d) \in \mathbb{R}^d.$$

**Remark:** If the original copula is *max-stable*, i.e.

$$C_{(n)}(u_1,\cdots,u_d) = C^n\left(u_1^{1/n},\cdots,u_d^{1/n}\right) = C(u_1,\cdots,u_d) \text{ for all } 0 \le u_1,\cdots,u_d \le 1,$$

then the condition $F_k \in \mathcal{D}(G_k)$ for $k = 1,\cdots,d$ is already sufficient for the weak convergence of normalized vector maxima $\mathbf{A}_n \cdot (\mathbf{X}_{n:n} - \mathbf{B}_n)$, $n \in \mathbb{N}$.

**Record values and record times:**

Let $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables with a continuous cdf $F$.

The sequence $\{U_n\}_{n\in\mathbb{Z}^+}$ of (upper) record times is recursively defined by

$$U_0 := 1,\ U_n := \inf\{k > U_{n-1} | X_k > X_{U_{n-1}}\} \text{ for } n \in \mathbb{N}.$$

The $\{X_{U_n}\}_{n\in\mathbb{Z}^+}$ are called (upper) record values.

**Essential properties:**

- $U_n$ is measurable and a.s. finite for all $n \in \mathbb{N}$
- $X_{U_n}$ is measurable and a.s. finite for all $n \in \mathbb{N}$
- $\{U_n\}_{n\in\mathbb{Z}^+}$ is a homogeneous Markov chain with transition probabilities

$$P(U_n = k | U_{n-1} = j) = \frac{j}{k(k-1)} \text{ for } 1 \le j < k \in \mathbb{N},\ n \in \mathbb{N}.$$

**Record values and record times:**

Let $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables with a continuous cdf $F$. The sequence $\{U_n\}_{n\in\mathbb{Z}^+}$ of (upper) record times is recursively defined by

$$U_0 := 1,\ U_n := \inf\{k > U_{n-1} | X_k > X_{U_{n-1}}\} \text{ for } n \in \mathbb{N}.$$

The $\{X_{U_n}\}_{n\in\mathbb{Z}^+}$ are called (upper) record values.

**Essential properties:**

- $\{(U_n, X_{U_n})\}_{n\in\mathbb{Z}^+}$ is a homogeneous Markov chain with transition probabilities

$$P(U_n = k, X_{U_n} > y | U_{n-1} = j, X_{U_{n-1}} = x) = F^{k-j-1}(x)\cdot(1-F(y)) \text{ a.s.}, 1 \le j < k \in \mathbb{N}, x \le y \in \mathbb{R}$$

- $\{X_{U_n}\}_{n\in\mathbb{Z}^+}$ is a homogeneous Markov chain with transition probabilities

$$P(X_{U_n} > y | X_{U_{n-1}} = x) = \frac{1-F(y)}{1-F(x)},\ x \le y \in \mathbb{R}$$

**Limit relations for record times:**

$$\lim_{n\to\infty} E(\ln U_n) - (n+1-\gamma) = 0 \qquad (\gamma = \text{Euler's constant})$$

$$\lim_{n\to\infty} Var(\ln U_n) - \left(n+1-\frac{\pi^2}{6}\right) = 0$$

$$\lim_{n\to\infty} P\left(\frac{\ln U_n - n}{\sqrt{n}} \le x\right) = \Phi(x), \ x \in \mathbb{R}$$

$$\lim_{n\to\infty} \frac{1}{n} \ln U_n = 1 \text{ a.s.}$$

Sketch of proof: Let $\{Y_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of $\mathcal{E}(1)$-distributed random variables. Define $V_0 := 1$, $V_n := \lceil V_{n-1} \cdot e^{Y_n} \rceil$, $n \in \mathbb{N}$ (rounding up). Then $\{V_n\}_{n\in\mathbb{Z}^+}$ is a homogeneous Markov chain with transition probabilities

$$P(V_n = k \,|\, V_{n-1} = j) = P\left(\lceil V_{n-1} \cdot e^{Y_n} \rceil = k \,|\, V_{n-1} = j\right) = P\left(\lceil j \cdot e^{Y_n} \rceil = k\right) = P\left(k-1 < j \cdot e^{Y_n} \le k\right)$$

$$= P\left(\ln\left(\frac{k-1}{j}\right) < Y_n \le \ln\left(\frac{k-1}{j}\right)\right) = \frac{j}{k-1} - \frac{j}{k} = \frac{j}{k(k-1)}, \ 1 \le j < k \in \mathbb{N}$$

**Limit relations for record times:**

$$\lim_{n\to\infty} E(\ln U_n) - (n+1-\gamma) = 0 \qquad (\gamma = \text{Euler's constant})$$

$$\lim_{n\to\infty} Var(\ln U_n) - \left(n+1-\frac{\pi^2}{6}\right) = 0$$

$$\lim_{n\to\infty} P\left(\frac{\ln U_n - n}{\sqrt{n}} \le x\right) = \Phi(x), \quad x \in \mathbb{R}$$

$$\lim_{n\to\infty} \frac{1}{n} \ln U_n = 1 \text{ a.s.}$$

which means that $\{V_n\}_{n\in\mathbb{Z}^+}$ has the same distribution as $\{U_n\}_{n\in\mathbb{Z}^+}$. Now

$$V_{n-1}\cdot e^{Y_n} < V_n \le V_{n-1}\cdot e^{Y_n}\cdot\left(1+\frac{e^{-Y_n}}{V_{n-1}}\right) \le V_{n-1}\cdot e^{Y_n}\cdot\left(1+\frac{1}{n}\right) \text{ for all } n\in\mathbb{N}$$

implying that $\ln V_{k-1} + Y_k < \ln V_k \le \ln V_{k-1} + Y_k + \ln\left(1+\frac{1}{k}\right)$ for all $k\in\mathbb{N}$, hence

$$\sum_{k=1}^{n} Y_k < \ln V_n = \sum_{k=1}^{n} (\ln V_k - \ln V_{k-1}) \le \sum_{k=1}^{n} Y_k + \sum_{k=1}^{n} \ln\left(1+\frac{1}{k}\right) = \sum_{k=1}^{n} Y_k + \ln n \text{ for all } n\in\mathbb{N}.$$

**Limit relations for record times:**

$$\lim_{n\to\infty} E(\ln U_n) - (n+1-\gamma) = 0 \qquad (\gamma = \text{Euler's constant})$$

$$\lim_{n\to\infty} Var(\ln U_n) - \left(n+1-\frac{\pi^2}{6}\right) = 0$$

$$\lim_{n\to\infty} P\left(\frac{\ln U_n - n}{\sqrt{n}} \le x\right) = \Phi(x), \ x \in \mathbb{R}$$

$$\lim_{n\to\infty} \frac{1}{n} \ln U_n = 1 \text{ a.s.}$$

The last two limit relations now follow from the Central Limit Theorem and the Law of Large Numbers, applied to the sequence $\{Y_n\}_{n\in\mathbb{N}}$.

For the first two relations, a more sophisticated estimation for the difference $\ln\left|V_{n-1}\cdot e^{Y_n}\right| - \ln V_{n-1}$, $n \in \mathbb{N}$ is need.

**Resnick's Theorem:** Let $G^*$ be a non-degenerate cdf and $\{X_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables with cdf $F$. If there exists $A_n > 0$, $B_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} P\left(A_n \cdot \left(X_{U_n} - B_n\right) \le x\right) = G^*(x) \text{ for all } x \in \mathcal{C}(G^*),$$

then $G^*$ is necessarily of the form

$$G^*(x) = \Phi\left(-2\ln\left(-\ln G(x)\right)\right), \; x \in \mathbb{R},$$

where $G$ is one of the three extreme value cdf's of the Extremal Types Theorem.

A necessary and sufficient condition for weak convergence of normalized record values is

$$F^* := 1 - \exp\left(-\sqrt{-\ln(1-F)}\right) \in \mathcal{D}(G).$$

# IV. Monte Carlo Methods

**Random Number Generation:** random numbers $\{u_n\}_{n\in\mathbb{Z}^+}$ (should) imitate the stochastic behaviour of a sequence of i.i.d. random variables from a continuous uniform distribution over the interval $[0,1]$. Critical aspects:

- *uniformity* (also in higher dimensions)
- *independence* (also of blocks built form disjoint sections).

A computer oriented generation of $\{u_n\}_{n\in\mathbb{Z}^+}$ is usually done by recursion. Example: multiplicative congruential method:

$$u_n := \frac{z_n}{m},\ z_{n+1} := a\cdot z_n \mod m \quad \text{for } n\in\mathbb{Z}^+$$

with appropriate natural numbers $a,m\in\mathbb{N}$ and a positive seed $0<z_0<m$. In order to exclude zero from the sequence, m is typically a prime or power of a prime; also, $a$ and $m$ must not have a common divisor.

Example 16: Choose $z_0 = 1$, hence $z_n = a^n \bmod m$. For the case $m = 13$ we obtain the following results:

| $a$ | $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | period length |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $z_n$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | $z_n$ | 2 | 4 | 8 | 3 | 6 | 12 | 11 | 9 | 5 | 10 | 7 | 1 | 12 |
| 3 | $z_n$ | 3 | 9 | 1 | 3 | 9 | 1 | 3 | 9 | 1 | 3 | 9 | 1 | 3 |
| 4 | $z_n$ | 4 | 3 | 12 | 9 | 10 | 1 | 4 | 3 | 12 | 9 | 10 | 1 | 6 |
| 5 | $z_n$ | 5 | 12 | 8 | 1 | 5 | 12 | 8 | 1 | 5 | 12 | 8 | 1 | 4 |
| 6 | $z_n$ | 6 | 10 | 8 | 9 | 2 | 12 | 7 | 3 | 5 | 4 | 11 | 1 | 12 |

Example 16: Choose $z_0 = 1$, hence $z_n = a^n \bmod m$. For the case $m = 13$ we obtain the following results:

| $a$ | $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | period length |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | $z_n$ | 7 | 10 | 5 | 9 | 11 | 12 | 6 | 3 | 8 | 4 | 2 | 1 | 12 |
| 8 | $z_n$ | 8 | 12 | 5 | 1 | 8 | 12 | 5 | 1 | 8 | 12 | 5 | 1 | 4 |
| 9 | $z_n$ | 9 | 3 | 1 | 9 | 3 | 1 | 9 | 3 | 1 | 9 | 3 | 1 | 3 |
| 10 | $z_n$ | 10 | 9 | 12 | 3 | 4 | 1 | 10 | 9 | 12 | 3 | 4 | 1 | 6 |
| 11 | $z_n$ | 11 | 4 | 5 | 3 | 7 | 12 | 2 | 9 | 8 | 10 | 6 | 1 | 12 |
| 12 | $z_n$ | 12 | 1 | 12 | 1 | 12 | 1 | 12 | 1 | 12 | 1 | 12 | 1 | 2 |

Remark:

- The number „1“ occurs in every sequence at some time.
- The period length correponds to the number of elements in the sequence until „1“ is reached, and is a divisor of $m-1$.
- The maximal period length is $m-1$.
- In case of $a=m-1$ the period length is always identical to 2 (since $a^2=m^2-2m+1=m(m-2)+1$, i.e. there is a remainder of 1 when dividing by $m$).

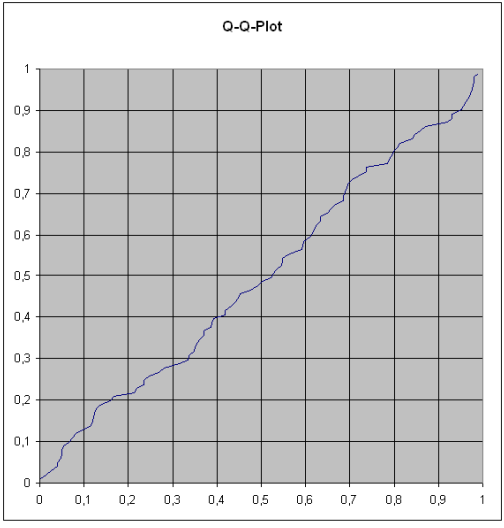
The background of these observations is a fundamental Theorem of Euler / Fermat / Lagrange in number theory in case that $m$ is a prime. More precisely, the period length corresponds to the order of the cyclical group generated by $a$.

**Consequences:**

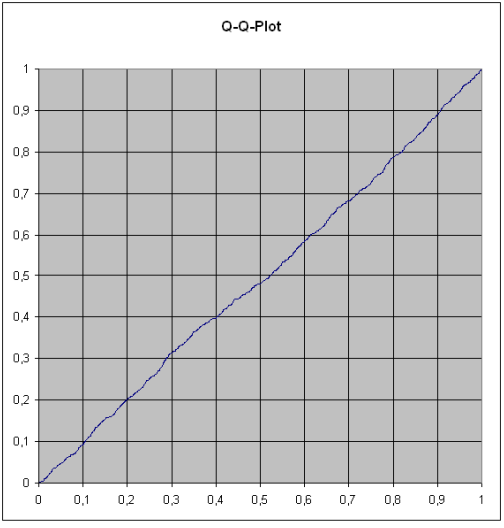
- In order to avoid too short period lengths and local dependencies the parameters of the multiplicative congruential method must be chosen very carefully.
- In case of $m = 2^K$, it is necessary to choose $a \equiv 3 \bmod 8$ or $a \equiv 5 \bmod 8$ in order to achieve the maximal period length of $2^{K-2} = m/4$.

A drawback of such simple generation methods is the fact that consecutive pairs, or more general, *m*-tuples built of such a sequence can concentrate on low-dimensional subspaces.

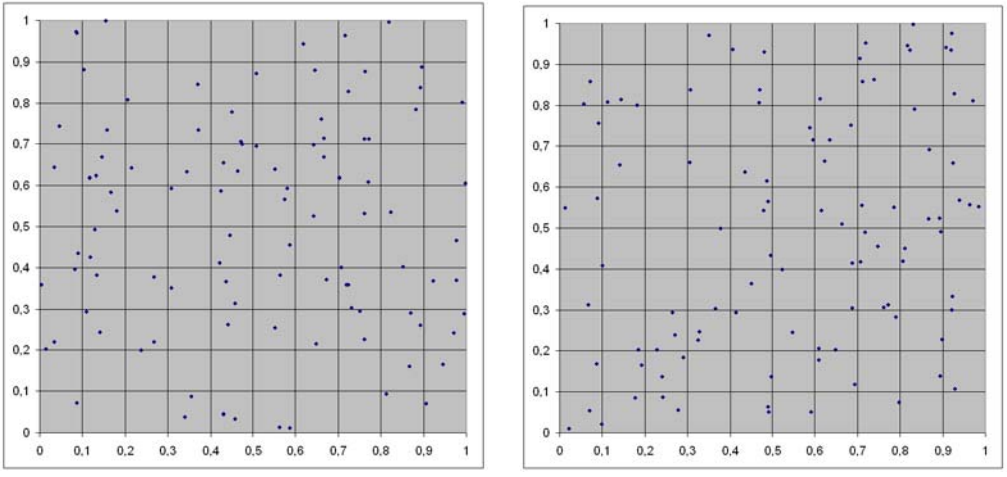
The quality of random numbers should generally be checked by statistical tests. The following graphs show Q-Q-plots for 100 and 1000 random numbers, resp., generated with EXCEL (German command ZUFALLSZAHL()).




100 random numbers

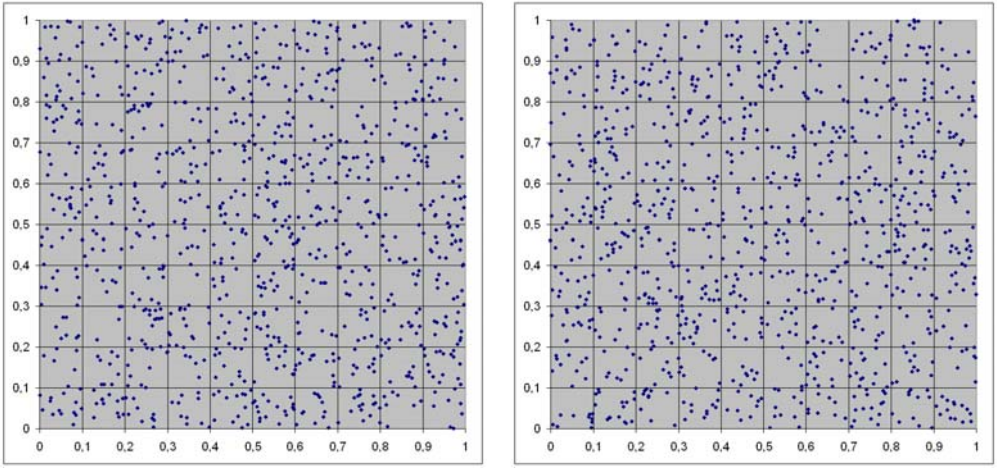



1000 random numbers




100 pairs of random numbers generated by EXCEL

Ideally, these points should be uniformly distributed over the unit square.




1000 pairs of random numbers generated by EXCEL

Ideally, these points should be uniformly distributed over the unit square.

**Random number generation for general distributions: Inversion Method:**

Reminder: Let $F$ denote the cdf of an arbitrary random variable $X$ and $F^{-1}$ the corresponding quantile function

$$F^{-1}(u) := \inf\{x \in \mathbb{R} | F(x) \geq u\}, \ 0 < u < 1,$$

then it holds:

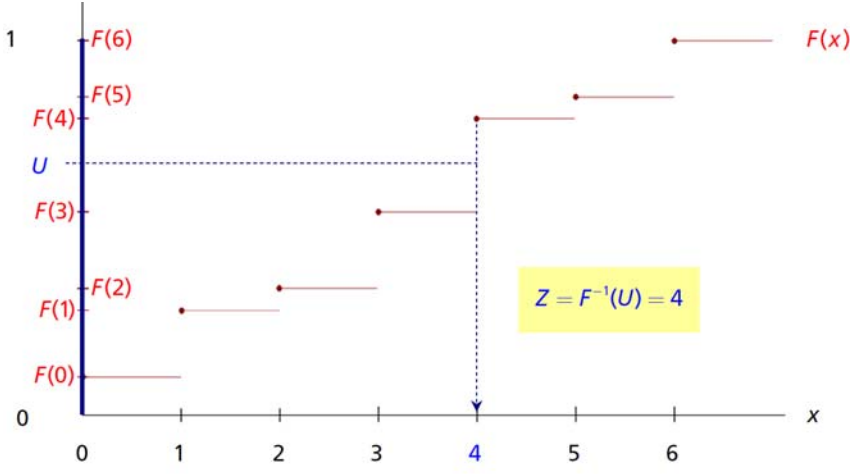
- $U := F(X)$ is a random number, if $F$ is *continuous*;
- $Z := F^{-1}(U)$ is – without restriction – distributed as $X$, if $U$ follows a continuous uniform distribution over the interval $[0,1]$.

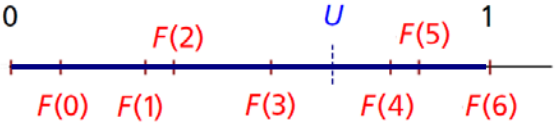
Note that $U$ and $1-U$ have the same distribution.

Application: Generate random numbers $\{u_n\}_{n\in\mathbb{Z}^+}$ and use $\{F^{-1}(u_n)\}_{n\in\mathbb{Z}^+}$ or $\{F^{-1}(1-u_n)\}_{n\in\mathbb{Z}^+}$ as i.i.d. "realizations" of the random variable $X$.

The Inversion Method is particularly suited for the generation of random numbers from discrete distributions and can be easily implemented in EXCEL.

Example 17: The following graph shows how the method works, for a random variable $X$ with values in the set $\{0,1,\cdots,6\}$.






Here the value of $U$ generates the discrete random value „4". If $U$ falls into the intervall $\big(F(k-1), F(k)\big]$ with $F(-1) = 0$, then let $Z := k$; in this case, we have

$$P(Z = k) = P\big(F(k-1) < U \leq F(k)\big) = F(k) - F(k-1).$$

Example 18: discrete uniform distribution over $\{1,2,\cdots,n\}$: here we have

$$F(k)=\frac{k}{n},\ k=0,\cdots,n$$

with $Z=k$ iff $\frac{k-1}{n}<U\leq\frac{k}{n}$. The inversion method here simplifies to

$$Z=\lceil n\cdot U\rceil \quad \text{or} \quad Z=\lfloor n\cdot U\rfloor+1, \quad \text{resp.,}$$

where

$$\lceil z\rceil:=\min\{m\in\mathbb{Z}\,|\,z\leq m\} \qquad \text{(rounding up)}$$
$$\lfloor z\rfloor:=\max\{m\in\mathbb{Z}\,|\,z\geq m\} \qquad \text{(rounding down).}$$

The inversion method works in general very well in all cases where the quantile function $Q = F^{-1}$ has a closed form. Examples:

- Exponential distribution $\mathcal{E}(\lambda)$: $X = -\frac{1}{\lambda}\ln(U)$ for $\lambda > 0$
- Cauchy distribution $\mathcal{C}$: $X = \cot(\pi U)$
- Pareto distribution $\mathcal{P}a(\alpha)$: $X = U^{-1/\alpha} - 1$ for $\alpha > 0$
- Logistic distribution $\mathcal{L}$: $X = \ln\left(\frac{U}{1-U}\right)$
- Loglogistic distribution $\mathcal{LL}$: $X = \frac{U}{1-U}$
- Fréchet distribution $\mathcal{F}(\alpha)$: $X = \{-\ln(U)\}^{-1/\alpha}$ for $\alpha > 0$
- Gumbel distribution $\mathcal{G}$: $X = -\ln(-\ln(U))$
- Weibull type min distribution $\mathcal{W}(\alpha)$: $X = \{-\ln(U)\}^{1/\alpha}$ for $\alpha > 0$.

In case that the quantile function $Q = F^{-1}$ cannot be simply determined, it might be useful to discretize the (continuous) cdf $F$ in an appropriate way. E.g., if $X$ denotes the underlying random variable, we can alternatively consider the rounded random variable $X_\Delta$:

$$X_\Delta := \Delta \cdot \left\lfloor \frac{X}{\Delta} \right\rfloor$$ with a (typically small) step size $\Delta > 0$.

Then we have

$$F_{X_\Delta}(x) = F\left(\Delta \cdot \left\lfloor \frac{x}{\Delta} \right\rfloor\right), \; x \in \mathbb{R} \text{ and } P(X_\Delta = k\Delta) = F(k\Delta) - F((k-1)\Delta), \; k \in \mathbb{Z}.$$

The inversion method for discrete distributions can easily be implemented in EXCEL using the (German) command SVERWEIS. We first discuss the case that the random variable $X$ takes only the values $0,1,2,\ldots,M$, with

$$f(k) = P(X = k) \text{ and } F(k) = P(X \leq k) = \sum_{i=0}^{k} f(i) \text{ for } 0 \leq k \leq M.$$

Construction:

1. Create a spreadsheet T (for Table) with the following entries:

| | A | B | C |
|---|---|---|---|
| 1 | $F(k)$ | $k+1$ | $f(k)$ |
| 2 | 0 | 0 | 0 |
| 3 | $F(0)$ | 1 | $f(0)$ |
| 4 | $F(1)$ | 2 | $f(1)$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $M+3$ | $F(M)$ | $M+1$ | $f(M)$ |

The inversion method for discrete distributions can easily be implemented in EXCEL using the (German) command SVERWEIS. We first discuss the case that the random variable $X$ takes only the values 0,1,2,...,$M$, with

$$f(k) = P(X = k) \text{ and } F(k) = P(X \leq k) = \sum_{i=0}^{k} f(i) \text{ for } 0 \leq k \leq M.$$

Construction:

2. Create a spreadsheet S (for Simulation) with the following output:

| | A | B |
|---|---|---|
| 1 | U | Z |
| 2 | =ZUFALLSZAHL() | =SVERWEIS(A2;T!$A$2:T!$B$[$M+3$];2) |
| 3 | =ZUFALLSZAHL() | =SVERWEIS(A3;T!$A$2:T!$B$[$M+3$];2) |
| 4 | =ZUFALLSZAHL() | =SVERWEIS(A4;T!$A$2:T!$B$[$M+3$];2) |
| ⋮ | ⋮ | ⋮ |

**Remarks:**

- In case that the random variable $X$ is unbounded (examples: Poisson distribution, negative binomial distribution), the distribution of $X$ can be truncated for a sufficiently large $M$, say if $P(X > M)$ is small in comparison with the number of simulations performed.
- In case that the random variable $X$ is distributed over an arithmetic set $\mathcal{M} = \{a + k\Delta | k = 0, \cdots, M\}$ for some $a \in \mathbb{R}$ and $\Delta > 0$, then the random variable $Y = \frac{X - a}{\Delta}$ is again distributed over the set $\{0, \cdots, M\}$. $Y$ can hence be simulated by the procedure described above, and $X$ can finally be simulated using the transformation $X = a + \Delta Y$. (This is e.g. of importance when using the discretization method.)
- The procedure described above is also suited to simulate multi-dimensional discrete distributions by using an appropriate denumeration technique.

Example 19: The following table describes the joint distribution of a discrete random vector $(X,Y)$:

| $P(X=x,\, Y=y)$ | | $x$ | | | | |
|---|---|---|---|---|---|---|
| | | 0 | 50 | 100 | $P(Y=y)$ | $P(Y\leq y)$ |
| $y$ | 0 | 0.100 | 0.340 | 0.000 | 0.440 | 0.440 |
| | 40 | 0.454 | 0.100 | 0.001 | 0.555 | **0.995** |
| | 50 | 0.000 | 0.001 | 0.004 | 0.005 | 1.000 |
| | $P(X=x)$ | 0.554 | 0.441 | 0.005 | | |
| | $P(X\leq x)$ | 0.554 | **0.995** | 1.000 | | |

The random variables $X$ and $Y$ are not stochastically independent; their correlation is given by $\rho(X,Y) = -0.5536$.

Construction:

1. We denumerate the 9 entries in the table above in the following way:

| $z$ | $x$ | $y$ |
|---|---|---|
| 1 | 0 | 0 |
| 2 | 50 | 0 |
| 3 | 100 | 0 |
| 4 | 0 | 40 |
| 5 | 50 | 40 |
| 6 | 100 | 40 |
| 7 | 0 | 50 |
| 8 | 50 | 50 |
| 9 | 100 | 50 |

table R

| $F(k)$ | $k+1$ | $f(k)$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 0.1 | 2 | 0.1 |
| 0.44 | 3 | 0.34 |
| 0.44 | 4 | 0 |
| 0.894 | 5 | 0.454 |
| 0.994 | 6 | 0.1 |
| 0.995 | 7 | 0.001 |
| 0.995 | 8 | 0 |
| 0.996 | 9 | 0.001 |
| 1 | 10 | 0.004 |

table T

Construction:

2. Using the command SVERWEIS again we obtain the following output table:

| | A | B | C | D |
|---|---|---|---|---|
| 1 | U | Z | X | Y |
| 2 | … | … | =SVERWEIS(B2;R!$A$2:$C$10;2) | =SVERWEIS(B2;R!$A$2:$C$10;3) |
| 3 | … | … | =SVERWEIS(B3;R!$A$2:$C$10;2) | =SVERWEIS(B3;R!$A$2:$C$10;3) |
| 4 | … | … | =SVERWEIS(B4;R!$A$2:$C$10;2) | =SVERWEIS(B4;R!$A$2:$C$10;3) |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 9 | … | … | =SVERWEIS(B9;R!$A$2:$C$10;2) | =SVERWEIS(B9;R!$A$2:$C$10;3) |

**Acceptance-Rejection Method:** In case that $X$ has a density $f$ which is concentrated on some (finite) interval $[a,b]$ and which is bounded by some positive constant $M>0$, then the so called acceptance-rejection method is applicable. It is based on two jointly i.i.d. sequences $\{U_n\}_{n\in\mathbb{N}}$ and $\{V_n\}_{n\in\mathbb{N}}$ where $U_n$ follows a continuous uniform distribution over the interval $[a,b]$ and $V_n$ follows a continuous uniform distribution over the interval $[0,M]$ for all $n\in\mathbb{N}$. Consider the stopping time

$$N := \inf\{n\in\mathbb{N} | V_n \leq f(U_n)\}.$$

Then

$$Z := U_N$$

is distributed as $X$.

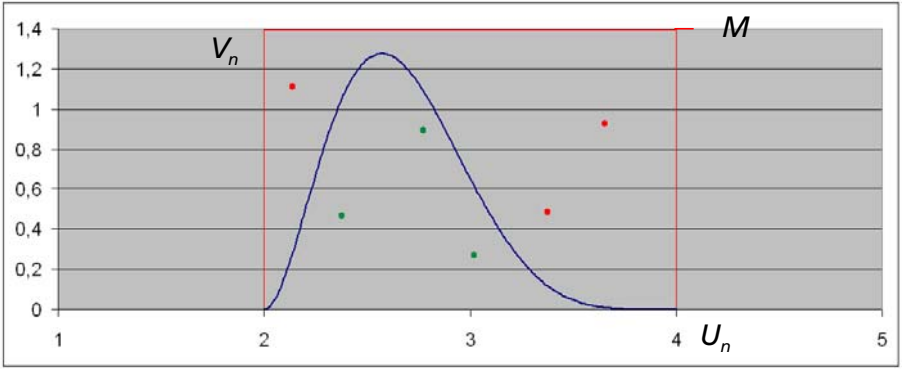
**Remarks:**

- $N$ is a.s. finite and follows a geometric distribution over $\mathbb{N}$ with $P(N=k)=p(1-p)^k$ for $k\in\mathbb{N}$ where $p=\frac{1}{M\cdot(b-a)}$ (the *acceptance rate*)
- $E(N)=\frac{1}{p}=M\cdot(b-a).$
- The random variable $U_N$ is a.s. well-defined.
- The random variables $U_n$ and $V_n$ can be obtained from i.i.d. random variables $\tilde{U}_n$ and $\tilde{V}_n$ which follow a continuous uniform distribution over the interval $[0,1]$ each by the following positive-linear transformations:

  $U_n=a+(b-a)\cdot\tilde{U}_n,\ V=M\cdot\tilde{V}_n$ for $n\in\mathbb{N}$.

  In practical applications, $\tilde{U}_n$ and $\tilde{V}_n$ are replaced by random numbers.

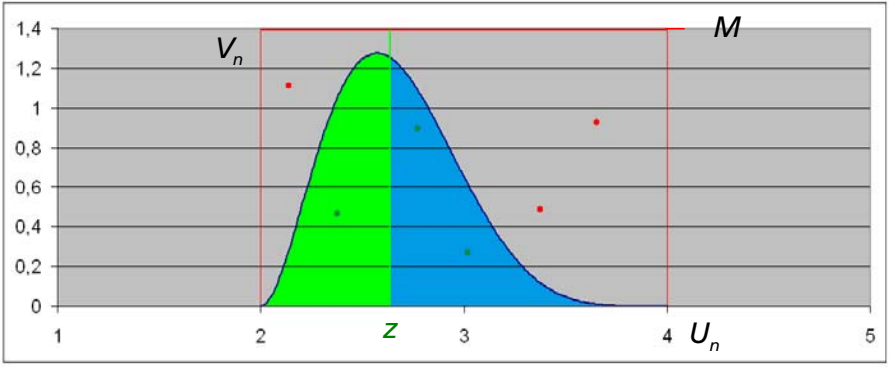
**Explanation:**




green points: $(U_n, V_n)$ with $V_n \leq f(U_n)$: acceptance; $Z = U_N$

red points: $(U_n, V_n)$ with $V_n > f(U_n)$: rejection

**Explanation:**




$$P(Z \le z) = P\big(U \le z \,|\, V \le f(U)\big) = \frac{P\big(U \le z, V \le f(U)\big)}{P\big(V \le f(U)\big)}$$

$$= \frac{\text{green area} / \{(b-a)\cdot M\}}{(\text{green + blue area}) / \{(b-a)\cdot M\}} = \frac{\text{green area}}{\text{green + blue area}} = \frac{\int_a^z f(x)dx}{\int_a^b f(x)dx} = F(z), \;\; a \le z \le b$$

**Remarks:**

- Instead of the density $f$ itself also an arbitrary multiple $g := c \cdot f$ of the density can be used with $c > 0$. This avoids the (sometimes tedious) calculation of the norming constant for the density $f$ (e.g. in case of a beta distribution).
- The acceptance-rejection method is not restricted to the one-dimensional case, but can be generalized to the multivariate case: let $\mathbf{X}$ be a $d$-dimensional random vector with a density $f$ that is concentrated on the interval $D = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d]$ and is bounded by some positive constant $M > 0$. Let further $\{\mathbf{U}_n\}_{n \in \mathbb{N}}$ and $\{V_n\}_{n \in \mathbb{N}}$ be jointly i.i.d. sequences where $\mathbf{U}_n$ follows a continuous uniform distribution over the interval $D$ and $V_n$ follows a continuous uniform distribution over the interval $[0, M]$ for all $n \in \mathbb{N}$. Consider again the stopping time $N := \inf\{n \in \mathbb{N} | V_n \leq f(\mathbf{U}_n)\}$. Then $\mathbf{Z} := \mathbf{U}_N$ is distributed as $\mathbf{X}$.

**Remarks:**

- $N$ is a.s. finite and follows a geometric distribution over $\mathbb{N}$ with

  $P(N=k)=p(1-p)^k$ for $k\in\mathbb{N}$ where $p=\dfrac{c}{M\cdot\prod_{i=1}^{d}(b_i-a_i)}$ (the *acceptance rate*)

- $E(N)=\dfrac{1}{p}=\dfrac{M}{c}\cdot\prod_{i=1}^{d}(b_i-a_i).$

- The random vector $\mathbf{U}_N$ is a.s. well-defined.

- The acceptance-rejection method can - also in arbitrary dimensions - be generalized to cases where the underlying density $f$ is not concentrated on a finite interval or is unbounded.

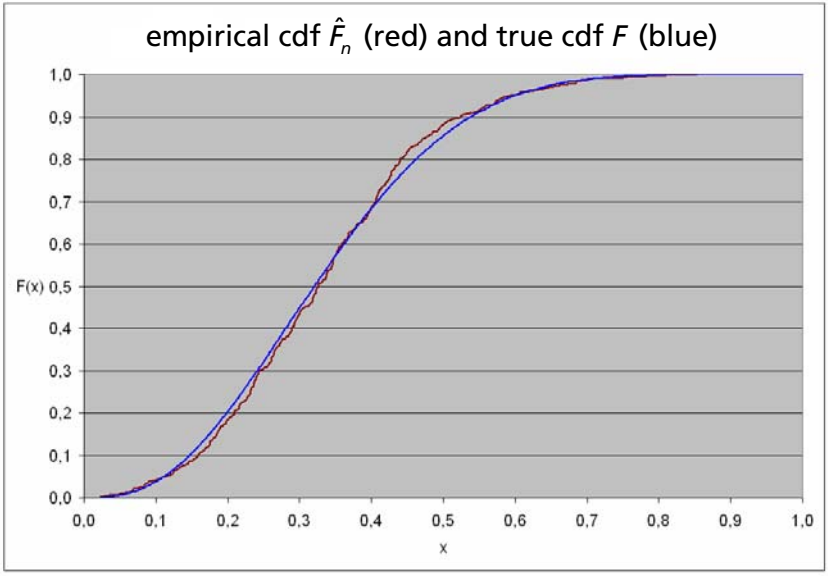
Example 20: The (bounded) density of the Beta-distribution $\mathcal{B}(a,b)$ for $a,b \geq 1$ is given by

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \text{ for } 0 \leq x \leq 1$$

with the inverse norming constant (Beta function)

$$B(a,b) := \int_0^1 x^{a-1}(1-x)^{b-1}\,dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

For an application of the acceptance-rejection method it suffices to consider just the function $g(x) = x^{a-1}(1-x)^{b-1}$ for $0 \leq x \leq 1$. Its maximum is attained for $x = \frac{a-1}{a+b-2}$, if $a+b > 2$. For $a = 3$ und $b = 6$ the minimal upper bound M is given by $M = \frac{2^2 \cdot 5^5}{7^7} = \frac{12500}{823543} < 0.0152$. The acceptance rate here is 39%.




Simulation with EXCEL on the basis of 1,000 pairs of random numbers; the empirical acceptance rate is 39.9%

Example 21: Comparison of three simulation methods:

Consider a random variable $X$ with density

$$f(x) = \frac{1}{b} \cdot (a+1) \cdot (a+b+1) \cdot x^a \cdot \left(1 - x^b\right) \text{ for } 0 \leq x \leq 1 \text{ and } a, b > 0.$$

The corresponding cdf is given by

$$F(x) = \frac{x^{a+1}}{b} \cdot \left((a+b+1) - (a+1) \cdot x^b\right) \text{ for } 0 \leq x \leq 1 \text{ and } a, b > 0.$$

For $a = 1$ and $b = 2$ we obtain:

Method 1 (inversion): Here we have

$$F(x) = 2x^2 - x^4 \text{ for } 0 \leq x \leq 1$$

with the quantile function

$$Q(u) = \sqrt{1 - \sqrt{1-u}} \text{ for } 0 \leq u \leq 1.$$

Example 21: Comparison of three simulation methods:

Consider a random variable $X$ with density

$$f(x) = \frac{1}{b} \cdot (a+1) \cdot (a+b+1) \cdot x^a \cdot \left(1 - x^b\right) \text{ for } 0 \le x \le 1 \text{ and } a, b > 0.$$

The corresponding cdf is given by

$$F(x) = \frac{x^{a+1}}{b} \cdot \left((a+b+1) - (a+1) \cdot x^b\right) \text{ for } 0 \le x \le 1 \text{ and } a, b > 0.$$

For $a = 1$ and $b = 2$ we obtain:

Method 2 (acceptance-rejection): Here we have

$$f(x) = 4x \cdot \left(1 - x^2\right) \text{ for } 0 \le x \le 1$$

with a maximal value of $\frac{8}{9}\sqrt{3} < 1.54 =: M$ and an acceptance rate of 64.9%.

Example 21: Comparison of three simulation methods:

Consider a random variable $X$ with density

$$f(x) = \frac{1}{b} \cdot (a+1) \cdot (a+b+1) \cdot x^a \cdot \left(1 - x^b\right) \text{ for } 0 \leq x \leq 1 \text{ and } a, b > 0.$$

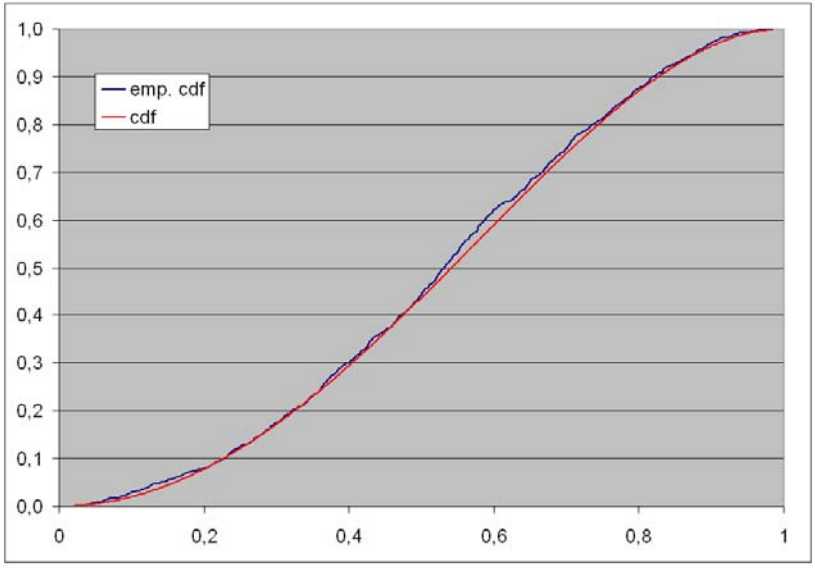
The corresponding cdf is given by

$$F(x) = \frac{x^{a+1}}{b} \cdot \left((a+b+1) - (a+1) \cdot x^b\right) \text{ for } 0 \leq x \leq 1 \text{ and } a, b > 0.$$

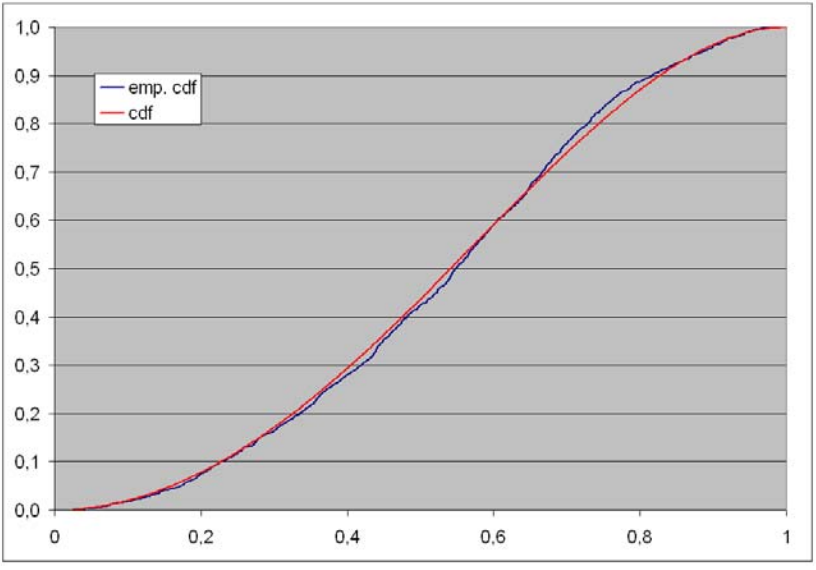
For $a = 1$ and $b = 2$ we obtain:

Method 3 (discretization): With $\Delta = 0.0005$ we obtain

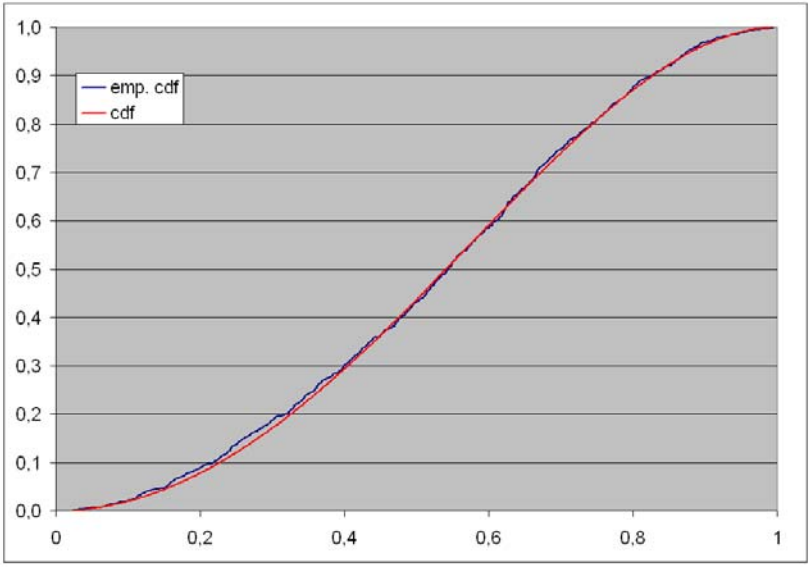
$$f(k\Delta) = F(k\Delta) - F\left((k-1)\Delta\right) \text{ for } k = 1, 2, \cdots, 2000.$$




inversion



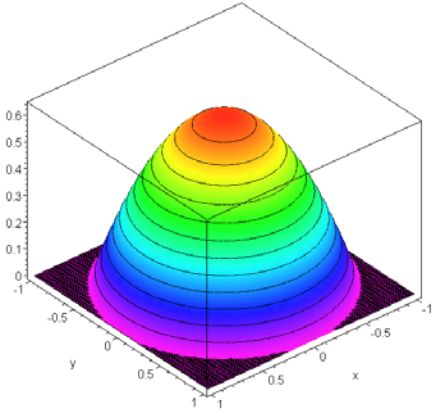

acceptance-rejection




discretization

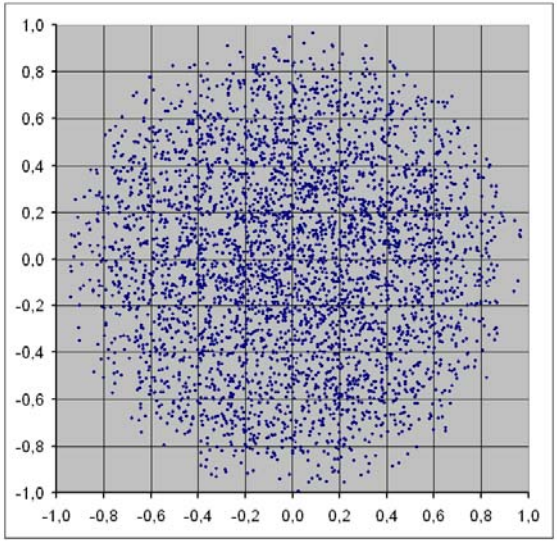
Example 22: Simulation of a two-dimensional random vector $(X,Y)$ with a radial symmetric density:

$$f(x,y)=\begin{cases}\frac{2}{\pi}\left(1-x^2-y^2\right) & \text{if } x^2+y^2\le 1\\ 0 & \text{otherwise}\end{cases} \quad \text{for } 0\le x,y\le 1.$$




Graph of the density $f(x,y)$

The density reaches its unique maximum in the origin with a value of $M = \frac{2}{\pi} = 0.6366...$ The acceptance rate is $\pi / 8 = 39.27\%$.




Simulation with 10000 triples of random numbers; emp. acceptance rate: 39.6%

EXCEL also provides special commands for the cdf and quantile function of the beta-, normal-, lognormal and other distributions so that simulations can be performed directly on the basis of random numbers.

Normal distribution $\mathcal{N}\left(\mu, \sigma^2\right)$:

| cdf $F(x)$ | quantile function $Q(u)$ |
|---|---|
| NORMVERT(x;$\mu$;$\sigma$;WAHR) | NORMINV(u;$\mu$;$\sigma$) |

Lognormal distribution $\mathcal{LN}\left(\mu, \sigma^2\right)$:

| cdf $F(x)$ | quantile function $Q(u)$ |
|---|---|
| LOGNORMVERT(x;$\mu$;$\sigma$) | LOGINV(u;$\mu$;$\sigma$) |

EXCEL also provides special commands for the cdf and quantile function of the beta-, normal-, lognormal and other distributions so that simulations can be performed directly on the basis of random numbers.

Beta distribution $\mathcal{B}(\alpha,\beta)$:

| cdf $F(x)$ | quantile function $Q(u)$ |
|---|---|
| BETAVERT($x;\alpha;\beta$) | BETAINV($u;\alpha;\beta$) |

Gamma distribution $\Gamma(\alpha,\lambda)$:

| cdf $F(x)$ | quantile function $Q(u)$ |
|---|---|
| GAMMAVERT($x;\alpha;1/\lambda$;WAHR) | GAMMAINV($u;\alpha;1/\lambda$) |

EXCEL also provides special commands for the cdf and quantile function of the beta-, normal-, lognormal and other distributions so that simulations can be performed directly on the basis of random numbers.

*t* distribution $t_d$ (with *d* degrees of freedom):

| cdf $F(x)$ | quantile function $Q(u)$ |
|---|---|
| WENN($x$ > 0;1 – TVERT($x$;$d$;1);TVERT(– $x$;$d$;1)) | =WENN($u$ < 0,5; – TINV(2*$u$;$d$);TINV(2–2*$u$;$d$)) |

$\chi^2$ distribution $\chi^2_d$ (with *d* degrees of freedom):

| cdf $F(x)$ | quantile function $Q(u)$ |
|---|---|
| 1 – CHIVERT($x$;$d$) | CHIINV(1–$u$;$d$) |

**Particular methods:** Let througout $\{U_n\}_{n\in\mathbb{N}}$ be an i.i.d. sequence of random variables with a continuous uniform distribution over the interval $[0,1]$.

1. Binomial distribution $B(n,p)$ for $n\in\mathbb{N}$, $p\in(0,1)$:

$$Z := \sum_{k=1}^{n} \lfloor U_k + p \rfloor$$

2. Negative binomial distribution $NB(n,p)$ for $n\in\mathbb{N}$, $p\in(0,1)$:

$$Z := \sum_{k=1}^{n} \left\lfloor \frac{\ln(U_k)}{\ln(1-p)} \right\rfloor$$

3. Poisson distribution $\mathcal{P}(\lambda)$ for $\lambda > 0$:

$$Z := \inf\left\{ n\in\mathbb{Z}^+ \,\middle|\, \prod_{k=1}^{n+1} U_k \le e^{-\lambda} \right\} = \inf\left\{ n\in\mathbb{Z}^+ \,\middle|\, -\frac{1}{\lambda}\sum_{k=1}^{n+1} \ln(U_k) > 1 \right\}$$

**Jöhnk's beta method:** Let $U$ and $V$ be i.i.d. random variables with a continuous uniform distribution over the interval $[0,1]$ and $\alpha, \beta > 0$. Define

$$T := U^{1/\alpha} + V^{1/\beta} \text{ and } S := \frac{U^{1/\alpha}}{U^{1/\alpha} + V^{1/\beta}} = \frac{U^{1/\alpha}}{T}. \quad (*)$$

Then we have

$$P^S(\bullet \mid T \leq 1) = B(\alpha, \beta).$$

This relation can be used to simulate a $\mathcal{B}(\alpha, \beta)$ distribution by a modified acceptance-rejection method: Let $\{U_n\}_{n\in\mathbb{N}}$ and $\{V_n\}_{n\in\mathbb{N}}$ be jointly i.i.d. where $U_n$ and $V_n$ both follow a continuous uniform distribution over the interval $[0,1]$. Define $T_n$ and $S_n$ according to (*) for $n \in \mathbb{N}$ and let

$$N := \inf\{n \in \mathbb{N} \mid T_n \leq 1\}.$$

Then $S_N$ follows a $\mathcal{B}(\alpha, \beta)$ distribution.

This method is appropriate for "small" values of $\alpha, \beta > 0$.

**Jöhnk's gamma method** to simulate a $\Gamma(\alpha,\lambda)$ distribution for $\alpha,\lambda>0$:

Case I: $\alpha<1$: Let $X$ and $Y$ be independent with $X$ following a $\mathcal{B}(\alpha,1-\alpha)$ distribution and $Y$ following a $\Gamma(1,\lambda)=\mathcal{E}(\lambda)$ distribution. Then $Z:=X\cdot Y$ follows a $\Gamma(\alpha,\lambda)$ distribution.

Case II: $\alpha\geq 1$: If $\alpha$ is integer then $Z:=-\frac{1}{\lambda}\sum_{k=1}^{\alpha}\ln(U_k)$ follows a $\Gamma(\alpha,\lambda)=\mathcal{E}(\alpha,\lambda)$ distribution (Erlang distribution). If $\alpha$ is not integer consider $m:=\lfloor\alpha\rfloor$ and $\beta:=\alpha-m<1$. Let $X$ and $Y$ be independent with $X$ following a $\Gamma(m,\lambda)$ distribution and $Y$ following a $\Gamma(\beta,\lambda)$ distribution. Then $Z:=X+Y$ follows a $\Gamma(\alpha,\lambda)$ distribution.

Simulation of a $\mathcal{B}(\alpha,\beta)$ distribution with "large" values of $\alpha,\beta>0$: Let $X$ and $Y$ be independent with $X$ following a $\Gamma(\alpha,1)$ distribution and $Y$ following a $\Gamma(\beta,1)$ distribution. Then $Z:=\frac{X}{X+Y}$ follows a $\mathcal{B}(\alpha,\beta)$ distribution.

Particular topic: Estimation of $\mathrm{VaR}_\alpha(X)$ by simulation (continuous case):

Method 1:

- Simulate $n \in \mathbb{N}$ independent copies $X_1, \cdots, X_n$ of $X$
- form the order statistics $X_{1:n}, \cdots, X_{n:n}$ by sorting
- choose $X_{n-k_n:n}$ as an estimate for $\mathrm{VaR}_\alpha(X)$ with $k_n := \lfloor \alpha \cdot n \rfloor$ (cf. slide 114).

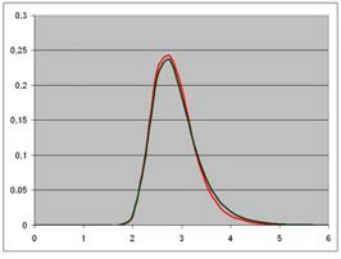
Method 2:

- Fix $n \in \mathbb{N}$ and simulate a $\mathcal{B}(n-k_n, k_n+1)$-distributed random variable $U$
- choose $Q_X(U) = F_X^{-1}(U)$ as an estimate for $\mathrm{VaR}_\alpha(X)$.

Remark: $Q_X(U)$ from method 2 has the same distribution as $X_{n-k_n:n}$ from method 1 (cf. slide 84), but requires only 1 simulation instead of $n$.

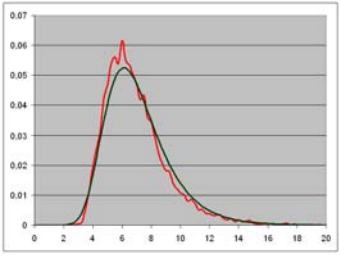
Example 23a: The following graphs show empirical densities (red, with a manual fit to a Gumbel density, green) for 10000 estimates each for $\mathrm{VaR}_{0.005}(X)$ after method 2 where $X$ follows a $\mathcal{LN}\left(-\frac{\sigma^2}{2}, \sigma^2\right)$-distribution (cf. slide 26, Example 4).

$$n = 200,\ k_n = 1$$



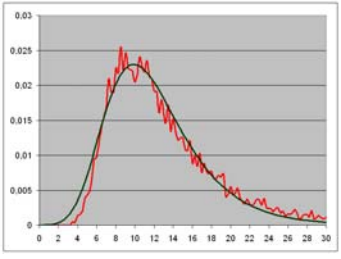

$\sigma = 0.5$

$\mathrm{VaR}_{0.005}(X) = 3.1992$




$\sigma = 1.0$

$\mathrm{VaR}_{0.005}(X) = 7.9711$

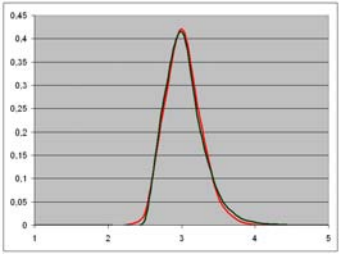



$\sigma = 1.5$

$\mathrm{VaR}_{0.005}(X) = 15.4675$

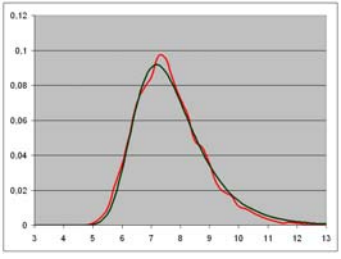
Example 23a: The following graphs show empirical densities (red, with a manual fit to a Gumbel density, green) for 10000 estimates each for $\mathrm{VaR}_{0.005}(X)$ after method 2 where $X$ follows a $\mathcal{LN}\left(-\frac{\sigma^2}{2}, \sigma^2\right)$-distribution (cf. slide 26, Example 4).

$$n = 1000,\ k_n = 5$$



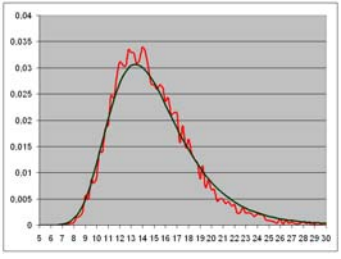

$\sigma = 0.5$

$\mathrm{VaR}_{0.005}(X) = 3.1992$




$\sigma = 1.0$

$\mathrm{VaR}_{0.005}(X) = 7.9711$




$\sigma = 1.5$

$\mathrm{VaR}_{0.005}(X) = 15.4675$

Particular topic: Estimation of $\mathrm{ES}_\alpha(X)$ by simulation (continuous case):

Method 1:

- Simulate $n \in \mathbb{N}$ independent copies $X_1, \cdots, X_n$ of $X$
- form the order statistics $X_{1:n}, \cdots, X_{n:n}$ by sorting
- choose $\dfrac{1}{k_n+1}\sum_{k=0}^{k_n} X_{n-k:n}$ as an estimate for $\mathrm{ES}_\alpha(X)$ with $k_n := \lfloor \alpha \cdot n \rfloor$.

Method 2:

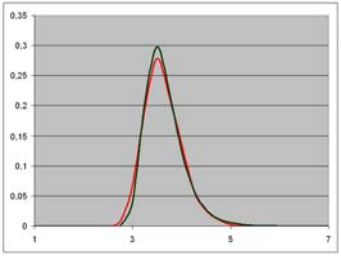
- Fix $n \in \mathbb{N}$; simulate $k_n+1$ independent random variables $U_0, \cdots, U_{k_n}$ with a continuous uniform distribution over the interval $[0,1]$
- choose $\dfrac{1}{k_n+1}\sum_{k=0}^{k_n} Q_X\left(\prod_{j=0}^{k} U_j^{1/(n-j)}\right)$ as an estimate for $\mathrm{ES}_\alpha(X)$.

Remark1: The random variable $\frac{1}{k_n+1}\sum_{k=0}^{k_n} Q_X\left(\prod_{j=0}^{k} U_j^{1/(n-j)}\right)$ from method 2 has the same distribution as the random variable $\frac{1}{k_n+1}\sum_{k=0}^{k_n} X_{n-k:n}$ from method 1, but requires only $k_n+1$ simulations instead of $n$. This is due to the Markov property of order statistics in the continuous case (cf. slide 84).

Remark 2: The random variables $U_j^{1/(n-j)}$ from method 2 follow a $\mathcal{B}(n-j,1)$ beta distribution for $j=0,\cdots,k_n$. This corresponds to the fact that for i.i.d. random variables $V_1,\cdots,V_n$ distributed as $V$ with a continuous uniform distribution over the interval $[0,1]$, the conditional distribution of an order statistic $V_{r-1:n}$ given $V_{r:n}=v_{rn}$ for a fixed $v_{rn}\in(0,1)$ is identical with the distribution of the maximum of $r-1$ independent copies of $v_{rn}\cdot V$, for $1<r\leq n$.

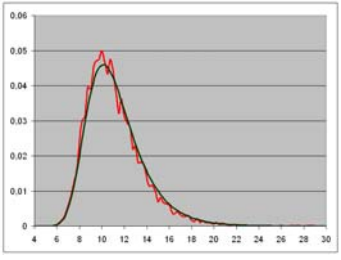
Example 23b: The following graphs show empirical densities (red, with a manual fit to a Gumbel density, green) for 10000 estimates each for $\mathrm{ES}_{0.005}(X)$ after method 2 where $X$ follows a $\mathcal{LN}\left(-\frac{\sigma^2}{2}, \sigma^2\right)$-distribution (cf. slide 26, Example 4).

$$n = 1000,\ k_n = 5$$



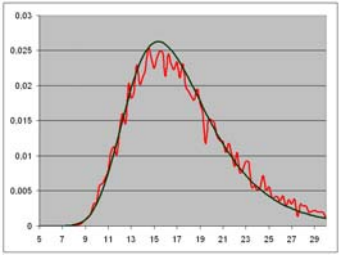

$\sigma = 0.5$

$\mathrm{ES}_{0.005}(X) = 3.7910$




$\sigma = 1.0$

$\mathrm{ES}_{0.005}(X) = 11.5065$




$\sigma = 1.25$

$\mathrm{ES}_{0.005}(X) = 18.4896$

Simulation of $d$-dimensional random vectors $\mathbf{X} = (X_1, \cdots, X_d)^T$:

Multivariate normal distribution:

$P^{\mathbf{X}} = \mathcal{N}(\mathbf{0}, \Sigma)$ with a non-negative definite variance-covariance matrix $\Sigma$:

- decompose $\Sigma = A \cdot A^T$ with a suitable $d \times d$ matrix $A$ (Cholesky or spectral decomposition)
- generate a random vector $Y$ with $P^{\mathbf{Y}} = \mathcal{N}(\mathbf{0}, \mathbf{I})$ (independent components)
- generate $\mathbf{X}$ via $\mathbf{X} = A\mathbf{Y}$.

Simulation of $d$-dimensional random vectors $\mathbf{X} = (X_1, \cdots, X_d)^T$ :

Multivariate normal distribution:

Special case $d = 2$: Box-Muller transformation:

If U and V are independent continuous uniformly distributed over the interval [0,1] and

$$X = \sqrt{-2\ln(U)}\cos(2\pi V)$$
$$Y = \sqrt{-2\ln(U)}\cos(2\pi V + \alpha)$$

with $\alpha \in [0, \pi]$, then $\mathbf{Z} = (X, Y)^T$ follows a $\mathcal{N}(\mathbf{0}, \Sigma)$-distribution with

$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ where $\rho = \cos(\alpha)$ (cf. also Chapter II).

Simulation of $d$-dimensional random vectors $\mathbf{X} = (X_1, \cdots, X_d)^T$ :

Multivariate $t_\nu$ -distribution:

$P^{\mathbf{X}} = t_\nu(\mathbf{0}, \Sigma)$ with a non-negative definite matrix $\Sigma$:

- decompose $\Sigma = A \cdot A^T$ with a suitable $d \times d$ matrix $A$ (Cholesky or spectral decomposition)
- generate a random vector $\mathbf{Y}$ with $P^{\mathbf{Y}} = \mathcal{N}(\mathbf{0}, \Sigma)$
- generate a random variable $W$, independent of $\mathbf{Y}$, with $P^W = \chi^2_\nu$
- generate $\mathbf{X}$ via $\mathbf{X} = \sqrt{\frac{\nu}{W}} A\mathbf{Y}$.

By linear transformations, general normal and $t$-distributions can be simulated. Correspondingly, also Gauß- and $t$- copulas can be simulated (cf. Chapter II).

Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with a Gauß copula $C_\Sigma^G$ with a positive definite correlation matrix $\Sigma$:

- decompose $\Sigma = A \cdot A^T$ with a suitable $d \times d$ matrix $A$ (Cholesky or spectral decomposition)
- generate a random vector $\mathbf{Y}$ with $P^{\mathbf{Y}} = \mathcal{N}(\mathbf{0}, \mathbf{I})$
- generate a random vector $\mathbf{Z}$ via $\mathbf{Z} = A\mathbf{Y}$
- generate the random vector $\mathbf{U}$ by $U_i := \Phi(Z_i)$ for $i = 1, \cdots, d$.

Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with a $t$-copula $C_\Sigma^{t_\nu}$ with a positive definite correlation matrix $\Sigma$:

- decompose $\Sigma = A \cdot A^T$ with a suitable $d \times d$ matrix $A$ (Cholesky or spectral decomposition)
- generate a random vector $\mathbf{Y}$ with $P^{\mathbf{Y}} = \mathcal{N}(\mathbf{0}, \mathbf{I})$
- generate a random variable $W$, independent of $\mathbf{Y}$, with $P^W = \chi^2_\nu$
- generate $\mathbf{Z}$ via $\mathbf{Z} = \sqrt{\frac{\nu}{W}} A\mathbf{Y}$
- generate the random vector $\mathbf{U}$ by $U_i := F_\nu(Z_i)$ for $i = 1, \cdots, d$, where $F_\nu$ denotes the cdf of a univariate $t_\nu$-distribution.

Note that in EXCEL, a matrix multiplication function exists which allows for an easy calculation of $\mathbf{Z} = A\mathbf{Y}$ both for the Gauß as well as for the $t$-copula.

Example 24: Simulation of a Gauß and $t$-copula with a given correlation matrix

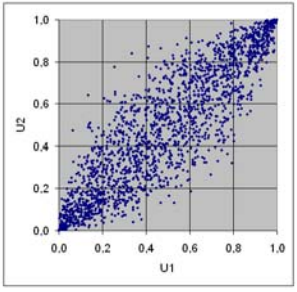
$$\Sigma = \begin{bmatrix} 1.00 & 0.90 & 0.85 & 0.80 \\ 0.90 & 1.00 & 0.80 & 0.70 \\ 0.85 & 0.80 & 1.00 & 0.50 \\ 0.80 & 0.70 & 0.50 & 1.00 \end{bmatrix}.$$

The Cholesky decomposition $\Sigma = A \cdot A^T$ gives (rounded to two decimal places)

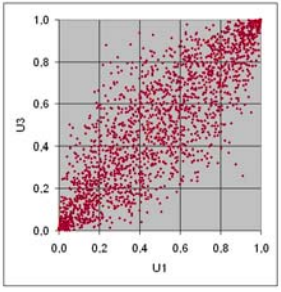
$$A = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.95 & 0.44 & 0.00 & 0.00 \\ 0.90 & 0.08 & 0.52 & 0.00 \\ 0.85 & -0.05 & -0.34 & 0.49 \end{bmatrix}.$$

The following graphs show the pairwise simulated copula random vectors $(U_i, U_j)$ with $1 \leq i < j \leq 4$, from a sample of size 2000:

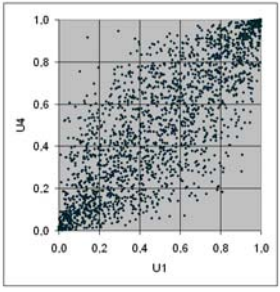
Gauß copula:



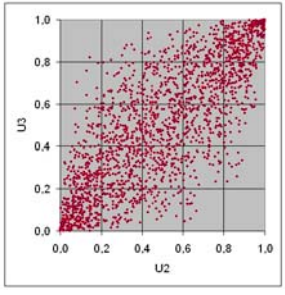

$\rho_{12} = 0.90$



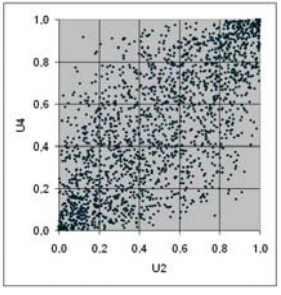

$\rho_{13} = 0.85$



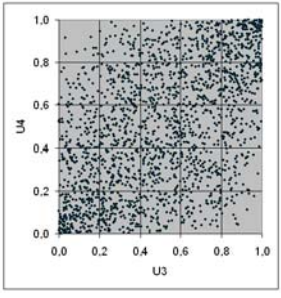

$\rho_{14} = 0.80$




$\rho_{23} = 0.80$

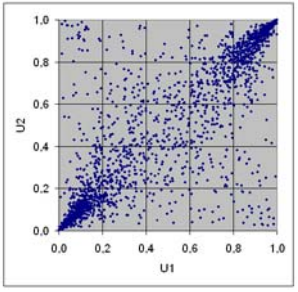



$\rho_{24} = 0.70$

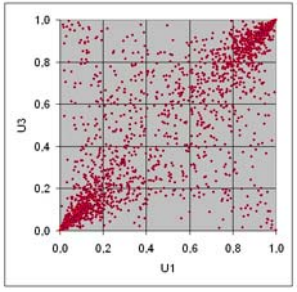



$\rho_{34} = 0.50$

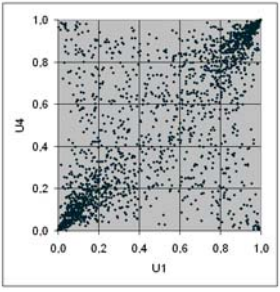
$t_1$-copula:



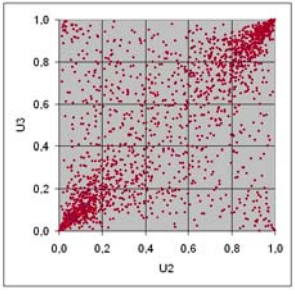

$\rho_{12} = 0.90$



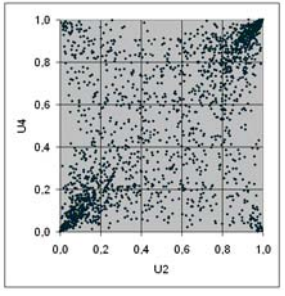

$\rho_{13} = 0.85$



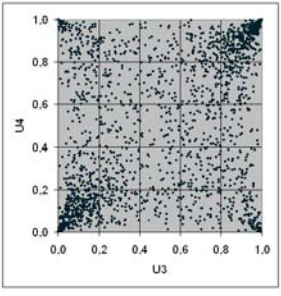

$\rho_{14} = 0.80$




$\rho_{23} = 0.80$

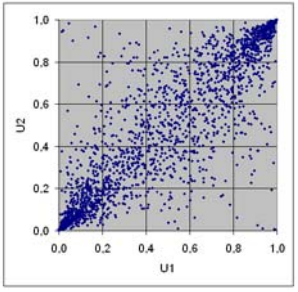



$\rho_{24} = 0.70$

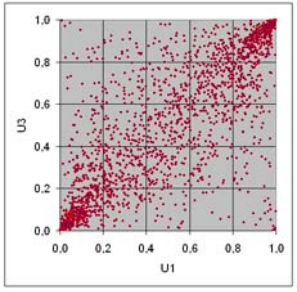



$\rho_{34} = 0.50$

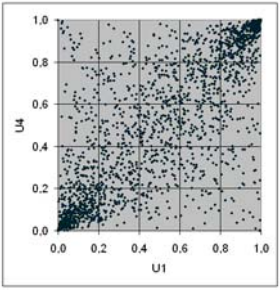
$t_2$-copula:



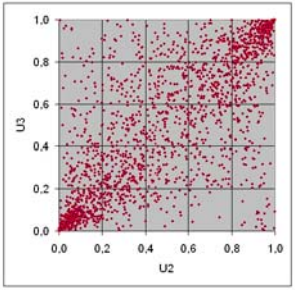

$\rho_{12} = 0.90$



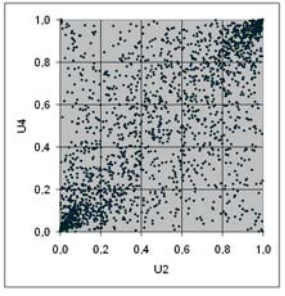

$\rho_{13} = 0.85$



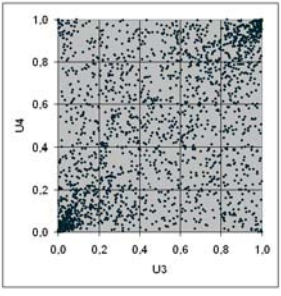

$\rho_{14} = 0.80$




$\rho_{23} = 0.80$

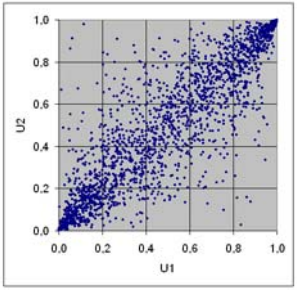



$\rho_{24} = 0.70$

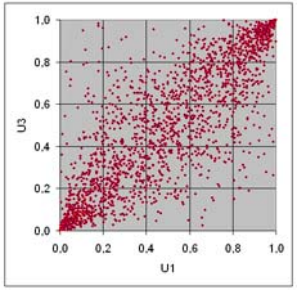



$\rho_{34} = 0.50$

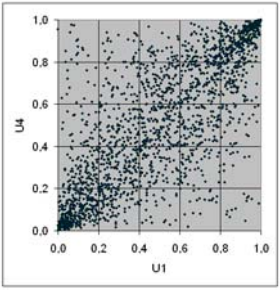
$t_3$-copula:



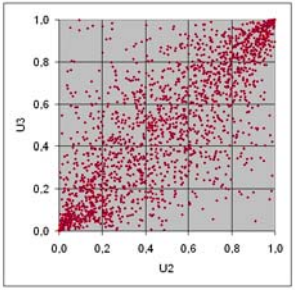

$\rho_{12} = 0.90$



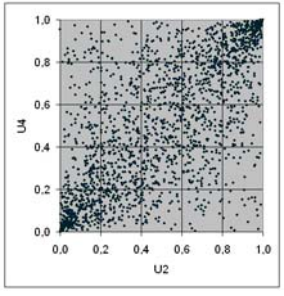

$\rho_{13} = 0.85$



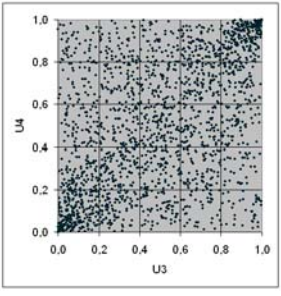

$\rho_{14} = 0.80$




$\rho_{23} = 0.80$




$\rho_{24} = 0.70$




$\rho_{34} = 0.50$

I6 =NORMINV('random numbers'!A6,0;1)

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | empirical variances | | | | empirical variances | | | | Sigma | | | | Cholesky A | | | | A^T | | | |
| 2 | | | | | 1,0538 | 1,0215 | 1,0172 | 1,0468 | 1,0538 | 0,9747 | 1,0109 | 0,9592 | 1,00 | 0,90 | 0,85 | 0,80 | 1,00 | 0,00 | 0,00 | 0,00 | 1,00 | 0,90 | 0,85 | 0,80 |
| 3 | | | | | | | | | | | | | 0,90 | 1,00 | 0,80 | 0,70 | 0,90 | 0,44 | 0,00 | 0,00 | 0,00 | 0,44 | 0,08 | -0,05 |
| 4 | U1 | U2 | U3 | U4 | Z1 | Z2 | Z3 | Z4 | Y1 | Y2 | Y3 | Y4 | 0,85 | 0,80 | 1,00 | 0,50 | 0,85 | 0,08 | 0,52 | 0,00 | 0,00 | 0,00 | 0,52 | -0,34 |
| 5 | | | | | | | | | | | | | 0,80 | 0,70 | 0,50 | 1,00 | 0,80 | -0,05 | -0,34 | 0,49 | 0,00 | 0,00 | 0,00 | 0,49 |
| 6 | 0,1786 | 0,1246 | 0,1142 | 0,2592 | -0,9208 | -1,1523 | -1,2046 | -0,6459 | -0,9208 | -0,7424 | -0,6960 | -0,3630 | | | | | | | | | | | | |
| 7 | 0,3128 | 0,2641 | 0,3670 | 0,2548 | -0,4880 | -0,6309 | -0,3398 | -0,6594 | -0,4880 | -0,4397 | 0,2118 | -0,4409 | | | | | | | | | | | | |
| 8 | 0,5610 | 0,5499 | 0,6342 | 0,2136 | 0,1534 | 0,1255 | 0,3429 | -0,7940 | 0,1534 | -0,0288 | 0,4126 | -1,5782 | | | | | | | | | | | | |
| 9 | 0,7613 | 0,6678 | 0,5013 | 0,7907 | 0,7103 | 0,4339 | 0,0032 | 0,8088 | 0,7103 | -0,4712 | -1,0809 | -0,2984 | | | | | | | | | | | | |
| 10 | 0,0472 | 0,1575 | 0,2945 | 0,0577 | -1,6723 | -1,0046 | -0,5402 | -1,5741 | -1,6723 | 1,1481 | 1,5156 | 0,6685 | | | | | | | | | | | | |

E6 {=MMULT(I6:L6;$U$2:$X$5)}

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | empirical variances | | | | empirical variances | | | | Sigma | | | | Cholesky A | | | | A^T | | | |
| 2 | | | | | 1,0538 | 1,0215 | 1,0172 | 1,0468 | 1,0538 | 0,9747 | 1,0109 | 0,9592 | 1,00 | 0,90 | 0,85 | 0,80 | 1,00 | 0,00 | 0,00 | 0,00 | 1,00 | 0,90 | 0,85 | 0,80 |
| 3 | | | | | | | | | | | | | 0,90 | 1,00 | 0,80 | 0,70 | 0,90 | 0,44 | 0,00 | 0,00 | 0,00 | 0,44 | 0,08 | -0,05 |
| 4 | U1 | U2 | U3 | U4 | Z1 | Z2 | Z3 | Z4 | Y1 | Y2 | Y3 | Y4 | 0,85 | 0,80 | 1,00 | 0,50 | 0,85 | 0,08 | 0,52 | 0,00 | 0,00 | 0,00 | 0,52 | -0,34 |
| 5 | | | | | | | | | | | | | 0,80 | 0,70 | 0,50 | 1,00 | 0,80 | -0,05 | -0,34 | 0,49 | 0,00 | 0,00 | 0,00 | 0,49 |
| 6 | 0,1786 | 0,1246 | 0,1142 | 0,2592 | -0,9208 | -1,1523 | -1,2046 | -0,6459 | -0,9208 | -0,7424 | -0,6960 | -0,3630 | | | | | | | | | | | | |
| 7 | 0,3128 | 0,2641 | 0,3670 | 0,2548 | -0,4880 | -0,6309 | -0,3398 | -0,6594 | -0,4880 | -0,4397 | 0,2118 | -0,4409 | | | | | | | | | | | | |
| 8 | 0,5610 | 0,5499 | 0,6342 | 0,2136 | 0,1534 | 0,1255 | 0,3429 | -0,7940 | 0,1534 | -0,0288 | 0,4126 | -1,5782 | | | | | | | | | | | | |
| 9 | 0,7613 | 0,6678 | 0,5013 | 0,7907 | 0,7103 | 0,4339 | 0,0032 | 0,8088 | 0,7103 | -0,4712 | -1,0809 | -0,2984 | | | | | | | | | | | | |
| 10 | 0,0472 | 0,1575 | 0,2945 | 0,0577 | -1,6723 | -1,0046 | -0,5402 | -1,5741 | -1,6723 | 1,1481 | 1,5156 | 0,6685 | | | | | | | | | | | | |

A6 =NORMVERT(E6;0;1;WAHR)

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | empirical variances | | | | empirical variances | | | | Sigma | | | | Cholesky A | | | | A^T | | | |
| 2 | | | | | 1,0538 | 1,0215 | 1,0172 | 1,0468 | 1,0538 | 0,9747 | 1,0109 | 0,9592 | 1,00 | 0,90 | 0,85 | 0,80 | 1,00 | 0,00 | 0,00 | 0,00 | 1,00 | 0,90 | 0,85 | 0,80 |
| 3 | | | | | | | | | | | | | 0,90 | 1,00 | 0,80 | 0,70 | 0,90 | 0,44 | 0,00 | 0,00 | 0,00 | 0,44 | 0,08 | -0,05 |
| 4 | U1 | U2 | U3 | U4 | Z1 | Z2 | Z3 | Z4 | Y1 | Y2 | Y3 | Y4 | 0,85 | 0,80 | 1,00 | 0,50 | 0,85 | 0,08 | 0,52 | 0,00 | 0,00 | 0,00 | 0,52 | -0,34 |
| 5 | | | | | | | | | | | | | 0,80 | 0,70 | 0,50 | 1,00 | 0,80 | -0,05 | -0,34 | 0,49 | 0,00 | 0,00 | 0,00 | 0,49 |
| 6 | 0,1786 | 0,1246 | 0,1142 | 0,2592 | -0,9208 | -1,1523 | -1,2046 | -0,6459 | -0,9208 | -0,7424 | -0,6960 | -0,3630 | | | | | | | | | | | | |
| 7 | 0,3128 | 0,2641 | 0,3670 | 0,2548 | -0,4880 | -0,6309 | -0,3398 | -0,6594 | -0,4880 | -0,4397 | 0,2118 | -0,4409 | | | | | | | | | | | | |
| 8 | 0,5610 | 0,5499 | 0,6342 | 0,2136 | 0,1534 | 0,1255 | 0,3429 | -0,7940 | 0,1534 | -0,0288 | 0,4126 | -1,5782 | | | | | | | | | | | | |
| 9 | 0,7613 | 0,6678 | 0,5013 | 0,7907 | 0,7103 | 0,4339 | 0,0032 | 0,8088 | 0,7103 | -0,4712 | -1,0809 | -0,2984 | | | | | | | | | | | | |
| 10 | 0,0472 | 0,1575 | 0,2945 | 0,0577 | -1,6723 | -1,0046 | -0,5402 | -1,5741 | -1,6723 | 1,1481 | 1,5156 | 0,6685 | | | | | | | | | | | | |

EXCEL worksheet structure for the Gauß copula

N6 =NORMINV('random numbers'!A6;0;1)

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | empirical variances | | | | | empirical variances | | | | empirical variances | | | |
| 2 | | | | | 4,6615 | 3,7464 | 4,2451 | 4,2622 | | 0,9992 | 0,9525 | 1,0229 | 0,9695 | 0,9992 | 0,9335 | 1,0174 | 1,0543 |
| 3 | | | | | | | | | | | | | | | | | |
| 4 | U1 | U2 | U3 | U4 | Z1 | Z2 | Z3 | Z4 | W | AY1 | AY2 | AY3 | AY4 | Y1 | Y2 | Y3 | Y4 |
| 5 | | | | | | | | | | | | | | | | | |
| 6 | 0,3365 | 0,2334 | 0,2149 | 0,6201 | -0,4659 | -0,8314 | -0,9102 | 0,3349 | 2,8774 | -0,3951 | -0,7051 | -0,7720 | 0,2841 | -0,3951 | -0,8018 | -0,7140 | 0,6520 |
| 7 | 0,8193 | 0,8764 | 0,8576 | 0,5338 | 1,0745 | 1,4332 | 1,2991 | 0,0922 | 4,7538 | 1,1714 | 1,5624 | 1,4162 | 0,1005 | 1,1714 | 1,1657 | 0,6279 | -1,1568 |
| 8 | 0,0488 | 0,1198 | 0,3182 | 0,0363 | -2,3801 | -1,4628 | -0,5242 | -2,7179 | 2,0844 | -1,7181 | -1,0560 | -0,3784 | -1,9620 | -1,7181 | 1,1249 | 1,9047 | 0,2213 |
| 9 | 0,1406 | 0,8321 | 0,9194 | 0,0203 | -1,3106 | 1,1434 | 1,8515 | -3,4613 | 0,6733 | -0,5377 | 0,4691 | 0,7596 | -1,4201 | -0,5377 | 2,1865 | 1,9998 | -0,4306 |
| 10 | 0,2502 | 0,1135 | 0,4518 | 0,0822 | -0,7642 | -1,5147 | -0,1315 | -1,8311 | 2,4153 | -0,5939 | -1,1771 | -0,1022 | -1,4229 | -0,5939 | -1,4742 | 1,0006 | -1,3720 |

J6 {=MMULT(N6:Q6;$Z$2:$AC$5)}

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | empirical variances | | | | | empirical variances | | | | empirical variances | | | |
| 2 | | | | | 4,6615 | 3,7464 | 4,2451 | 4,2622 | | 0,9992 | 0,9525 | 1,0229 | 0,9695 | 0,9992 | 0,9335 | 1,0174 | 1,0543 |
| 3 | | | | | | | | | | | | | | | | | |
| 4 | U1 | U2 | U3 | U4 | Z1 | Z2 | Z3 | Z4 | W | AY1 | AY2 | AY3 | AY4 | Y1 | Y2 | Y3 | Y4 |
| 5 | | | | | | | | | | | | | | | | | |
| 6 | 0,3365 | 0,2334 | 0,2149 | 0,6201 | -0,4659 | -0,8314 | -0,9102 | 0,3349 | 2,8774 | -0,3951 | -0,7051 | -0,7720 | 0,2841 | -0,3951 | -0,8018 | -0,7140 | 0,6520 |
| 7 | 0,8193 | 0,8764 | 0,8576 | 0,5338 | 1,0745 | 1,4332 | 1,2991 | 0,0922 | 4,7538 | 1,1714 | 1,5624 | 1,4162 | 0,1005 | 1,1714 | 1,1657 | 0,6279 | -1,1568 |
| 8 | 0,0488 | 0,1198 | 0,3182 | 0,0363 | -2,3801 | -1,4628 | -0,5242 | -2,7179 | 2,0844 | -1,7181 | -1,0560 | -0,3784 | -1,9620 | -1,7181 | 1,1249 | 1,9047 | 0,2213 |
| 9 | 0,1406 | 0,8321 | 0,9194 | 0,0203 | -1,3106 | 1,1434 | 1,8515 | -3,4613 | 0,6733 | -0,5377 | 0,4691 | 0,7596 | -1,4201 | -0,5377 | 2,1865 | 1,9998 | -0,4306 |
| 10 | 0,2502 | 0,1135 | 0,4518 | 0,0822 | -0,7642 | -1,5147 | -0,1315 | -1,8311 | 2,4153 | -0,5939 | -1,1771 | -0,1022 | -1,4229 | -0,5939 | -1,4742 | 1,0006 | -1,3720 |

EXCEL worksheet structure for the *t*-copula (I)

Arial 11 F K U Ω Bearbeitung zurückgenden... Bearbeitung beenden...

I6 =CHIINV('random numbers'!E6,$AH$2)

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | empirical variances | | | | | empirical variances | | | | empirical variances | | | |
| 2 | | | | | 4,6615 | 3,7464 | 4,2451 | 4,2622 | | 0,9992 | 0,9525 | 1,0229 | 0,9695 | 0,9992 | 0,9335 | 1,0174 | 1,0543 |
| 3 | | | | | | | | | | | | | | | | | |
| 4 | U1 | U2 | U3 | U4 | Z1 | Z2 | Z3 | Z4 | W | AY1 | AY2 | AY3 | AY4 | Y1 | Y2 | Y3 | Y4 |
| 5 | | | | | | | | | | | | | | | | | |
| 6 | 0,3365 | 0,2334 | 0,2149 | 0,6201 | -0,4659 | -0,8314 | -0,9102 | 0,3349 | 2,8774 | -0,3951 | -0,7051 | -0,7720 | 0,2841 | -0,3951 | -0,8018 | -0,7140 | 0,6520 |
| 7 | 0,8193 | 0,8764 | 0,8576 | 0,5338 | 1,0745 | 1,4332 | 1,2991 | 0,0922 | 4,7538 | 1,1714 | 1,5624 | 1,4162 | 0,1005 | 1,1714 | 1,1657 | 0,6279 | -1,1568 |
| 8 | 0,0488 | 0,1198 | 0,3182 | 0,0363 | -2,3801 | -1,4628 | -0,5242 | -2,7179 | 2,0844 | -1,7181 | -1,0560 | -0,3784 | -1,9620 | -1,7181 | 1,1249 | 1,9047 | 0,2213 |
| 9 | 0,1406 | 0,8321 | 0,9194 | 0,0203 | -1,3106 | 1,1434 | 1,8515 | -3,4613 | 0,6733 | -0,5377 | 0,4691 | 0,7596 | -1,4201 | -0,5377 | 2,1865 | 1,9998 | -0,4306 |
| 10 | 0,2502 | 0,1135 | 0,4518 | 0,0822 | -0,7642 | -1,5147 | -0,1315 | -1,8311 | 2,4153 | -0,5939 | -1,1771 | -0,1022 | -1,4229 | -0,5939 | -1,4742 | 1,0006 | -1,3720 |

Arial 11 F K U Ω Bearbeitung zurückgenden... Bearbeitung beenden...

A6 =WENN(E6>0;1-TVERT(E6;$AH$2;1);TVERT(-E6;$AH$2;1))

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | empirical variances | | | | | empirical variances | | | | empirical variances | | | |
| 2 | | | | | 4,6615 | 3,7464 | 4,2451 | 4,2622 | | 0,9992 | 0,9525 | 1,0229 | 0,9695 | 0,9992 | 0,9335 | 1,0174 | 1,0543 |
| 3 | | | | | | | | | | | | | | | | | |
| 4 | U1 | U2 | U3 | U4 | Z1 | Z2 | Z3 | Z4 | W | AY1 | AY2 | AY3 | AY4 | Y1 | Y2 | Y3 | Y4 |
| 5 | | | | | | | | | | | | | | | | | |
| 6 | 0,3365 | 0,2334 | 0,2149 | 0,6201 | -0,4659 | -0,8314 | -0,9102 | 0,3349 | 2,8774 | -0,3951 | -0,7051 | -0,7720 | 0,2841 | -0,3951 | -0,8018 | -0,7140 | 0,6520 |
| 7 | 0,8193 | 0,8764 | 0,8576 | 0,5338 | 1,0745 | 1,4332 | 1,2991 | 0,0922 | 4,7538 | 1,1714 | 1,5624 | 1,4162 | 0,1005 | 1,1714 | 1,1657 | 0,6279 | -1,1568 |
| 8 | 0,0488 | 0,1198 | 0,3182 | 0,0363 | -2,3801 | -1,4628 | -0,5242 | -2,7179 | 2,0844 | -1,7181 | -1,0560 | -0,3784 | -1,9620 | -1,7181 | 1,1249 | 1,9047 | 0,2213 |
| 9 | 0,1406 | 0,8321 | 0,9194 | 0,0203 | -1,3106 | 1,1434 | 1,8515 | -3,4613 | 0,6733 | -0,5377 | 0,4691 | 0,7596 | -1,4201 | -0,5377 | 2,1865 | 1,9998 | -0,4306 |
| 10 | 0,2502 | 0,1135 | 0,4518 | 0,0822 | -0,7642 | -1,5147 | -0,1315 | -1,8311 | 2,4153 | -0,5939 | -1,1771 | -0,1022 | -1,4229 | -0,5939 | -1,4742 | 1,0006 | -1,3720 |

EXCEL worksheet structure for the *t*-copula (II)

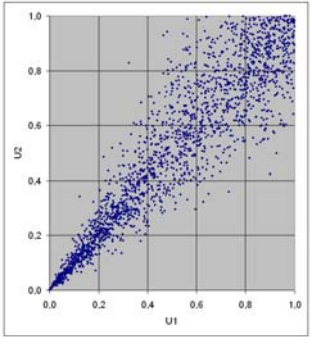
Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with an Archimedean copula with generator $\varphi$:

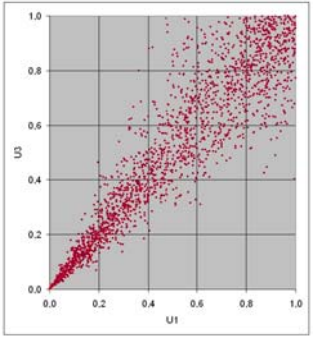
General approach:

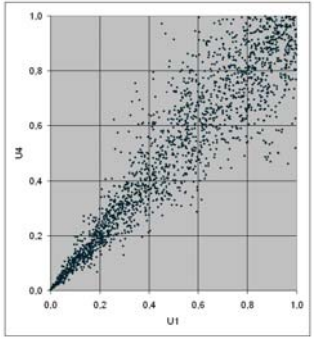
- Simulate $d$ independent copies $Y_1, \cdots, Y_d$ of a random variable $Y$ following a continuous uniform distribution over the interval $[0,1]$
- simulate a positive random variable $Z$ whose distribution has the Laplace transform $\varphi^{-1}$, independent of $Y_1, \cdots, Y_d$
- generate the random vector $\mathbf{U}$ by $U_i := \varphi^{-1}\left(\frac{-\ln(Y_i)}{Z}\right)$ for $i = 1, \cdots, d$.

Example 25: Simulation of a Clayton copula with $\theta = 7$ (cf. slide 49):

Here $Z$ has a $\Gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$-distribution. The following graphs show the pairwise simulated copula random vectors $(U_i, U_j)$ with $1 \le i < j \le 4$, from a sample of size 2000:

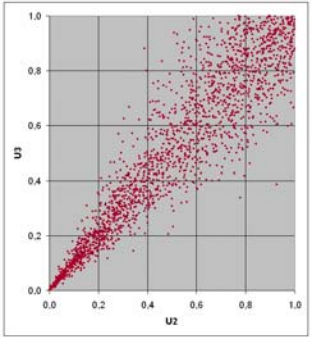


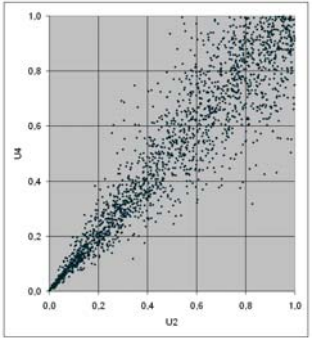



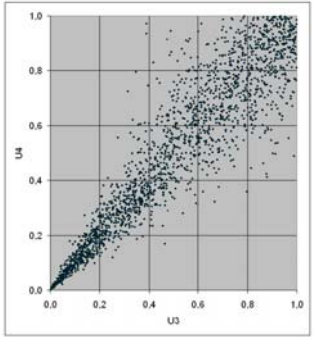



Example 25: Simulation of a Clayton copula with $\theta = 7$ (cf. slide 49):

Here $Z$ has a $\Gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$-distribution. The following graphs show the pairwise simulated copula random vectors $(U_i, U_j)$ with $1 \leq i < j \leq 4$, from a sample of size 2000:



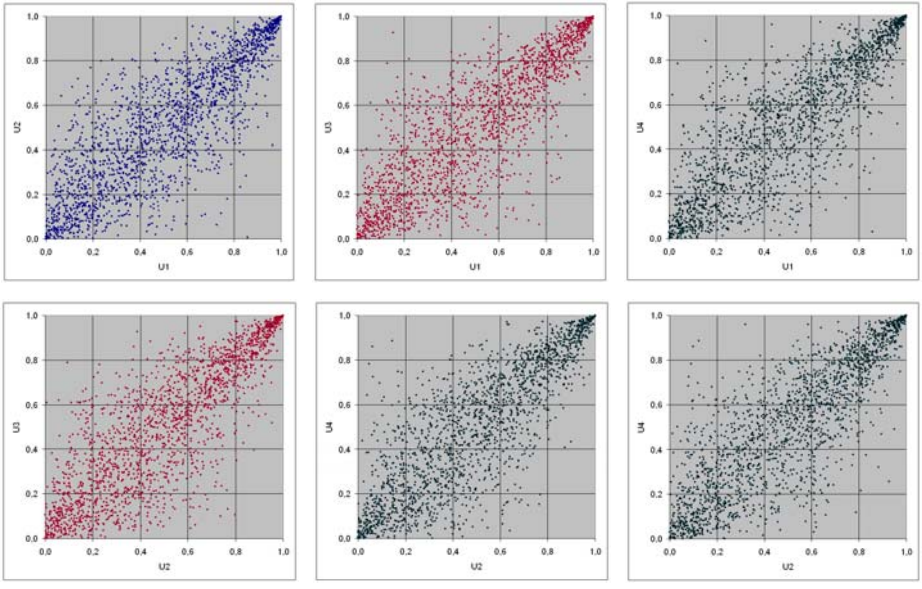
Example 26: Simulation of a Gumbel copula with $\theta = 3$ (cf. slide 49): A simulation of the required random variable $Z$ following a stable distribution with parameter $\alpha = \frac{1}{\theta} < 1$ can be performed using the representation

$$Z = \frac{\sin(\alpha\pi V)}{(\sin(\pi V))^{\theta}} \cdot \left( \frac{\sin((1-\alpha)\pi V)}{-\ln(W)} \right)^{\theta-1}$$

with $V, W$ being independent and continuous uniformly distributed over the interval $[0,1]$ (modified after MAI AND SCHERER (2012), Section 6.13).

The following graphs show the pairwise simulated copula random vectors $(U_i, U_j)$ with $1 \leq i < j \leq 4$, from a sample of size 2000:

Example 26: Simulation of a Gumbel copula with $\theta = 3$ (cf. slide 49):




pairwise simulated copula random vectors $(U_i, U_j)$ with $1 \leq i < j \leq 4$

**Particular methods:** Simulation of bivariate Archimedean copulas with generator $\varphi$:

- Generate independent random variables $U_1, V$ with a continuous uniform distribution over the interval $[0,1]$
- put

$$W := (\varphi')^{-1}\left(\frac{\varphi'(U_1)}{V}\right),\ U_2 := \varphi^{-1}\left(\varphi(W) - \varphi(U_1)\right)$$

(see NELSEN (1999), Exercise 4.16).

Then $(U_1, U_2)$ has the desired copula as cdf.

Example 27: Simulation of a bivariate Clayton copula (cf. slide 49):

Here we have

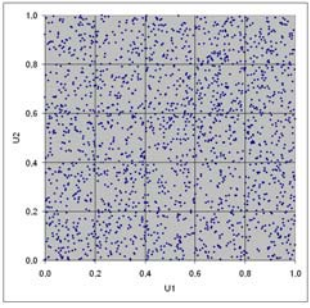
$$\varphi(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right),\ \varphi'(t) = -t^{-(1+\theta)},\ \left(\varphi'\right)^{-1}(s) = (-s)^{-1/(1+\theta)}, \varphi^{-1}(s) = (1+\theta s)^{-1/\theta}$$

giving

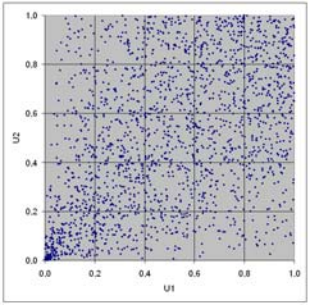
$$W = U_1 \cdot V^{1/(1+\theta)},\ U_2 = \left(1 - U_1^{-\theta} + W^{-\theta}\right)^{-1/\theta}$$

The following graphs show simulated bivariate copula random vectors $(U_1, U_2)$ from a sample of size 2000, with different parameters $\theta > 0$:

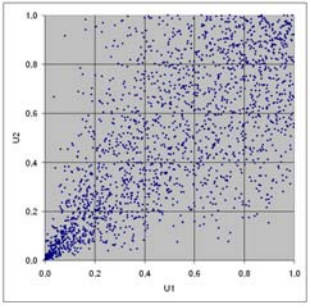
Example 27: Simulation of a bivariate Clayton copula (cf. slide 49):



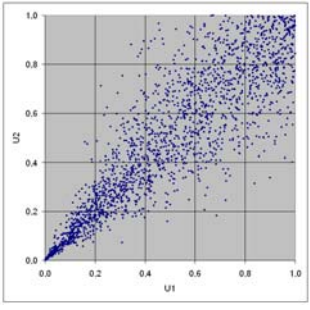

$\theta = 0.1$



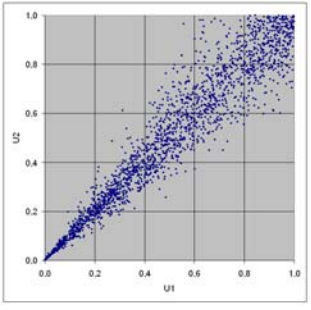

$\theta = 1$



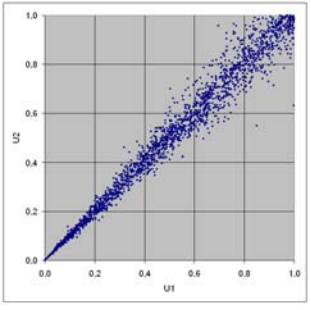

$\theta = 2$




$\theta = 5$




$\theta = 10$




$\theta = 20$

Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with a checkerboard copula:

Let $\mathbf{V} = (V_1, \cdots, V_d)$ be a random vector whose components $V_i$ follow a discrete uniform distribution over the set $T := \{0, 1, \cdots, m-1\}$ with $m \in \mathbb{N}$ for $i = 1, \cdots, d$. Let further denote

$$p_m(k_1, \cdots, k_d) := P\left(\bigcap_{i=1}^{d} \{V_i = k_i\}\right) \text{ for all } (k_1, \cdots, k_d) \in T^d$$

the joint probabilities of $\mathbf{V}$ (forming a $d$-dimensional contingency table) and

$$I_{k_1, \cdots, k_d} := \bigtimes_{j=1}^{d} \left[\frac{k_j}{m}, \frac{k_j + 1}{m}\right] \text{ for } (k_1, \cdots, k_d) \in T^d$$

giving all possible subcubes of $(0,1]^d$ with edge length $1/m$.

Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with a checkerboard copula:

The simulation works in the following steps:

- Convert the contingency table into a matrix with $m^d$ rows and $m+1$ columns, where the first column contains the entries of the contingency table in a suitably denumerated way (so called vectorization), and the remaining columns contain the indices of the attached subcube
- generate a $d$-tuple $(k_1, \cdots, k_d) \in T^d$ as a realization of $\mathbf{V}$ by the inversion method for discrete distributions
- generate a random vector $\mathbf{U}$ independent of $\mathbf{V}$ which follows a continuous uniform distribution over the subcube $I_{k_1,\cdots,k_d} = \bigtimes_{j=1}^{d} \left( \frac{k_j}{m}, \frac{k_j+1}{m} \right]$.

Example 28a: Simulation of a bivariate checkerboard copula with the following contingency table:

| $i \setminus j$ | 0 | 1 | 2 | 3 | |
|---|---|---|---|---|---|
| 3 | 0.00 | 0.02 | 0.05 | 0.18 | 0.25 |
| 2 | 0.06 | 0.05 | 0.10 | 0.04 | 0.25 |
| 1 | 0.07 | 0.10 | 0.06 | 0.02 | 0.25 |
| 0 | 0.12 | 0.08 | 0.04 | 0.01 | 0.25 |
| | 0.25 | 0.25 | 0.25 | 0.25 | |

The numbers in the shaded area are supposed to indicate the indices of the subsquares; note that this notation is consistent with the position of subsquares in the plane, i.e. the pair $(i,j)$ with $i,j \in \{0,1,2,3\}$ denotes the subsquare $I_{ij} = \left(\frac{i}{4}, \frac{i+1}{4}\right] \times \left(\frac{j}{4}, \frac{j+1}{4}\right]$ (this notation differs from the one used earlier, see e.g. slide 51).

Example 28a: Simulation of a bivariate checkerboard copula with the following contingency table:

**Vectorization step:**

| $i \setminus j$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 3 | 0.00 | 0.02 | 0.05 | 0.18 |
| 2 | 0.06 | 0.05 | 0.10 | 0.04 |
| 1 | 0.07 | 0.10 | 0.06 | 0.02 |
| 0 | 0.12 | 0.08 | 0.04 | 0.01 |

contingency table

| $F(k-1)$ | $k$ | $f(k-1)$ | $i$ | $j$ |
|---|---|---|---|---|
| 0.00 | 1 | 0.00 | 0 | 0 |
| 0.12 | 2 | 0.12 | 1 | 0 |
| 0.20 | 3 | 0.08 | 2 | 0 |
| 0.24 | 4 | 0.04 | 3 | 0 |
| 0.25 | 5 | 0.01 | 0 | 1 |
| 0.32 | 6 | 0.07 | 1 | 1 |
| 0.42 | 7 | 0.10 | 2 | 1 |
| 0.48 | 8 | 0.06 | 3 | 1 |
| 0.50 | 9 | 0.02 | 0 | 2 |
| 0.56 | 10 | 0.06 | 1 | 2 |
| 0.61 | 11 | 0.05 | 2 | 2 |
| 0.71 | 12 | 0.10 | 3 | 2 |
| 0.75 | 13 | 0.04 | 0 | 3 |
| 0.75 | 14 | 0.00 | 1 | 3 |
| 0.77 | 15 | 0.02 | 2 | 3 |
| 0.82 | 16 | 0.05 | 3 | 3 |

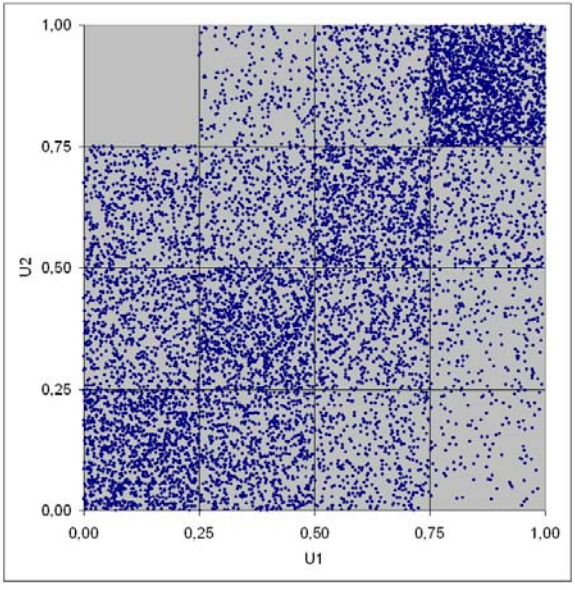
vectorization

Example 28a: Simulation of a bivariate checkerboard copula with the following contingency table:

**Simulation step:**

| $i \setminus j$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 3 | 0.00 | 0.02 | 0.05 | 0.18 |
| 2 | 0.06 | 0.05 | 0.10 | 0.04 |
| 1 | 0.07 | 0.10 | 0.06 | 0.02 |
| 0 | 0.12 | 0.08 | 0.04 | 0.01 |

contingency table



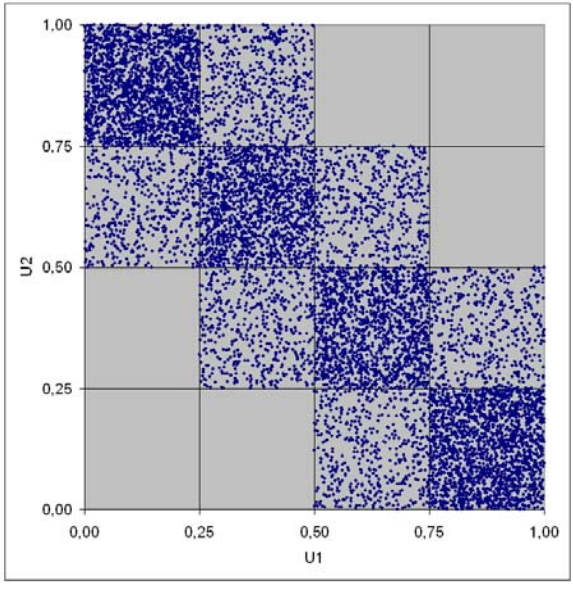

simulated bivariate copula random vectors $(U_1, U_2)$ from a sample of size 10000

Example 28b: Simulation of a bivariate checkerboard copula with the following alternative contingency table:

**Simulation step:**

| $i \setminus j$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 3 | 0.20 | 0.05 | 0.00 | 0.00 |
| 2 | 0.05 | 0.15 | 0.05 | 0.00 |
| 1 | 0.00 | 0.05 | 0.15 | 0.05 |
| 0 | 0.00 | 0.00 | 0.05 | 0.20 |

contingency table

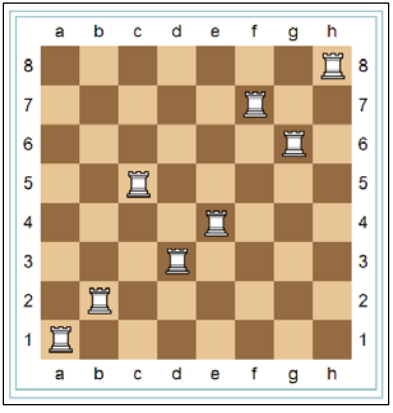



simulated bivariate copula random vectors $(U_1, U_2)$ from a sample of size 10000

**Particular methods:** Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with a rook copula (cf. slide 54):

Let

$$M := \begin{bmatrix} \sigma_{01} & \sigma_{02} & \cdots & \sigma_{0,d-1} & \sigma_{0d} \\ \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,d-1} & \sigma_{1d} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{m-2,1} & \sigma_{m-2,2} & \cdots & \sigma_{m-2,d-1} & \sigma_{m-2,d} \\ \sigma_{m-1,1} & \sigma_{m-1,2} & \cdots & \sigma_{m-1,d-1} & \sigma_{m-1,d} \end{bmatrix}$$



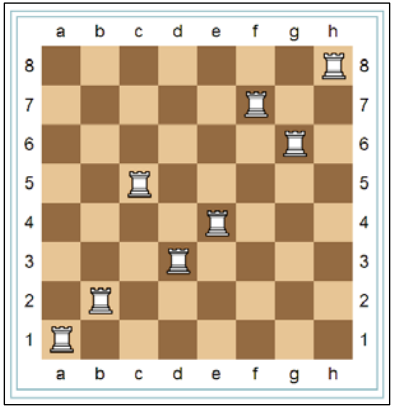

denote a matrix of transposed permutations of the set $T := \{0, 1, \cdots, m-1\}$ for $k = 1, \cdots, d$. For the non-zero probabilities of a rook copula there holds

$$p_m(k_1, \cdots, k_d) = P\left(\bigcap_{i=1}^{d} \{U_i = k_i\}\right) = \frac{1}{m} \Leftrightarrow (k_1, \cdots, k_d) = (\sigma_{t1}, \sigma_{t2}, \cdots, \sigma_{t,d}) \text{ for some } t \in T.$$

**Particular methods:** Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with a rook copula (cf. slide 54):

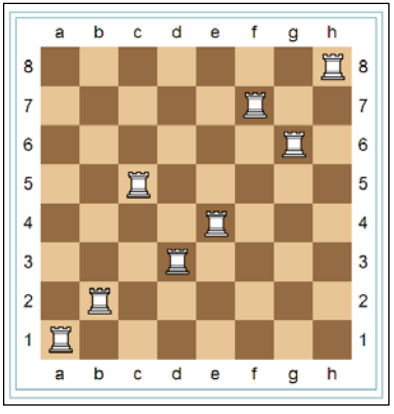
The simulation works in the following steps:

- choose an index $t$ according to a discrete uniform distribution over $T := \{0, 1, \cdots, m-1\}$
- generate $\mathbf{U}$ from a continuous uniform distribution over the subcube $I_{\sigma_{t1}, \sigma_{t2}, \cdots, \sigma_{t,d}}$.



Example 29a (cf. Example 8, slide 55):

Here $M = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 2 & 4 \\ 3 & 2 \\ 4 & 3 \\ 5 & 6 \\ 6 & 5 \\ 7 & 7 \end{bmatrix}$:

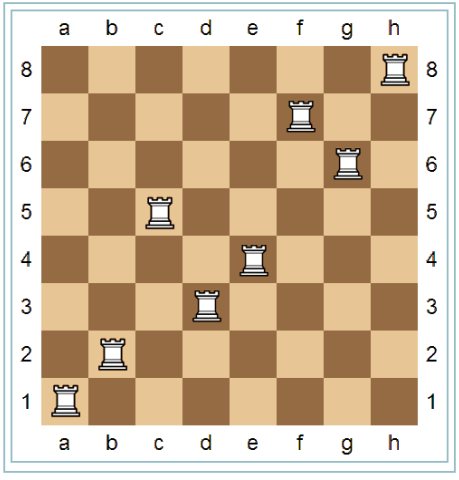


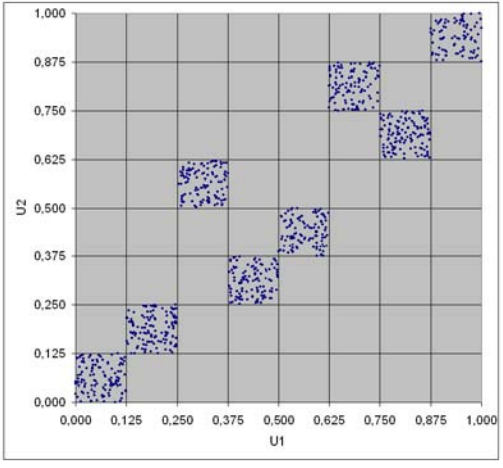

$(k_1, k_2) = (7,7)$

$(k_1, k_2) = (4,3)$

$(k_1, k_2) = (0,0)$

Example 29a (cf. Example 8, slide 55):




simulated rook copula random vectors $(U_1, U_2)$ from a sample of size 1000

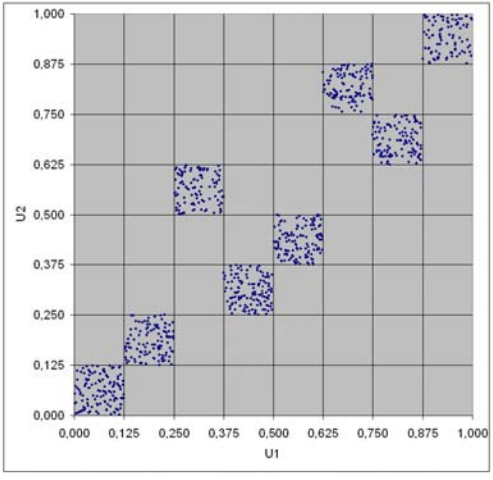
**Particular methods:** Simulation of a $d$-dimensional random vector $\mathbf{U} = (U_1, \cdots, U_d)^T$ with a Bernstein copula derived from a rook copula (cf. slide 56):

Let $M = [\sigma_{ij}]$ denote the underlying matrix of transposed permutations for the rook copula.

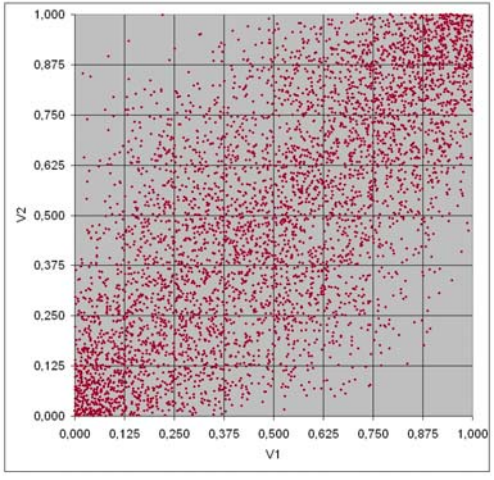
The simulation works in the following steps:

- choose an index $t$ according to a discrete uniform distribution over $T := \{0, 1, \cdots, m-1\}$
- generate $\mathbf{U} = (U_1, \cdots, U_d)$ with independent components where $U_k$ follows a $\mathcal{B}(\sigma_{tk} + 1, m - \sigma_{tk})$ beta distribution for $k = 1, \cdots, d$.

Example 29b (continuation of Example 29a):




simulated rook copula




simulated Bernstein copula

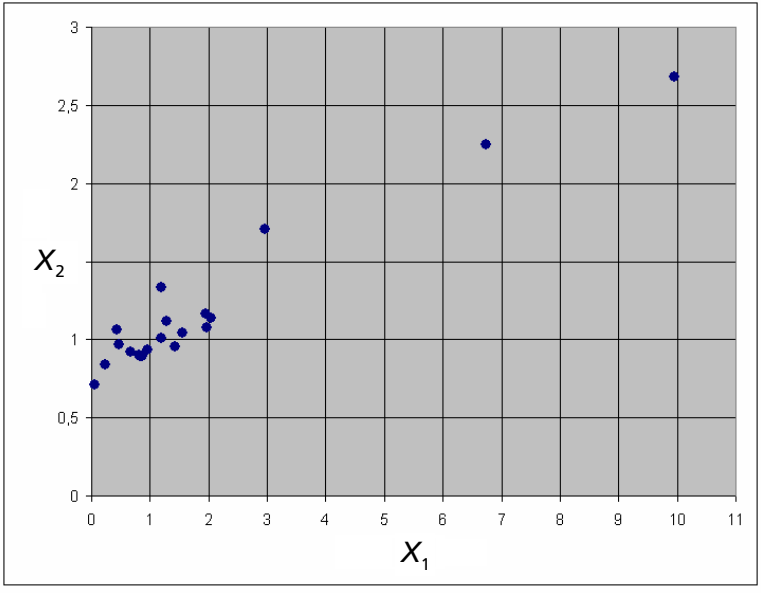
**From data to non-parametric copula models:** General approach for a given matrix $\mathbf{x} = \left[x_{ij}\right]$ of data, where $i = 1, \cdots, n$ is the *i*-th out of *n* (independent) *d*-dimensional obervation row vectors and $j = 1, \cdots, d$ is the corresponding component (dimension) index:

- For each *j*, calculate the rank $r_{ij}$ of the observation $x_{ij}$ among $x_{1j}, \cdots, x_{nj}$ for $i = 1, \cdots, n$
- form the matrix $M := \left[\left(r_{ij} - 1\right)\right]$ of transposed permutations for the empirical rook copula (possibly after a suitable re-ordering of the rows).

Example 30: From data to copula models:

original data:

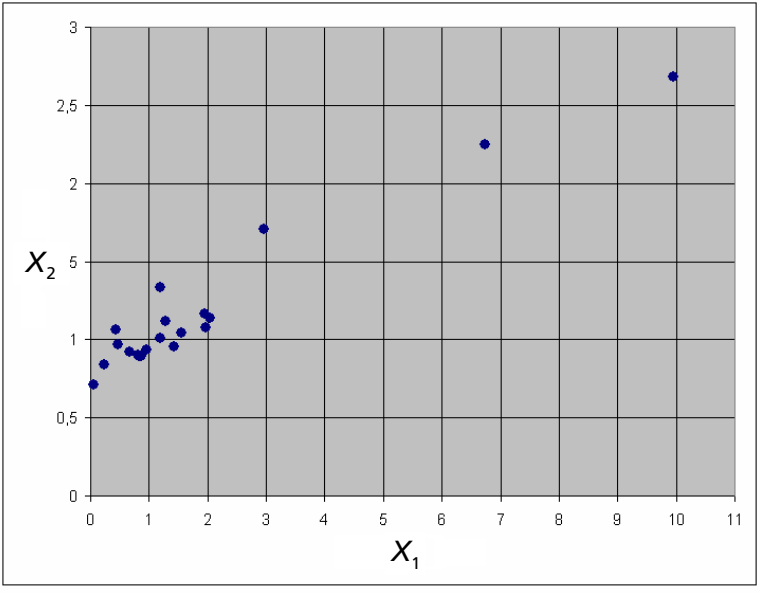
| $i$ | $x_{i1}$ | $x_{i2}$ |
|---|---|---|
| 1 | 0.468 | 0.966 |
| 2 | 9.951 | 2.679 |
| 3 | 0.866 | 0.897 |
| 4 | 6.731 | 2.249 |
| 5 | 1.421 | 0.956 |
| 6 | 2.040 | 1.141 |
| 7 | 2.967 | 1.707 |
| 8 | 1.200 | 1.008 |
| 9 | 0.426 | 1.065 |
| 10 | 1.946 | 1.162 |
| 11 | 0.676 | 0.918 |
| 12 | 1.184 | 1.336 |
| 13 | 0.960 | 0.933 |
| 14 | 1.972 | 1.077 |
| 15 | 1.549 | 1.041 |
| 16 | 0.819 | 0.899 |
| 17 | 0.063 | 0.710 |
| 18 | 1.280 | 1.118 |
| 19 | 0.824 | 0.894 |
| 20 | 0.227 | 0.837 |



Example 30: From data to copula models:

ranks:

| $i$ | $r_{i1}$ | $r_{i2}$ | $x_{i1}$ | $x_{i2}$ |
|---|---|---|---|---|
| 1 | 4 | 9 | 0.468 | 0.966 |
| 2 | 20 | 20 | 9.951 | 2.679 |
| 3 | 8 | 4 | 0.866 | 0.897 |
| 4 | 19 | 19 | 6.731 | 2.249 |
| 5 | 13 | 8 | 1.421 | 0.956 |
| 6 | 17 | 15 | 2.040 | 1.141 |
| 7 | 18 | 18 | 2.967 | 1.707 |
| 8 | 11 | 10 | 1.200 | 1.008 |
| 9 | 3 | 12 | 0.426 | 1.065 |
| 10 | 15 | 16 | 1.946 | 1.162 |
| 11 | 5 | 6 | 0.676 | 0.918 |
| 12 | 10 | 17 | 1.184 | 1.336 |
| 13 | 9 | 7 | 0.960 | 0.933 |
| 14 | 16 | 13 | 1.972 | 1.077 |
| 15 | 14 | 11 | 1.549 | 1.041 |
| 16 | 6 | 5 | 0.819 | 0.899 |
| 17 | 1 | 1 | 0.063 | 0.710 |
| 18 | 12 | 14 | 1.280 | 1.118 |
| 19 | 7 | 3 | 0.824 | 0.894 |
| 20 | 2 | 2 | 0.227 | 0.837 |

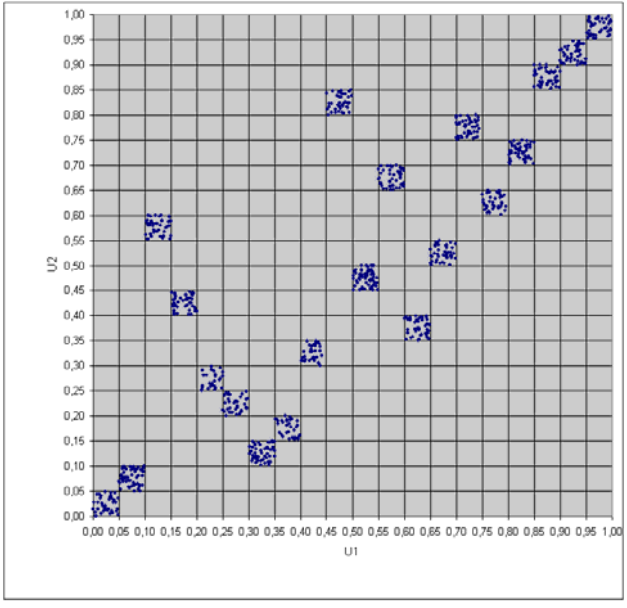


Example 30: From data to copula models:

*M*-matrix:

| $i$ | $M$ | | $x_{i1}$ | $x_{i2}$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.063 | 0.71 |
| 2 | 1 | 1 | 0.227 | 0.837 |
| 3 | 2 | 11 | 0.426 | 1.065 |
| 4 | 3 | 8 | 0.468 | 0.966 |
| 5 | 4 | 5 | 0.676 | 0.918 |
| 6 | 5 | 4 | 0.819 | 0.899 |
| 7 | 6 | 2 | 0.824 | 0.894 |
| 8 | 7 | 3 | 0.866 | 0.897 |
| 9 | 8 | 6 | 0.96 | 0.933 |
| 10 | 9 | 16 | 1.184 | 1.336 |
| 11 | 10 | 9 | 1.2 | 1.008 |
| 12 | 11 | 13 | 1.28 | 1.118 |
| 13 | 12 | 7 | 1.421 | 0.956 |
| 14 | 13 | 10 | 1.549 | 1.041 |
| 15 | 14 | 15 | 1.946 | 1.162 |
| 16 | 15 | 12 | 1.972 | 1.077 |
| 17 | 16 | 14 | 2.04 | 1.141 |
| 18 | 17 | 17 | 2.967 | 1.707 |
| 19 | 18 | 18 | 6.731 | 2.249 |
| 20 | 19 | 19 | 9.951 | 2.679 |

simulated rook copula:

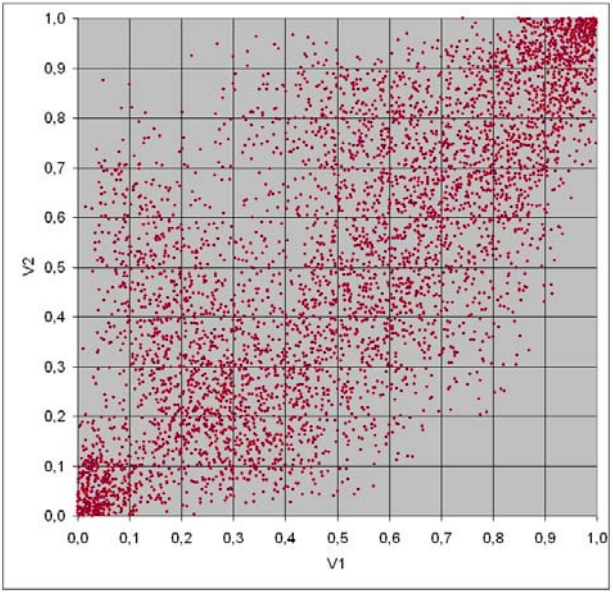


Example 30: From data to copula models:

*M*-matrix:

| $i$ | $M$ | | $x_{i1}$ | $x_{i2}$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.063 | 0.71 |
| 2 | 1 | 1 | 0.227 | 0.837 |
| 3 | 2 | 11 | 0.426 | 1.065 |
| 4 | 3 | 8 | 0.468 | 0.966 |
| 5 | 4 | 5 | 0.676 | 0.918 |
| 6 | 5 | 4 | 0.819 | 0.899 |
| 7 | 6 | 2 | 0.824 | 0.894 |
| 8 | 7 | 3 | 0.866 | 0.897 |
| 9 | 8 | 6 | 0.96 | 0.933 |
| 10 | 9 | 16 | 1.184 | 1.336 |
| 11 | 10 | 9 | 1.2 | 1.008 |
| 12 | 11 | 13 | 1.28 | 1.118 |
| 13 | 12 | 7 | 1.421 | 0.956 |
| 14 | 13 | 10 | 1.549 | 1.041 |
| 15 | 14 | 15 | 1.946 | 1.162 |
| 16 | 15 | 12 | 1.972 | 1.077 |
| 17 | 16 | 14 | 2.04 | 1.141 |
| 18 | 17 | 17 | 2.967 | 1.707 |
| 19 | 18 | 18 | 6.731 | 2.249 |
| 20 | 19 | 19 | 9.951 | 2.679 |

simulated Bernstein copula:



## Simulations of Stochastic Processes of importance in Finance:

Standard brownian motion (or Wiener process) $\{B_t\}_{t\geq 0}$ :

1. $\{B_t\}_{t\geq 0}$ has independent increments $B_{t_k} - B_{t_{k-1}}$ for $0 \leq t_0 < t_1 < t_2 < \cdots < t_n$
2. $\{B_t\}_{t\geq 0}$ has stationary increments, i.e. the distribution of $B_t - B_s$ for $0 \leq s < t$ depends only of $t - s$
3. $\{B_t\}_{t\geq 0}$ has Gaussian increments, i.e. the distribution of $B_t - B_s$ for $0 \leq s < t$ is a $\mathcal{N}(0, t-s)$ normal distribution
4. $\{B_t\}_{t\geq 0}$ has a.s. continuous paths
5. $B_0 = 0$ a.s.

Note that conditions 3 and 4 are *equivalent*, i.e. one of these can be cancelled w.l.o.g.

## Simulations of Stochastic Processes of importance in Finance:

Properties of a standard brownian motion (Wiener process) $\{B_t\}_{t\geq 0}$

- Each $B_t$ follows a $\mathcal{N}(0,t)$ normal distribution for $t>0$
- each path of $\{B_t\}_{t\geq 0}$ is a.s. *nowhere differentiable* and of *unbounded variation* in any finite time interval
- for every increasing sequence $\{t_k\}_{k=1}^n$ of positive time points the random vector $\left(B_{t_1}, B_{t_2}, \cdots, B_{t_n}\right)$ follows a multivariate $\mathcal{N}(\mathbf{0}, \Sigma)$ normal distribution with

  $$\sigma_{ij} = \min\{t_i, t_j\},\ 1\leq i,j\leq n\in\mathbb{N}$$

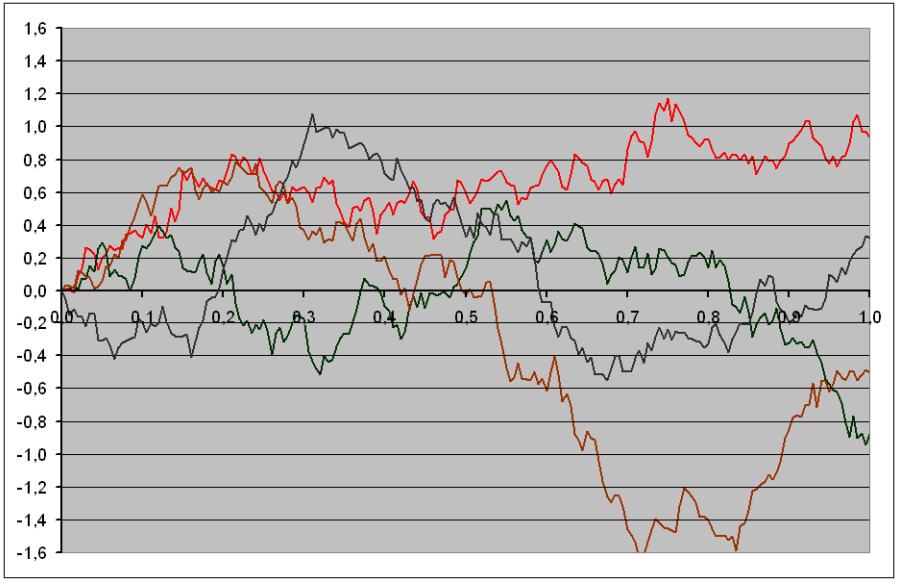
- the process $\left\{\frac{1}{\sqrt{c}}B_{ct}\right\}_{t\geq 0}$ also is a standard brownian motion for every $c>0$ (*fractal self-similarity*).

## Simulations of Stochastic Processes of importance in Finance:

**Paley-Wiener-construction** (1934) of standard brownian motion:

$$B_t = \frac{Y_0}{\sqrt{T}} t + \frac{\sqrt{2T}}{\pi} \sum_{n=1}^{\infty} \frac{\sin\left(\frac{n\pi t}{T}\right)}{n} Y_n, \quad 0 \le t \le T$$

where $\{Y_n\}_{n\in\mathbb{N}}$ is an i.i.d. sequence of standard normally distributed random variables.

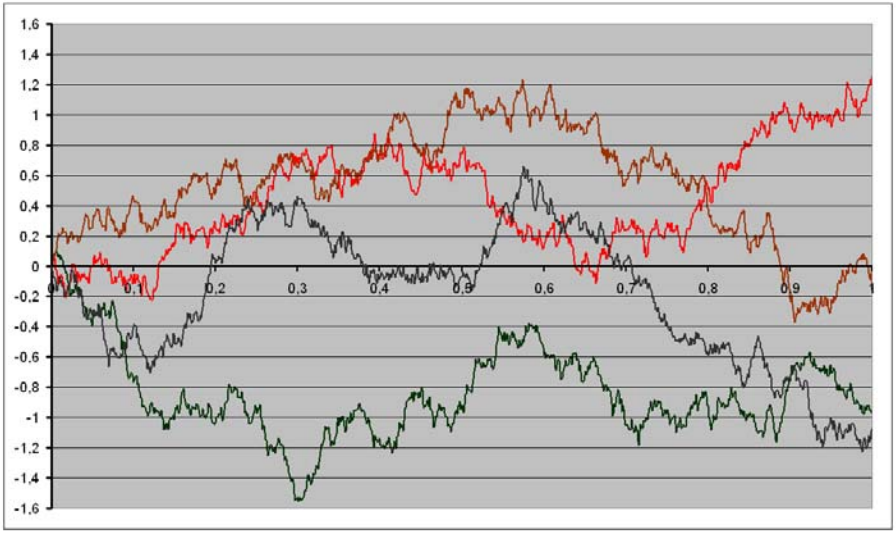



Simulations of an approximate Paley-Wiener-construction for $T = 1$ with 1000 summands und time steps $\Delta t = 0.005$

## Simulations of Stochastic Processes of importance in Finance:

### Alternative approximate simulation:

- Diskretization of time: choose a uniform step size $\Delta = \frac{T}{n} > 0$
- put $B_{k\Delta} = \sum_{i=1}^{k} Y_i$ for $k = 1,2,\cdots,n$ with i.i.d. random variables $Y_i$ following a $\mathcal{N}(0,\Delta)$ distribution (i.e. $\sigma = \sqrt{\Delta}$ !)
- plot the pairs $(0,0)$ and $(k\Delta, B_{k\Delta})$ for $k = 1,2,\cdots,n$.




Approximate simulations for $T = 1$ with $n = 1000$, i.e. $\Delta = 0.001$

## Simulations of Stochastic Processes of importance in Finance:

A general brownian motion (Wiener process) $\left\{B_t^{\mu,\sigma}\right\}_{t\geq 0}$ is obtained from the standard Wiener process $\left\{B_t\right\}_{t\geq 0}$ by a linear transformation:

$$B_t^{\mu,\sigma} = B_0^{\mu,\sigma} + \sigma B_t + \mu t,\ t \geq 0 \qquad (\sigma > 0,\ \mu \in \mathbb{R})$$

with an arbitrary real starting value $B_0^{\mu,\sigma} \in \mathbb{R}$. we call

- $\sigma$ the *diffusion coefficient* or *volatility* (particularly in financial applications)
- $\mu$ the *drift*

of the process.

## Simulations of Stochastic Processes of importance in Finance:

Properties of a general brownian motion (Wiener process):

- Each $B_t^{\mu,\sigma}$ follows a $\mathcal{N}(B_0^{\mu,\sigma} + \mu t, \sigma^2 t)$ normal distribution for $t > 0$
- each path of $\left\{B_t^{\mu,\sigma}\right\}_{t\geq 0}$ is a.s. *nowhere differentiable* and of *unbounded variation* in any finite time interval
- for every increasing sequence $\left\{t_k\right\}_{k=1}^n$ of positive time points the random vector $\left(B_{t_1}^{\mu,\sigma}, B_{t_2}^{\mu,\sigma}, \cdots, B_{t_n}^{\mu,\sigma}\right)$ follows a multivariate $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ normal distribution with

  $\boldsymbol{\mu} = \left(B_0^{\mu,\sigma} + t_1\mu, \cdots, B_0^{\mu,\sigma} + t_n\mu\right)^T$ and

  $\sigma_{ij} = \sigma^2 \min\left\{t_i, t_j\right\},\ 1 \leq i, j \leq n \in \mathbb{N}.$

## Simulations of Stochastic Processes of importance in Finance:

*Stochastic differential equations* (SDE's) are formally given by

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t$$

with a standard brownian motion $\{B_t\}_{t\geq 0}$, suitable (generally continuous) functions $\mu, \sigma : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ and a known starting value $X_0 \in \mathbb{R}$. Strictly speaking, a SDE is not a differential equation, buth rather a (Itô-)integral equation of the form

$$X_t = \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s.$$

## Simulations of Stochastic Processes of importance in Finance:

Simulation of a solution of a SDE:

Heuristically, we have

$$\int_t^{t+h} \sigma(s, X_s)\, dB_s \approx \sigma(t, X_t) \cdot (B_{t+h} - B_t) = \sigma(t, X_t) \cdot \sqrt{h} \cdot Z \quad \text{for small } h > 0,$$

where $Z$ is standard-normally distributed (remember that $B_{t+h} - B_t$ follows a $\mathcal{N}(0, h)$ distribution with a standand deviation of $\sqrt{h}$).

## Simulations of Stochastic Processes of importance in Finance:

Simulation of a solution of a SDE:

- Divide $[0,T]$ into $n$ equidistant subintervals $[t_{k-1}, t_k]$ with $t_k = k\Delta$ and $\Delta = \frac{T}{n},\ k = 0, \cdots n$

- define recursivlely

$$X_{t_{k+1}} - X_{t_k} = \Delta \cdot \mu\left(t_k, X_{t_k}\right) + \sqrt{\Delta} \cdot \sigma\left(t_k, X_{t_k}\right) Z_{k+1},\ k = 0, 1, \cdots, n-1$$

  with i.i.d. standard normally distributed $Z_1, \cdots, Z_n$

- plot the pairs $(0, X_0)$ and $(k\Delta, X_{k\Delta})$ for $k = 1, 2, \cdots, n$.

## Simulations of Stochastic Processes of importance in Finance:

A particular SDE is given by

$$dX_t = \mu X_t\, dt + \sigma X_t\, dB_t$$

with constants $\mu \in \mathbb{R}$, $\sigma > 0$. The solution with the Itô-integral is different from intuition, namely

$$X_t = X_0 \cdot \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right),\ t \geq 0.$$

This stochastic process is also called *geometric brownian motion* (→ Black-Scholes-model for the stock market).

## Simulations of Stochastic Processes of importance in Finance:

**Short rate models** deal with stochastic changes of interest rates on financial markets on the basis of a random function $r(t)$, $0 \leq t \leq T$ which describes the growth of a bank account with initial capital $K_0$ according to the formula

$$K_t = \exp\left( \int_0^t r(u)\,du \right) \quad \text{für } 0 \leq t \leq T.$$

## Simulations of Stochastic Processes of importance in Finance:

Popular short rate models are:

- The *Vasiček model*:

$$dr(t) = a\big[b - r(t)\big]dt + \sigma dB_t$$

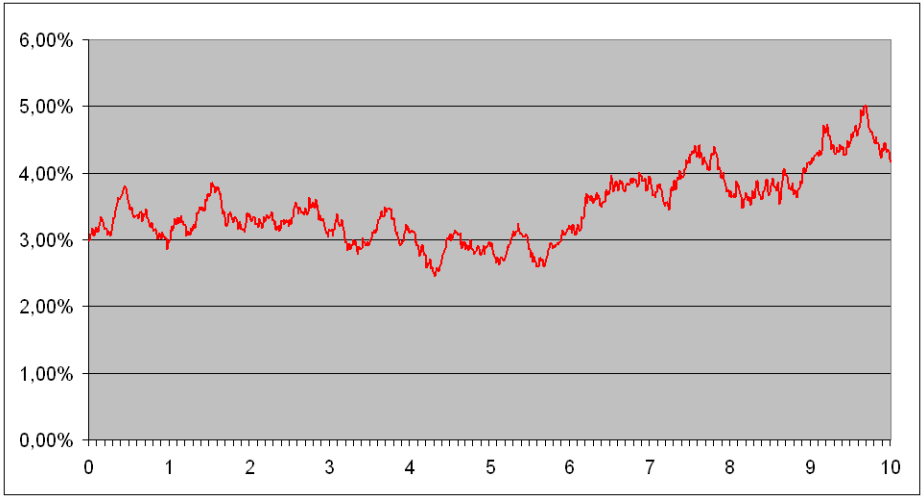
- The *Cox-Ingersoll-Ross model*:

$$dr(t) = a\big[b - r(t)\big]dt + \sigma\sqrt{r(t)}\, dB_t$$

- The *Black-Karasinski model*:

$$d\ln r(t) = a\big[b - \ln r(t)\big]dt + \sigma dB_t$$

The Black-Karasinski model is an exponential Vasiček-Modell, i.e. $\ln r(t)$ is the solution of a usual Vasiček model.




Simulation of a Black-Karasinski model over a time period of 10 years with $a = 0.20$ $b = -3.02$ $\sigma = 0.19$ and an initial interest rate of 3% (formerly suggested as a model for interest rate risk in Solvency II)

# V. Principles of Capital Allocation

The total risk capital $RC$ of a company (bank or insurance company) is in general a function of the *aggregated risk*

$$RC = RC(X),$$

where $X$ is additively composed of individual risks of $n \in \mathbb{N}$ sub-entities (e.g., lines of business):

$$X = \sum_{i=1}^{n} X_i.$$

For what follows we shall assume that all risks are evaluated according to a common risk measure $R$, for instance Value at Risk or Expected Shortfall.

The quantities

$$RC(X_i),\ i = 1, \cdots, n$$

are called *stand-alone risk capital*.

A Capital Allocation Principle is a method to attribute (allocate) the total risk $R\left(\sum_{i=1}^{n} X_i\right)$ to the individual sub-entites in an economically "reasonable" way.

Remember that in case that the risk measure is *coherent* (for the class of underlying risk distributions), we have

$$R\left(\sum_{i=1}^{n} X_i\right) \leq \sum_{i=1}^{n} R(X_i),$$

and the resulting difference

$$D_R(X_1, \cdots, X_n) = \sum_{i=1}^{n} R(X_i) - R\left(\sum_{i=1}^{n} X_i\right) \geq 0$$

is called *diversification gap*. This implies in particular that the allocated capital must be less than the value of its risk measure for *at least one* sub-entity if the diversification gap is positive.

According to *Denault*, a capital allocation principle RC is called *coherent*, if it satisfies the following four conditions:

$$\sum_{i=1}^{n} RC(X_i) = R\left(\sum_{i=1}^{n} X_i\right)$$ [total allocation]

$$\sum_{i\in J} RC(X_i) \le R\left(\sum_{i\in J} X_i\right) \text{ for all subsets } J \subseteq \{1,\cdots,n\}$$ [no undercut]

$$R\left(X_i + \sum_{k\in M_{ij}} X_k\right) - R\left(\sum_{k\in M_{ij}} X_k\right) = R\left(X_j + \sum_{k\in M_{ij}} X_k\right) - R\left(\sum_{k\in M_{ij}} X_k\right) \text{ for all}$$

$$M_{ij} \subseteq \{1,\cdots,n\}\setminus\{i,j\} \quad \Rightarrow \quad RC(X_i) = RC(X_j) \quad \text{for all } 1 \le i < j \le n$$ [symmetry]

$$RC(c) = R(c) = c \text{ for constant risks } c \in \mathbb{R}$$ [riskless allocation]

**Remark 1:** The assumption of a ***total allocation*** implies that

$$\sum_{i=1}^{n} RC(X_i) = R\left(\sum_{i=1}^{n} X_i\right) = RC\left(\sum_{i=1}^{n} X_i\right)$$

(for the equality on the r.h.s., consider the case $n = 1$).

This means that the part that is allocated to the $k$-th sub-entity $x_k$ (allocation factor) can be expressed as

$$x_k = \frac{RC(X_k)}{RC\left(\sum_{i=1}^{n} X_i\right)} = \frac{RC(X_k)}{R\left(\sum_{i=1}^{n} X_i\right)} \quad \text{with} \quad \sum_{k=1}^{n} x_k = 1.$$

**Remark 2:** The condition of a ***riskless allocation*** is not crucial, because constant risks can w.l.o.g. be excluded from the portfolio in advance.

**Remark 3:** The property of ***symmetry*** is obviously equivalent to the condition

$$R\left(X_i + \sum_{k \in M_{ij}} X_k\right) = R\left(X_j + \sum_{k \in M_{ij}} X_k\right) \text{ for all } M_{ij} \subseteq \{1, \cdots, n\} \setminus \{i, j\} \Rightarrow$$

$$RC(X_i) = RC(X_j) \quad \text{for all } 1 \leq i < j \leq n.$$

**Remark 4:** The condition of ***no undercut*** is in most cases difficult to achieve in general. If it is fulfilled only for $J = \{i\}$ with $i \in N$, the we speak of the ***individual no undercut*** condition.

**Remark 5:** A drawback of most of the proposed allocation concepts in practice is that they do not pay attention to the (individual or collective) default probabilities

$$P\big(X_i > RC(X_i)\big) \text{ for } i = 1,\cdots,n \quad \text{or} \quad P\left(\bigcup_{i=1}^{n}\{X_i > RC(X_i)\}\right)$$

that risk $X_i$ or at least one of the $X_i$ exceeds its allocated capital. This is in particular important when VaR or ES is the underlying risk measure.

## Commonly used allocation principles:

### I. Stand-Alone proportional allocation:

The allocation factors are here given by

$$x_k = \frac{R(X_k)}{\sum_{i=1}^{n} R(X_i)} \quad \text{or} \quad RC(X_k) = x_k \cdot RC\left(\sum_{i=1}^{n} X_i\right) = R(X_k) \cdot \frac{R\left(\sum_{i=1}^{n} X_i\right)}{\sum_{i=1}^{n} R(X_i)} \quad \text{for } k = 1, \cdots, n.$$

This principle fulfils the conditions of

- *total allocation*
- *symmetry*
- *individual no undercut*.

It does in general not fulfil the conditions of *no undercut* or *riskless allocation*.

## Commonly used allocation principles:

### II. Covariance-based allocation:

In case that the individual risks possess second moments we have

$$Var(X) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov\left(X_i, X_j\right) = \sum_{i=1}^{n} Cov\left(X_i, \sum_{j=1}^{n} X_j\right) = \sum_{i=1}^{n} Cov\left(X_i, X\right).$$

The allocation factors are here given by

$$x_k = \frac{Cov\left(X_k, X\right)}{Var(X)} \quad \text{or} \quad RC\left(X_k\right) = x_k \cdot RC\left(\sum_{i=1}^{n} X_i\right) = \frac{Cov\left(X_k, X\right)}{Var(X)} \cdot R(X) \quad \text{for } k = 1, \cdots, n.$$

In the case of *stochastic independence* this simplifies to

$$x_k = \frac{Cov\left(X_k, X\right)}{Var(X)} = \frac{Var\left(X_k\right)}{Var(X)} \quad \text{for } k = 1, \cdots, n.$$

## Commonly used allocation principles:

### II. Covariance-based allocation:

The allocation factors are here given by

$$x_k = \frac{Cov(X_k, X)}{Var(X)} \quad \text{or} \quad RC(X_k) = x_k \cdot RC\left(\sum_{i=1}^{n} X_i\right) = \frac{Cov(X_k, X)}{Var(X)} \cdot R(X) \quad \text{for } k = 1, \cdots, n.$$

This principle fulfils in general only the condition of

- *total allocation*.

It does in general not fulfil the conditions of *symmetry* or *no undercut* or *riskless allocation*.

## Commonly used allocation principles:

### II. Covariance-based allocation:

**Counterexample:** Let $n=3$ and assume that the underlying risk measure is *expectation*. Assume further that the risks are independent and that $E(X_i)=6$ and $Var(X_i)=i$ for $i=1,2,3$. Then $x_i=\frac{Var(X_i)}{Var(X)}=\frac{i}{6}$ for $i=1,2,3$, i.e. all risk allocations are different. But trivially, for $1\le i<j\le 3$ and $M_{ij}\subseteq\{1,\cdots,n\}\setminus\{i,j\}$, we have

$$R\left(X_i+\sum_{k\in M_{ij}}X_k\right)-R\left(\sum_{k\in M_{ij}}X_k\right)=E\left(X_i+\sum_{k\in M_{ij}}X_k\right)-E\left(\sum_{k\in M_{ij}}X_k\right)=E(X_i)$$

$$=E(X_j)=E\left(X_j+\sum_{k\in M_{ij}}X_k\right)-E\left(\sum_{k\in M_{ij}}X_k\right)=R\left(X_j+\sum_{k\in M_{ij}}X_k\right)-R\left(\sum_{k\in M_{ij}}X_k\right).$$

## Commonly used allocation principles:

### II. Covariance-based allocation:

**Remark:** It is potentially possible that some of the individual risks have a *negative correlation* with the total risk. This would lead to a *negative allocation* factor which might not be desirable. In order to avoid such cases, the covariance-based allocation is sometimes modified in the following way:

$$RC(X_k) = E(X_k) + x_k\big(R(X) - E(X)\big) \quad \text{with} \quad x_k = \frac{Cov(X_k, X)}{Var(X)} \quad \text{for } k = 1, \cdots, n.$$

Note that this allocation principle also fulfils the condition of *total allocation*, but does not generally avoid *negative* capital allocations.

Example 31: Let the joint distribution of the risks $X$ and $Y$ be given by the following table (where $\alpha = 0.005$ (Solvency II standard); cf. also Example 1):

| $P(X=x, Y=y)$ | | $x$ | | | | |
|---|---|---|---|---|---|---|
| | | 0 | 50 | 100 | $P(Y=y)$ | $P(Y \le y)$ |
| $y$ | 0 | $\beta$ | $0.441-\beta$ | 0.000 | 0.441 | 0.441 |
| | 40 | $0.553-\beta$ | $0.001+\beta$ | 0.000 | 0.554 | **0.995** |
| | 50 | 0.000 | 0.000 | 0.005 | 0.005 | 1.000 |
| | $P(X=x)$ | 0.553 | 0.442 | 0.005 | | |
| | $P(X \le x)$ | 0.553 | **0.995** | 1.000 | | |

with $0 \le \beta \le 0.441$, giving $\mathrm{VaR}_\alpha(X) = 50$, $\mathrm{VaR}_\alpha(Y) = 40$.

The range of possible correlations is given by [–0.9484; 0.7921], independence is obtained for $\beta = \beta_0 := 0.239733$.

The following table shows the distribution of the aggregated risk $S = X + Y$

| $s$ | 0 | 40 | 50 | 90 | 100 | 140 | 150 |
|---|---|---|---|---|---|---|---|
| $P(S = s)$ | $\beta$ | $0.553 - \beta$ | $0.441 - \beta$ | $0.001 + \beta$ | 0.000 | 0.000 | 0.005 |
| $P(S \leq s)$ | $\beta$ | 0.553 | $0.994 - \beta$ | **0.995** | 0.995 | 0.995 | 1.000 |

with

$$\mathrm{VaR}_\alpha(S) = 90 = \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y),$$

independent of the parameter $\beta$ and hence also independent of the possible correlations between $X$ and $Y$.

## Stand-Alone proportional allocation:

$$x_1 = \frac{\mathrm{VaR}_\alpha(X)}{\mathrm{VaR}_\alpha(X)+\mathrm{VaR}_\alpha(Y)} = \frac{5}{9} = 0.\bar{5},\ x_2 = \frac{\mathrm{VaR}_\alpha(Y)}{\mathrm{VaR}_\alpha(X)+\mathrm{VaR}_\alpha(Y)} = \frac{4}{9} = 0.\bar{4},$$

$$RC(X) = x_1 \cdot RC(X+Y) = \frac{5}{9}\cdot 90 = 50,\ RC(Y) = x_2 \cdot RC(X+Y) = \frac{4}{9}\cdot 90 = 40,$$

i.e. the stand-alone-proportional allocation attributes exactly the individual risk measures to the risks (due to the fact that there is a non-strict diversification effect in the portfolio), independent of the joint correlation.

## Covariance-based allocation:

$$x_1(\beta) = \frac{Cov(X, X+Y)}{Var(X+Y)} = \frac{Var(X)+Cov(X,Y)}{Var(X)+Var(Y)+2\cdot Cov(X,Y)} = \frac{\beta + 0.082387}{2\beta + 0.04099995}$$

$$x_2(\beta) = \frac{Cov(Y, X+Y)}{Var(X+Y)} = \frac{Var(Y)+Cov(X,Y)}{Var(X)+Var(Y)+2\cdot Cov(X,Y)} = \frac{\beta - 0.04138705}{2\beta + 0.04099995}$$

$$RC(X,\beta) = x_1(\beta)\cdot RC(X+Y) = x_1(\beta)\cdot 90, \;\; RC(Y,\beta) = x_2(\beta)\cdot RC(X+Y) = x_2(\beta)\cdot 90$$

For the extreme choices of $\beta$, we obtain:

$x_1(0) = 2.009441475$ $\qquad$ $x_2(0) = -1.009441475$

$x_1(0.441) = 0.5670498682$ $\qquad$ $x_2(0.441) = 0.4329501318$

$x_1(\beta_0) = 0.6189069621$ $\qquad$ $x_2(\beta_0) = 0.3810930379$

## Covariance-based allocation:

Note that the allocation factor $x_2(\beta)$ remains negative as along as

$\beta < \beta^- := 0.04138705.$

Especially, we obtain

$RC(X, \beta^-) = 90$ $\quad$ $RC(Y, \beta^-) = 0$

$RC(X, \beta_0) = 55.702$ $\quad$ $RC(Y, \beta_0) = 34.298$

$RC(X, 0.441) = 51.034$ $\quad$ $RC(Y, 0.441) = 38.966.$

In all three cases, the Lob connected to the risk $X$ is allocated more money than necessary, i.e. the condition of *no undercut* is violated.

## Covariance-based allocation:

For the modified covariance-based allocation, we obtain instead

$$RC_{\text{mod}}(X) = E(X) + x_1(\beta)\big(\text{VaR}_\alpha(X) - E(X)\big)$$
$$RC_{\text{mod}}(Y) = E(Y) + x_2(\beta)\big(\text{VaR}_\alpha(Y) - E(Y)\big)$$

with

$$RC_{\text{mod}}(X,0) = 113.005 \qquad RC_{\text{mod}}(Y,0) = -23.005 \text{ (!)}$$

$$RC_{\text{mod}}(X,\beta^-) = 67.59 \qquad RC_{\text{mod}}(Y,\beta^-) = 22.41$$

$$RC_{\text{mod}}(X,\beta_0) = 50.445 \qquad RC_{\text{mod}}(Y,\beta_0) = 39.555$$

$$RC_{\text{mod}}(X,0.441) = 48.112 \qquad RC_{\text{mod}}(Y,0.441) = 41.888.$$

Also here, the condition of *no undercut* is violated in all four cases.

## Commonly used allocation principles:

### III. The Shapley value allocation:

The mathematical foundations of the Shapley value go back to the theory of *coalition games*. There one considers a set $N = \{1, \cdots, n\}$ of „players“, who can form coalitions $K \subseteq N$, which are rated by a “value” that is given by a suitable mapping

$$W : \mathfrak{P}_0(N) := \mathfrak{P}(N) \setminus \{\varnothing\} \to \mathbb{R}.$$

Typically $W$ is assumed to be *super-additive*, i.e.

$$W(K_1 \oplus K_2) \geq W(K_1) + W(K_2) \qquad (*)$$

for disjoint coalitions $K_1, K_2 \subseteq N$.

## Commonly used allocation principles:

### III. The Shapley value allocation:

This condition enforces the players to form coalitions (rather than playing alone), because by induction, we get from (*)

$$W\left(\bigoplus_{i=1}^{m} K_i\right) \geq \sum_{i=1}^{m} W(K_i)$$

for pairwise disjoint coalitions $K_1, \cdots, K_m \subseteq N$. As a special case we obtain

$$W(N) \geq \sum_{i=1}^{n} W(\{i\}).$$

Example 32: Majority game:

$$W(K) = \begin{cases} 1, & \#K > \dfrac{\#N}{2} \\ 0, & \#K \leq \dfrac{\#N}{2} \end{cases} \quad \text{for } K \subseteq N;$$

note that for two arbitrary disjoint coalitions $K_1, K_2 \subseteq N$ we cannot have $W(K_1) = W(K_2) = 1$ simultaneously, otherwise $\#(K_1 \oplus K_2) > \#N$, a contradiction to $K_1 \oplus K_2 \subseteq N$. The inequality (*) hence is fulfilled, if $W(K_1 \oplus K_2) = 1$; for $W(K_1 \oplus K_2) = 0$, (*) is always fulfilled.

## Commonly used allocation principles:

### III. The Shapley value allocation:

In the sequel we denote with $\mathcal{W}$ the set of all functions defined on $\mathfrak{P}_0(N)$ which fulfil the condition (*). Obviously, $\mathcal{W}$ is a *convex cone*, i.e.

$\gamma W \in \mathcal{W}$ for all $W \in \mathcal{W}$ and $\gamma \geq 0$ and

$\alpha W_1 + (1-\alpha) W_2 \in \mathcal{W}$ for all $W_1, W_2 \in \mathcal{W}$ and $0 \leq \alpha \leq 1$.

We call a coalition $K \subseteq N$ *support* of $W \in \mathcal{W}$, if

$W(H) = W(K \cap H)$ for all $H \in \mathfrak{P}_0(N)$.

Further, for any permutation $\sigma \in \Sigma(N)$, we set

$W_\sigma(K) := W\left(\sigma^{-1}(K)\right)$, $K \in \mathfrak{P}_0(N)$.

## Commonly used allocation principles:

### III. The Shapley value allocation:

**Definition:** A mapping $\varphi : \mathcal{W} \to \mathbb{R}^n$ is called *Shapley value (function)*, if

- $\sum_{i \in K} \varphi_i(W) = W(K)$ for every support $K$ of $W \in \mathcal{W}$
- $\varphi_{\sigma(i)}(W) = \varphi_i\left(W_\sigma\right)$ for all permutations $\sigma \in \Sigma(N)$ and all $i \in N$
- $\varphi_i\left(W_1 + W_2\right) = \varphi_i\left(W_1\right) + \varphi_i\left(W_2\right)$ for all $W_1, W_2 \in \mathcal{W}$.

It can be proven that these conditions specify the mapping $\varphi$ uniquely, with

$$\varphi_i(W) = \frac{1}{n} \sum_{i \in H \in \mathfrak{P}_0(N)} \frac{W(H) - W\left(H \setminus \{i\}\right)}{\binom{n-1}{\#H - 1}} \quad \text{for all } W \in \mathcal{W} \text{ and } i \in N.$$

**Commonly used allocation principles:**

III. The Shapley value allocation:

**Interpretation:** $\varphi_i(W)$ corresponds to an average increment of the value of a coalition $H \in \mathcal{W}$ that player *i* contributes through joining the coalition.

Note that for every player *i*, there exist exactly *n* coalitions of varying size *h* from 1 (solo player) to n (full coalition) that player *i* can belong to, and that $\binom{n-1}{h-1}$ is the number of possible choices for coalitions with *h* members (including player *i*).

## Commonly used allocation principles:

### III. The Shapley value allocation:

**Interpretation:** This idea can be used to make the notion of an "average increment of the value of a coalition" more rigorous:

Define, for every $i \in N$, a random set $\mathcal{H}_i$ in two steps:

1. Choose a number $h \in N$ according to a Laplace distribution over $N$.

2. Choose a set $\mathcal{K} \subseteq N \setminus \{i\}$ with exactly $h-1$ elements at random (Laplace distribution over the set of all subsets of $N \setminus \{i\}$ with $\binom{n-1}{h-1}$ elements).

3. Let $\mathcal{H}_i := \mathcal{K} \oplus \{i\}$.

Then $\varphi_i(W) = E\left[W(\mathcal{H}_i) - W\left(\mathcal{H}_i \setminus \{i\}\right)\right]$.

## Commonly used allocation principles:

### III. The Shapley value allocation:

### Application to the risk capital allocation problem:

Let

$$W(K) := -R\left(\sum_{i\in K} X_i\right), \quad K \subseteq N,$$

then $W \in \mathcal{W}$ for every coherent risk measure $R$ because of

$$W(K_1 \oplus K_2) = -R\left(\sum_{i\in K_1\oplus K_2} X_i\right) = -R\left(\sum_{i\in K_1} X_i + \sum_{i\in K_2} X_i\right) \geq -R\left(\sum_{i\in K_1} X_i\right) - R\left(\sum_{i\in K_2} X_i\right) = W(K_1) + W(K_2).$$

## Commonly used allocation principles:

III. The Shapley value allocation:

**Application to the risk capital allocation problem:**

**Corollary:** For a coherent risk measure $R$,

$$RC_{\varphi}(X_i) := \frac{1}{n} \sum_{i \in H \in \mathfrak{P}_0(N)} \frac{R\left(\sum_{k \in H} X_k\right) - R\left(\sum_{k \in H \setminus \{i\}} X_k\right)}{\binom{n-1}{\#H-1}} = E\left[R\left(\sum_{k \in \mathcal{H}_i} X_k\right) - R\left(\sum_{k \in \mathcal{H}_i \setminus \{i\}} X_k\right)\right], \; i \in N$$

defines a capital allocation principle, called *Shapley value principle*.

## Commonly used allocation principles:

### III. The Shapley value allocation:

### Application to the risk capital allocation problem:

**Theorem:** For a coherent risk measure $R$, the Shapley value allocation principle fulfils Denault's conditions of *total allocation*, *individual no undercut*, *symmetry* and *riskless allocation*. In particular, if $R$ is *additive* for a risk vector $(X_1,\cdots,X_n)$, then

$$RC_\varphi(X_i) = R(X_i) \text{ for } i = 1,\cdots,n.$$

The Shapley value allocation also fulfils Denault's general condition of *no undercut* (and is hence coherent), if additionally there holds

$$\sum_{J\subseteq I}(-1)^{(\#I-\#J)}R\left(\sum_{j\in J}X_j\right)\leq 0 \text{ for all } I\subseteq N. \qquad (**)$$

## Commonly used allocation principles:

### III. The Shapley value allocation:

### Application to the risk capital allocation problem:

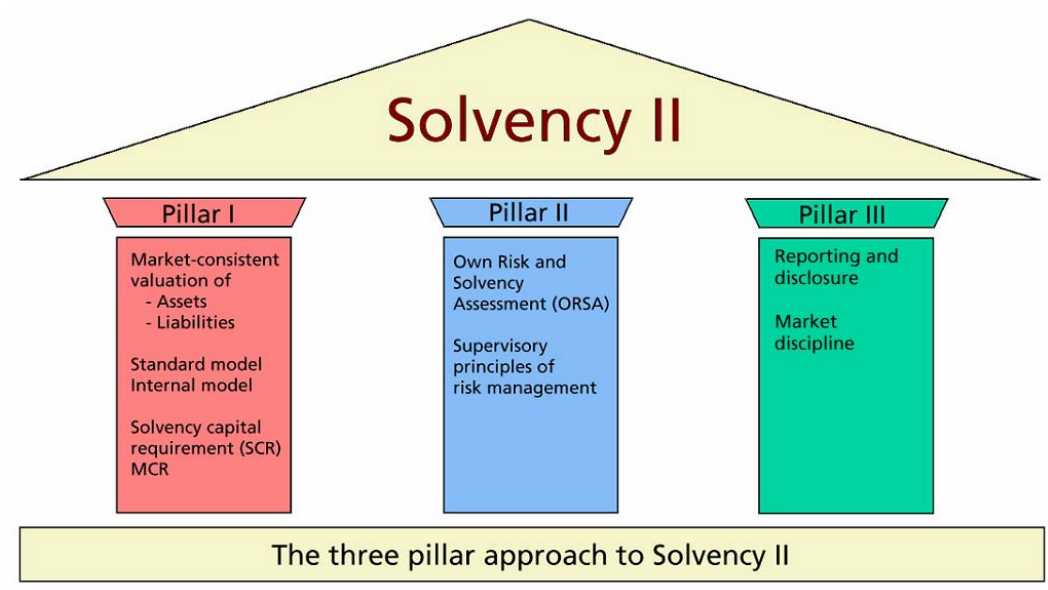
Note that the last condition (**) is always fulfilled for any subset $I = \{i_1, i_2\} \subseteq N$ with two elements, since then

$$\sum_{J \subseteq I} (-1)^{(\#I - \#J)} R\left(\sum_{j \in J} X_j\right) = R\left(X_{i_1} + X_{i_2}\right) - R\left(X_{i_1}\right) - R\left(X_{i_2}\right) \leq 0$$

by the coherence property of *R*. Condition (**) is, however, not a consequence of the coherence property of *R* in general.

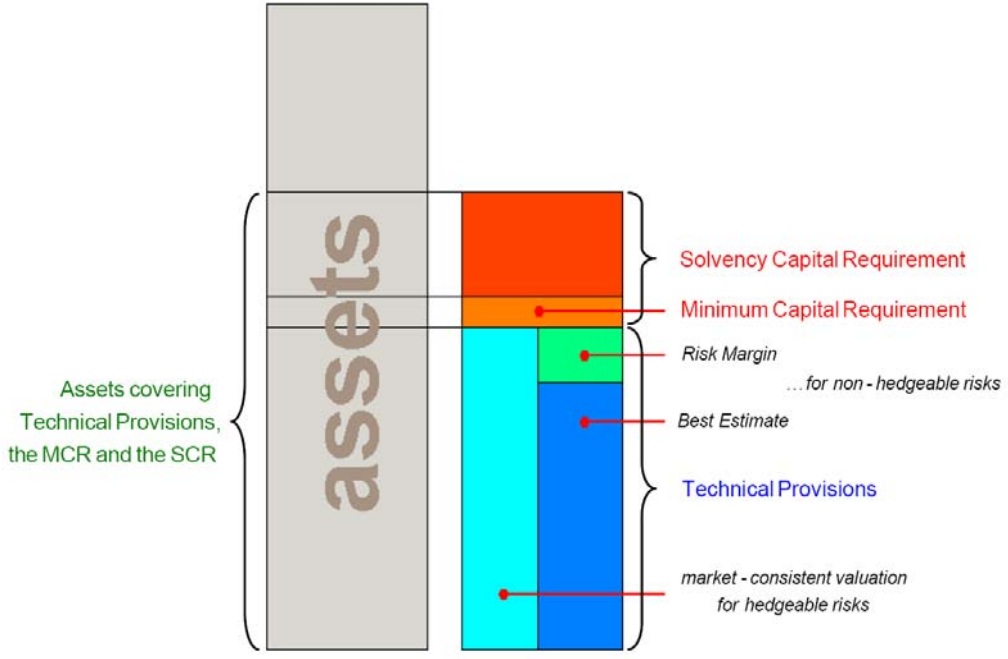
In case that *R* is throughout *additive*, condition (**) is trivially fulfilled in general, but then there is no strict diversification effect in the risk portfolio.

# VI. Risk Management under Solvency II and Basel III

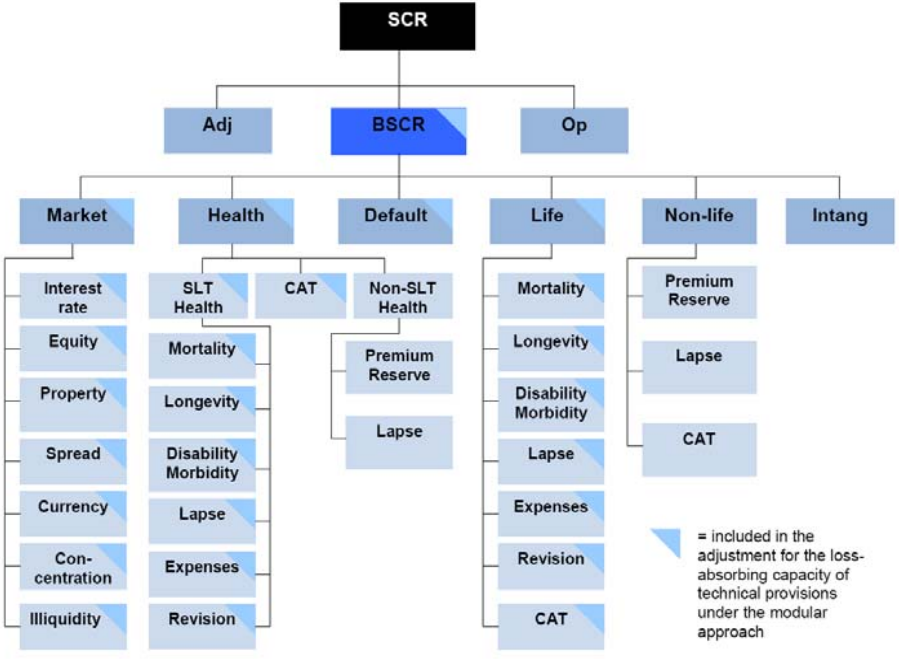



SCR: Solvency Capital Requirement

MCR: Minimum Capital Requirement




The economic balance sheet approach to Solvency II




Risk modules considered under Solvency II

Basic definition of the **S**olvency **C**apital **R**equirement (SCR) per risk module on the liability side, based on a *yearly time horizon*:

$$\mathrm{SCR}(X) = \mathrm{VaR}_{0.005}(X) - P(X)$$

where $X$ denotes the risk of the module (potential random payments to the clients) and $P(X)$ essentially the premium income.

The SCR for the market risk is mainly calculated by scenarios / factor models.

For both sides of the economic balance sheet, the volatility of the ***net asset value*** due to changes in the interest rate curve has to be taken into account, too.

The aggregation of the individual SCR's is performed with a covariance formula, based on given correlations $\rho_{ij}$ between the different modules $i$ and $j$:

$$\mathrm{SCR}_{\mathrm{total}} = \sqrt{\sum_{i,j} \rho_{ij} \cdot \mathrm{SCR}_i \cdot \mathrm{SCR}_j}$$

| $i \setminus j$ | Market | Default | Life | Health | Non-life |
|---|---|---|---|---|---|
| Market | 1 | 0.25 | 0.25 | 0.25 | 0.25 |
| Default | 0.25 | 1 | 0.25 | 0.25 | 0.5 |
| Life | 0.25 | 0.25 | 1 | 0.25 | 0 |
| Health | 0.25 | 0.25 | 0.25 | 1 | 0 |
| Non-life | 0.25 | 0.5 | 0 | 0 | 1 |

Example of a given correlation matrix from the Directive 2009/138/EC, Annex IV

**Motivation:** Standard Deviation Principle:

Suppose that the risks $X_1, \cdots, X_n$ have a pairwise correlation $\rho_{ij}$ for $1 \leq i, j \leq n$. If we identify the underlying risk measure with the *SDP* and the premium with the expectation, we obtain an individual $\mathrm{SCR}(X_i)$ accordingly as

$\mathrm{SCR}(X_i) = SDP(X_i) - E(X_i)$ for $1 \leq i \leq n$, with

**Motivation:** Standard Deviation Principle:

$$\begin{aligned}
\mathrm{SCR}\left(\sum_{i=1}^{n} X_i\right) &= SDP\left(\sum_{i=1}^{n} X_i\right) - \sum_{i=1}^{n} E(X_i) = \gamma\sqrt{Var\left(\sum_{i=1}^{n} X_i\right)} = \gamma\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)} \\
&= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij} \cdot \sqrt{\gamma Var(X_i) \cdot \gamma Var(X_j)}} \\
&= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij} \cdot \left\{SDP(X_i) - E(X_i)\right\} \cdot \left\{SDP(X_j) - E(X_j)\right\}} \\
&= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij} \cdot \mathrm{SCR}(X_i) \cdot \mathrm{SCR}(X_j)}
\end{aligned}$$

**Warning:** A simple application of this calculation scheme to the risk measure VaR is inconsistent in general!

Example 33: In what follows we consider Beta distributed risks $X$ with densities

$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m, \;\; 0 \le x \le 1; \;\; n,m \in \mathbb{Z}^+.$$

$F_X(x;n,m)$ will denote the corresponding cdf. Since the densities above are polynomials the convolution density for the aggregated risk $S = X + Y$ for independent summands with parameters $n_1, m_1, n_2, m_2$ is piecewise polynomial and can easily be calculated via the following formula:

$$f_S(x;n_1,m_1,n_2,m_2) = \begin{cases} \int_0^x f_X(y;n_1,m_1) \cdot f_Y(x-y;n_2,m_2)\,dy, & 0 \le x \le 1 \\ \int_{x-1}^1 f_X(y;n_1,m_1) \cdot f_Y(x-y;n_2,m_2)\,dy, & 1 \le x \le 2. \end{cases}$$

Example 33: In what follows we consider Beta distributed risks $X$ with densities

$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m, \ 0 \le x \le 1; \ n,m \in \mathbb{Z}^+.$$

Likewise, the cdf $F_S$ for the aggregated risk $S$ is also piecewise polynomial and can be calculated via

$$F_S(x;n_1,m_1,n_2,m_2) = \begin{cases} \int_0^x f_S(u;n_1,m_1,n_2,m_2)\,du, & 0 \le x \le 1 \\ F_S(1) + \int_1^x f_S(u;n_1,m_1,n_2,m_2)\,du, & 1 \le x \le 2. \end{cases}$$

Example 33: In what follows we consider Beta distributed risks $X$ with densities

$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m, \ 0 \le x \le 1; \ n,m \in \mathbb{Z}^+.$$

With the help of these formulas, it is possible to calculate (in the final step numerically) the true SCR's with $P(X) = E(X)$, for the individual risks as well as for the aggregated risk. Note that for a risk $X$ with density given above, we have

$$E(X) = \frac{n+1}{n+m+2}$$

and hence

$$\mathrm{SCR}(X) = \mathrm{VaR}_{0.005}(X) - E(X) = F_X^{-1}(0.995;n,m) - \frac{n+1}{n+m+2}.$$

Example 33: In what follows we consider Beta distributed risks $X$ with densities

$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m, \ 0 \le x \le 1; \ n,m \in \mathbb{Z}^+ = \mathbb{N} \cup \{0\}.$$

The following tables show some selected results.

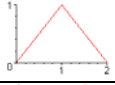
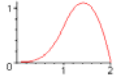
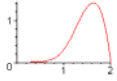
| $(n,m)$ | (0,0) | (1,0) | (2,0) | (3,0) | (0,1) | (0,2) | (0,3) |
|---|---|---|---|---|---|---|---|
| SCR($X$) | 0.4950 | 0.3308 | 0.2483 | 0.1987 | 0.5959 | 0.5790 | 0.5340 |

| $(n,m)$ | (1,2) | (1,3) | (1,4) | (2,1) | (3,1) | (4,1) |
|---|---|---|---|---|---|---|
| SCR($X$) | 0.4891 | 0.4816 | 0.4603 | 0.3706 | 0.3105 | 0.2670 |

Example 33: In what follows we consider Beta distributed risks $X$ with densities

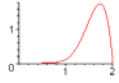
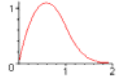
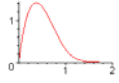
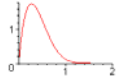
$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m, \ 0 \le x \le 1; \ n,m \in \mathbb{Z}^+ = \mathbb{N} \cup \{0\}.$$

The following tables contain the true SCR values for the aggregated risk $S = X + Y$, with *independent* Beta distributed risks $X$ and $Y$, in comparison to the values $\text{SCR}^{\sqrt{\ }}$ obtained via the covariance formula.

| $(n_1, m_1, n_2, m_2)$ | Density $f_s$ | SCR($S$) | $\text{SCR}^{\sqrt{\ }}(S)$ | Error in % |
|---|---|---|---|---|
| (0,0,0,0) |  | 0.9000 | 0.7000 | -22.21 |
| (1,0,1,0) |  | 0.6158 | 0.4678 | -24.02 |
| 2,0,2,0) |  | 0.4658 | 0.3511 | -24.61 |

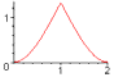
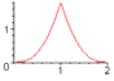
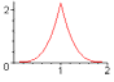
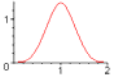
Example 33: In what follows we consider Beta distributed risks $X$ with densities

$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m, \ 0 \le x \le 1; \ n,m \in \mathbb{Z}^+ = \mathbb{N} \cup \{0\}.$$

| $(n_1, m_1, n_2, m_2)$ | Density $f_s$ | SCR($S$) | SCR$^{\sqrt{}}$($S$) | Error in % |
|---|---|---|---|---|
| (3,0,3,0) |  | 0.3743 | 0.2810 | -24.91 |
| (0,1,0,1) |  | 0.9171 | 0.8428 | -8.10 |
| (0,2,0,2) |  | 0.8187 | 0.8188 | 0.01 |
| (0,3,0,3) |  | 0.7229 | 0.7553 | 4.47 |

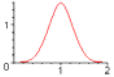
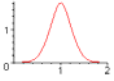
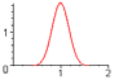
Example 33: In what follows we consider Beta distributed risks $X$ with densities

$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m,\ 0 \le x \le 1;\ n,m \in \mathbb{Z}^+ = \mathbb{N} \cup \{0\}.$$

| $(n_1, m_1, n_2, m_2)$ | Density $f_s$ | SCR($S$) | SCR$^{\sqrt{\ }}$($S$) | Error in % |
|---|---|---|---|---|
| (0,1,1,0) |  | 0.8008 | 0.6816 | -14.89 |
| (0,2,2,0) |  | 0.7056 | 0.6300 | -10.71 |
| (0,3,3,0) |  | 0.6239 | 0.5698 | -8.66 |
| (1,2,2,1) |  | 0.6851 | 0.6136 | -10.44 |

Example 33: In what follows we consider Beta distributed risks $X$ with densities

$$f_X(x;n,m) = (n+m+1)\binom{n+m}{n} x^n (1-x)^m,\ 0 \le x \le 1;\ n,m \in \mathbb{Z}^+ = \mathbb{N} \cup \{0\}.$$

| $(n_1, m_1, n_2, m_2)$ | Density $f_s$ | SCR($S$) | SCR$^{\sqrt{}}$($S$) | Error in % |
|---|---|---|---|---|
| (1,3,3,1) |  | 0.6276 | 0.5729 | -8.70 |
| (1,4,4,1) |  | 0.5760 | 0.5321 | -7.62 |
| … | … | … | … | … |
| (4,8,8,4) |  | 0.3816 | 0.4251 | 11.41 |

Similar examples can be presented for correlated risks.

The Calculation of the individual SCR in the non-life module is based on the assumption that the risk $X$ is of the form $X = Z \cdot V$ where $V$ is the (deterministic) premium volume of the following business year and $Z$ is a lognormally distributed random variable with expectation $m = E(Z) = 1$ and variance $s^2 = Var(Z) > 0$. In terms of the usual parametrization of the lognormal distribution as $\mathcal{LN}\left(\mu, \sigma^2\right)$ with $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, this means

$$1 = m = E(Z) = \exp\left(\mu + \frac{\sigma^2}{2}\right), \; s^2 = \exp\left(2\mu + \sigma^2\right) \cdot \left(e^{\sigma^2} - 1\right) = m^2 \cdot \left(e^{\sigma^2} - 1\right)$$

from which we obtain

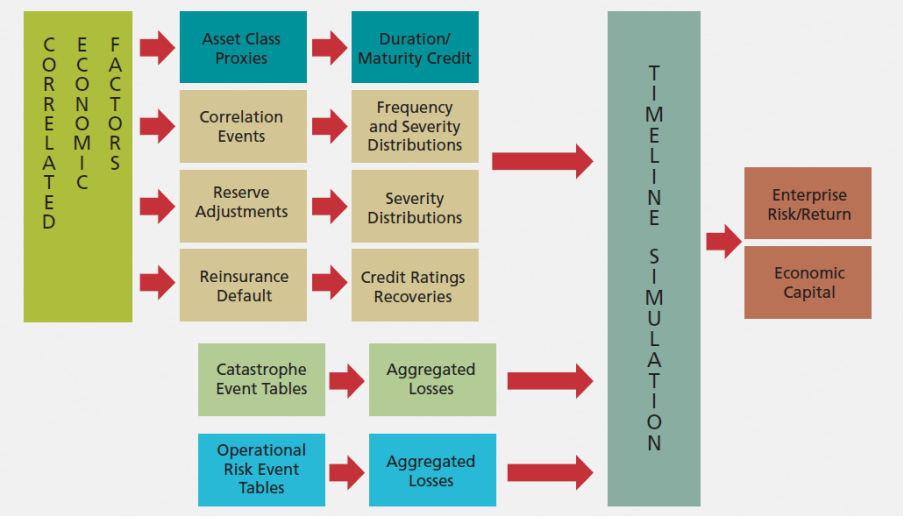
$$\sigma = \sqrt{\ln\left(1 + s^2\right)}, \; \mu = -\frac{\sigma^2}{2} = -\frac{1}{2}\ln\left(1 + s^2\right) = \ln\left(\frac{1}{\sqrt{1 + s^2}}\right).$$

By the definition of the SCR, we thus get the following formula:

$$\begin{aligned}\mathrm{SCR}(X) &= \mathrm{VaR}_{0.005}(X) - E(X) = \left(\mathrm{VaR}_{0.005}(Z) - E(Z)\right) \cdot V \\ &= \left(\exp\left(\mu + \sigma \cdot \Phi^{-1}(0.995)\right) - 1\right) \cdot V = \left( \frac{\Phi^{-1}(0.995) \cdot \exp\left(\sqrt{\ln\left(1+s^2\right)}\right)}{\sqrt{1+s^2}} - 1 \right) \cdot V\end{aligned}$$

for each non-life module.

## Internal and partial models:




Scheme of a fully implemented Internal Model

(source: Guy Carpenter & Company, LLC 2007)

**Internal and partial models:** Directive 2009/138/EC addresses the following topics:

***Test and Standards***

- Subsection 1: Use Test
- Subsection 2: Statistical Quality Standards
- Subsection 3: Calibration Standards
- Subsection 4: Integration of partial Internal Models
- Subsection 5: Profit and Loss Attribution
- Subsection 6: Validation Standards
- Subsection 7: Documentation Standards
- Subsection 8: External Models + Data

**Internal and partial models:** Directive 2009/138/EC Article 120: Use Test:

Insurance and reinsurance undertakings shall demonstrate that the internal model is widely used in and plays an important role in their system of governance, … in particular:

(a) their risk-management system … and their decision-making processes;

(b) their economic and solvency capital assessment and allocation processes, including the assessment …

In addition, insurance and reinsurance undertakings shall demonstrate that the frequency of calculation of the Solvency Capital Requirement using the internal model is consistent with the frequency with which they use their internal model for the other purposes covered by the first paragraph.

The administrative, management or supervisory body shall be responsible for ensuring the ongoing appropriateness of the design and operations of the internal model, and that the internal model continues to appropriately reflect the risk profile of the insurance and reinsurance undertakings concerned.

**Internal and partial models:** Article 121: Statistical Quality Standards:

… The methods used to calculate the probability distribution forecast shall be based on adequate, applicable and relevant actuarial and statistical techniques and shall be consistent with the methods used to calculate technical provisions.

The methods used to calculate the probability distribution forecast shall be based upon current and credible information and realistic assumptions.

…Data used for the internal model shall be accurate, complete and appropriate.

Insurance and reinsurance undertakings shall update the data sets used in the calculation of the probability distribution forecast at least annually.

No particular method for the calculation of the probability distribution forecast shall be prescribed.

Regardless of the calculation method chosen, the ability of the internal model to rank risk shall be sufficient to ensure that it is widely used in and plays an important role in the system of governance of insurance and reinsurance undertakings, in particular their risk-management system and decision-making processes, and capital allocation in accordance with Article 120.

**Internal and partial models:** Article 121: Statistical Quality Standards:

The internal model shall cover all of the material risks to which insurance and reinsurance undertakings are exposed. …

As regards diversification effects, insurance and reinsurance undertakings may take account in their internal model of dependencies within and across risk categories, provided that supervisory authorities are satisfied that the system used for measuring those diversification effects is adequate.

Insurance and reinsurance undertakings may take full account of the effect of risk-mitigation techniques in their internal model, as long as credit risk and other risks arising from the use of risk-mitigation techniques are properly reflected in the internal model.

…

**Internal and partial models:** Article 122: Calibration Standards:

Insurance and reinsurance undertakings may use a different time period or risk measure … for internal modelling purposes as long as the outputs of the internal model can be used by those undertakings to calculate the Solvency Capital Requirement in a manner that provides policy holders and beneficiaries with a level of protection equivalent to [99.5% over a one-year-period].

Where practicable, insurance and reinsurance undertakings shall derive the Solvency Capital Requirement directly from the probability distribution forecast generated by the internal model of those undertakings, using the Value-at-Risk measure …

**Internal and partial models:** Article 122: Calibration Standards:

Where insurance and reinsurance undertakings cannot derive the Solvency Capital Requirement directly from the probability distribution forecast generated by the internal model, the supervisory authorities may allow approximations to be used in the process to calculate the Solvency Capital Requirement, as long as those undertakings can demonstrate to the supervisory authorities that policy holders are provided with a level of protection equivalent to [99.5% over a one-year-period].

Supervisory authorities may require insurance and reinsurance undertakings to run their internal model on relevant benchmark portfolios and using assumptions based on external rather than internal data in order to verify the calibration of the internal model and to check that its specification is in line with generally accepted market practice.

**Internal and partial models:** Article 123: Validation Standards:

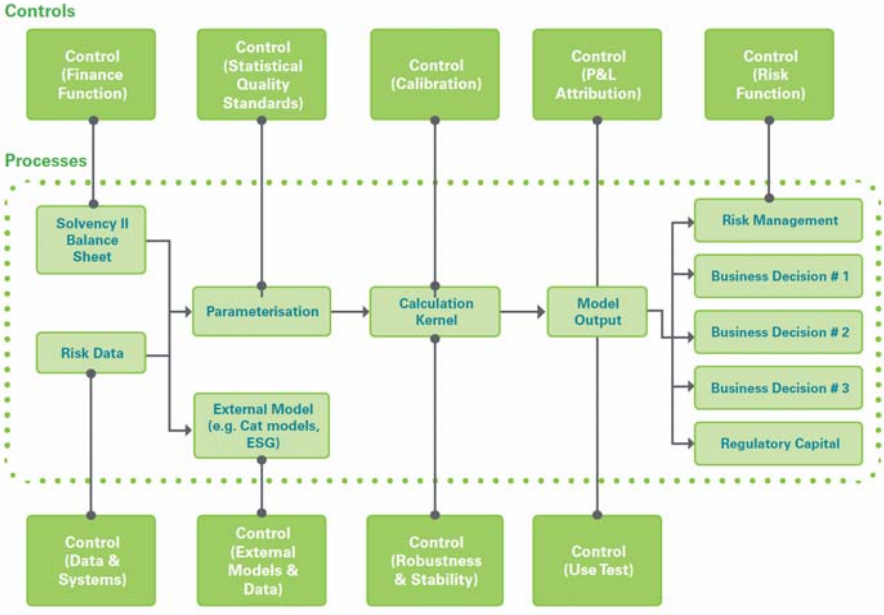
Insurance and reinsurance undertakings shall have a regular cycle of model validation which includes monitoring the performance of the internal model, reviewing the ongoing appropriateness of its specification, and testing its results against experience.

The model validation process shall include an effective statistical process for validating the internal model which enables the insurance and reinsurance undertakings to demonstrate to their supervisory authorities that the resulting capital requirements are appropriate.

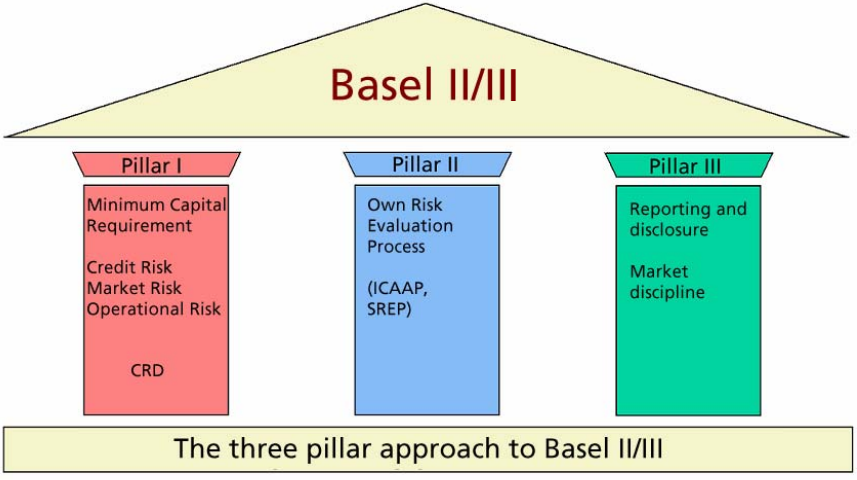
…

The model validation process shall include an analysis of the stability of the internal model and in particular the testing of the sensitivity of the results of the internal model to changes in key underlying assumptions. It shall also include an assessment of the accuracy, completeness and appropriateness of the data used by the internal model.

**Internal and partial models:** Article 123: Validation Standards:




(source: KPMG Europe LLP, 2011)




ICAAP: Internal Capital Adequacy Assessment Process

SREP: Supervisory Review and Evaluation Process

Risk Categories in Pillar I:

**Credit Risk:**

- relates to those assets and activities that a bank employs that have credit risk associated with them and do not form part of the trading activities of the bank. Otherwise it is part of the market risk.

- Credit risk capital requirements can be calculated in one of three ways:

1. standardized approach: attaches a specific risk weight to each credit-related instrument, depending on the credit quality of the assets on the bank's balance sheet.

## Risk Categories in Pillar I: Credit Risk: standardized approach:

| Risk weights for sovereigns | | | | | | |
|---|---|---|---|---|---|---|
| Credit rating | AAA to AA- | A+ to A- | BBB+ to BBB- | B+ to B- | Below B- | unrated |
| Risk weight | 0% | 20% | 50% | 100% | 150% | 100% |
| **Risk weights for banks and securities companies** | | | | | | |
| Credit rating | AAA to AA- | A+ to A- | BBB+ to BBB- | B+ to B- | Below B- | unrated |
| Risk weight | 20% | 50% | 100% | 100% | 150% | 100% |
| **Risk weights for corporates** | | | | | | |
| Credit rating | AAA to AA- | A+ to A- | BBB+ to BB- | | Below BB- | unrated |
| Risk weight | 20% | 50% | 100% | | 150% | 100% |

(source: F. de Weert)

Risk Categories in Pillar I:

**Credit Risk:**

- relates to those assets and activities of the bank employs that have credit risk associated with them and do not form part of the trading activities of the bank. Otherwise it is part of the market risk.

- Credit risk capital requirements can be calculated in one of three ways:

2. Foundation International Rating-Based (FIRB) approach: uses internal models for estimating the one-year probability of default (PD); regulator prescribed loss-given defaults (LGD's) and exposures at default (EaD's).

Risk Categories in Pillar I:

**Credit Risk:**

- relates to those assets and activities that a bank employs that have credit risk associated with them and do not form part of the trading activities of the bank. Otherwise it is part of the market risk.

- Credit risk capital requirements can be calculated in one of three ways:

3. Advanced International Rating-Based (AIRB) approach: uses internal models for estimating all of the one-year horizon parameters (PD's, LGD's, EaD's).

Risk Categories in Pillar I:

**Credit Risk:**

- relates to those assets and activities that a bank employs that have credit risk associated with them and do not form part of the trading activities of the bank. Otherwise it is part of the market risk.

- Credit risk capital requirement formulas:

Expected Loss: $EL = PD \times LGD \times EaD$

Unexpected Loss: $VaR_{0.001}(X) - EL = VaR_{0.001}(X) - PD \times LGD \times EaD$

(similar to the SCR formula in Solvency II)

Risk Categories in Pillar I:

**Market Risk:**

- relates to the trading activities of the bank and comprises
  - interest rate risk
  - foreign exchange rate risk
  - equity price risk
  - commodity price risk
  - option price risk.

- Market risk capital requirements can be calculated in different ways based on variants of the Value at Risk (parametric approach with a lognormal distribution for returns, Historic VaR, Monte Carlo VaR).

Risk Categories in Pillar I:

**Operational Risk:**

- relates to risks in executing the business, in particular
  - fraud risk (internal and external)
  - clients, products and business practice risk
  - business disruption and system failure risk
  - execution, delivery, and process management risk.

- Operational risk capital requirements can be calculated in different ways based on a simple gross factor approach or a more detailed consideration of eight specified lines of business, or a full internal model.

- The End -