Dependence of non-continuous random variables

Johanna Nešlehová

Department of Mathematics
ETH Zurich
Switzerland
www.math.ethz.ch/~johanna

August 31, 2005

Outline

Motivation

Consider

- Two possibly dependent loss types and

$$N_i(T) = \text{number of losses of type } i \text{ within the year } T$$

- Dependence structures of the pair of non continuous random variables

$$(N_1(T), N_2(T))$$

Other examples

- Counting random variables like claim/loss frequencies or number of defaults in a portfolio
- Variables with jumps like losses censored by a certain threshold

Distributions with continuous marginals

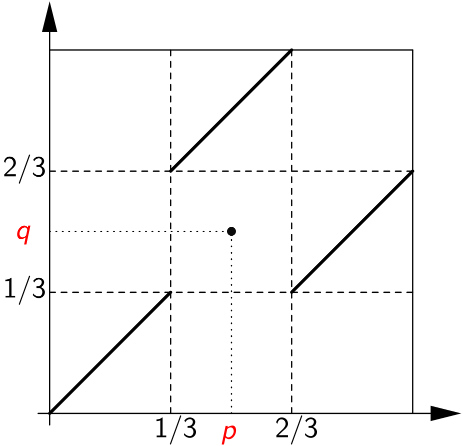
- Notation: $(X, Y) \sim F$ with marginals F_1 and F_2

F_1 and F_2 continuous

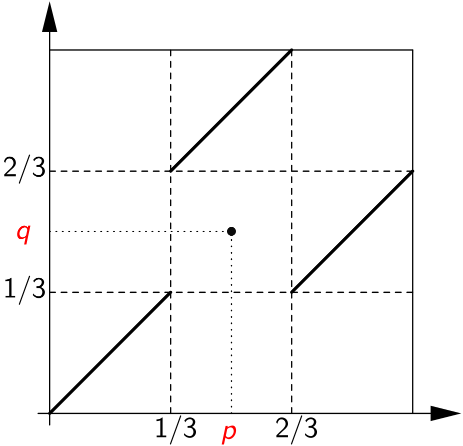
1. There exists a unique copula $\mathcal{C}$ such that $F(x, y) = \mathcal{C}(F_1(x), F_2(y))$ (Sklar's theorem)
2. Modeling of the marginals and the copula can be done separately
3. $\mathcal{C}$ captures dependence properties which are invariat under a.s. strictly increasing transformations of the marginals
4. Scale and translation invariant measures of dependence are functions of $\mathcal{C}$ alone

➠ $\mathcal{C}$ is the (scale and location invariant) dependence structure

Pathological Example



Pathological Example



- $(p, q) \in [0, 1/3] \times [0, 1/3]$: perfect positive dependence
- $(p, q) = (1/\sqrt{3}, 1/\sqrt{3})$: independence
- $(p, q) \in [2/3, 1] \times [2/3, 1]$: perfect negative dependence

Further Pitfalls

- The copula is not unique; i.e. several possible copulas exist
- The dependence structures of F are generally not the same as the dependence structures of the possible copulas

$$F(x,y) \leq F^*(x,y) \; \forall x,y \in \mathbb{R} \not\Rightarrow \mathcal{C}(u,v) \leq \mathcal{C}^*(u,v) \; \forall u,v \in [0,1]$$

- Possible copulas remain invariant under strictly increasing and continuous transformations, but do not necessarily change in the same way as the unique copula in the "continuous" case if at least one of the transformations is decreasing
- Weak convergence of F_n does not imply the point-wise convergence of the corresponding possible copulas
- Any measure of association which depends only on the copula of F is a constant

Kendall's Tau and Spearman's rho in the "Non-Continuous" Case?

"Naive" Idea: *work with some of the non unique fitting copula*

But: fitting copulas can differ considerably.

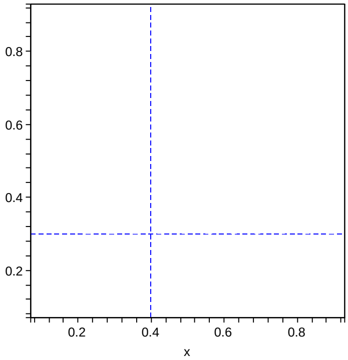
- For independent Bernoulli random variables X and Y
  - $\tau(X, Y) \in [-3/4, 3/4]$
  - $\rho(X, Y) \in [-13/16, 13/16]$
- for comonotonic Bernoulli random variables X and Y
  - $\tau(X, Y) \in [0, 1]$
  - $\rho(X, Y) \in [1/2, 1]$

Because: Difference between the probability of concordance and discordance $\neq 4 \int \mathcal{C}(u, v)\, d\mathcal{C}^*(u, v) - 1$

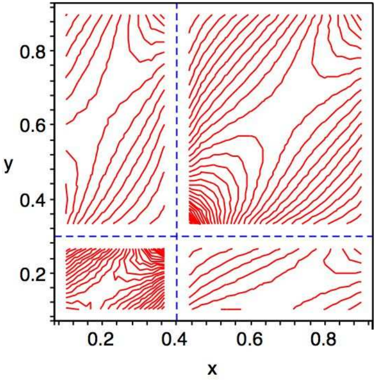
Some Starting Points

- Investigation of the family of possible copulas corresponding to a fixed bivariate distribution
  - In particular search for a suitable extension strategy which would produce a copula capturing the dependence structures of the joint distribution function
- Investigation of possible bivariate distributions obtained from a fixed copula and marginals which follow some specified distribution (up to parameters), such as Poisson, binomial or Bernoulli

Extension Strategy: Examples



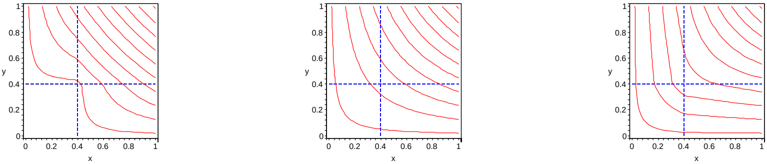
Subcopula for Bernoulli marginals:
$P[X = 0] = 0.4$, $P[Y = 0] = 0.3$
$P[X = 0, Y = 0] = 0.2$.



Extension by a Gauss copula with parameter $\varrho = 0.6$.

The "Standard" Extension Strategy

$\mathcal{C}(0.4, 0.4) = 0$ $\mathcal{C}(0.4, 0.4) = 0.16$ $\mathcal{C}(0.4, 0.4) = 0.3$



The Standard Extension $\mathcal{C}_S$

- The standard extension copula of Schweizer and Sklar
- $\mathcal{C}_S$ corresponds to the linear interpolation of the unique subcopula as well as the unique copula of the smoothed jdf

Some Selected Properties of $\mathcal{C}_S$

Pros

- X and Y are independent if and only if $\mathcal{C}_S$ is the independence copula
- (X^*, Y^*) more concordant than (X, Y) if and only if $\mathcal{C}_S(u, v) \leq \mathcal{C}_S^*(u, v)$ for all $u, v \in [0, 1]$
- $\mathcal{C}_S$ reacts on monotone transformations of the marginals as the unique copula in the "continuous" case

Cons

- If X and Y are perfect monotonic dependent, $\mathcal{C}_S$ does not coincide with the Fréchet-Hoeffding bounds
- Weak convergence does not imply the point-wise convergence of the $\mathcal{C}_S$'s

Towards Kendall's Tau and Spearman's Rho

Difference between the probabilities of concordance and discordance equals $4\int \mathcal{C}_S(u,v)\, d\mathcal{C}_S{}^*(u,v) - 1$

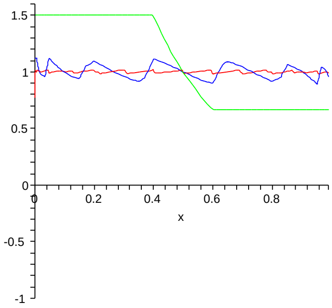
Consider

- Kendall's tau: $\widetilde{\tau}(\mathcal{C}_S) = 4\int_0^1\int_0^1 \mathcal{C}_S(u,v)\; d\mathcal{C}_S(u,v) - 1$
- Spearman's rho: $\widetilde{\rho}(\mathcal{C}_S) = 12\int_0^1\int_0^1 [\mathcal{C}_S(u,v) - uv]\; du\, dv$

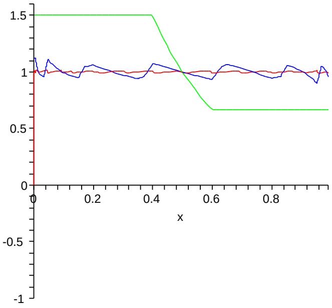
However

- $\widetilde{\tau}$ and $\widetilde{\rho}$ do not reach the bounds 1 and -1
- Exact bounds of $\widetilde{\tau}$ and $\widetilde{\rho}$ are complicated and do not have the same absolute value

Exact Bounds for Kendall's Tau and Spearman's Rho



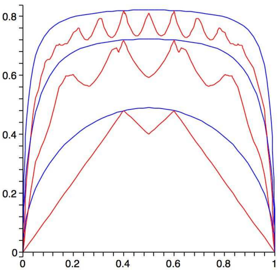
$|\widetilde{\tau}(\mathcal{M}_S)/\widetilde{\tau}(\mathcal{W}_S)|$ for Binomial distributions $F_1 = \mathcal{B}(n, 0.4)$ and $F_2 = \mathcal{B}(n, x)$ with $n =$ 1, 4 and 10.

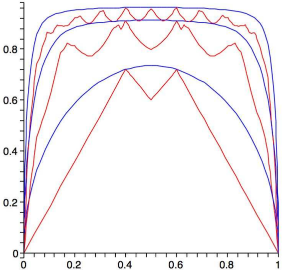


$|\widetilde{\rho}(\mathcal{M}_S)/\widetilde{\rho}(\mathcal{W}_S)|$ for Binomial distributions $F_1 = \mathcal{B}(n, 0.4)$ and $F_2 = \mathcal{B}(n, x)$ with $n =$ 1, 4 and 10.

Less Sharp Bounds for Kendall's Tau and Spearman's Rho

1. $|\widetilde{\tau}(X,Y)| \leq \sqrt{1 - \mathrm{E}\,\Delta F_1(X)}\sqrt{1 - \mathrm{E}\,\Delta F_2(Y)}$. The bounds are attained if $Y = T(X)$ a.s. where T is a strictly monotone and continuous transformation on the range of X.
2. $|\widetilde{\rho}(X,Y)| \leq \sqrt{1 - \mathrm{E}\,\Delta F_1(X)^2}\sqrt{1 - \mathrm{E}\,\Delta F_2(Y)^2}$. The bounds are attained if $Y = T(X)$ a.s. where T is a strictly monotone and continuous transformation on the range of X.





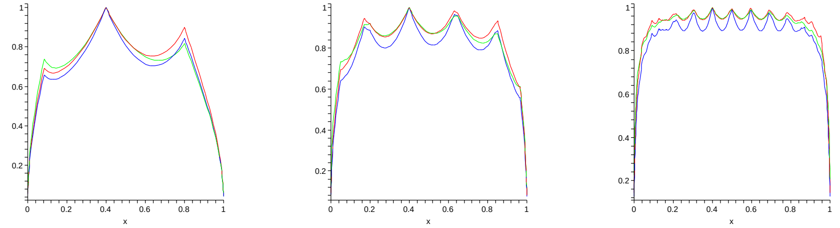
Kendall's Tau and Spearman's Rho for Non-Continuous Random Variables

$$\tau(X,Y) = \frac{4\int_0^1\int_0^1 \mathcal{C}_S(u,v)\, d\mathcal{C}_S(u,v) - 1}{\sqrt{[1 - \mathrm{E}\,\Delta F_1(X)][1 - \mathrm{E}\,\Delta F_2(Y)]}}$$

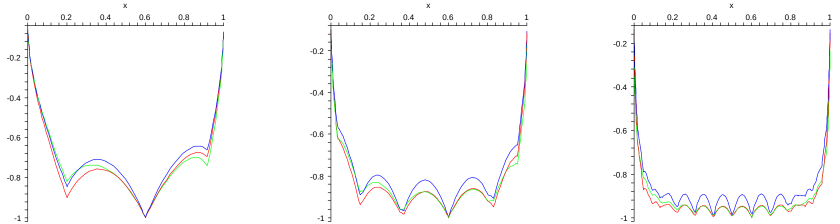
$$\rho(X,Y) = \frac{12\left(\int_0^1\int_0^1 \big(\mathcal{C}_S(u,v) - uv\big)\, du\, dv\right)}{\sqrt{[1 - \mathrm{E}\,\Delta F_1(X)^2][1 - \mathrm{E}\,\Delta F_2(Y)^2]}}$$

- τ and ρ satisfy (modified) axioms of concordance measures
- Bounds are attained if $X \overset{\text{a.s.}}{=} T(Y)$ for T continuous and strictly monotone
- τ and ρ are the sample versions for empirical distributions

τ, ρ and ϱ for Binomial Distributions

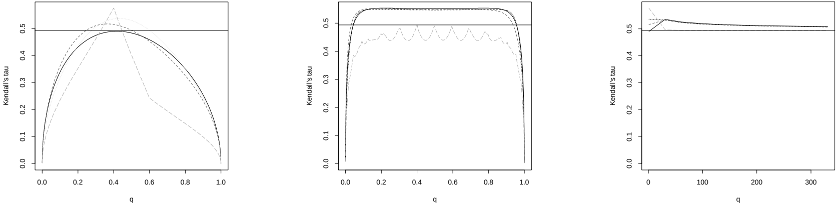


$F_1 = \mathcal{B}(2, 0.4)$, $F_2 = \mathcal{B}(2, x)$. $F_1 = \mathcal{B}(4, 0.4)$, $F_2 = \mathcal{B}(4, x)$. $F_1 = \mathcal{B}(10, 0.4)$, $F_2 = \mathcal{B}(10, x)$.



Some Problems

- Modeling: dependence properties of a family $\mathcal{H} = \{\mathcal{C}(F, G) : F \in \mathcal{F}, G \in \mathcal{G}\}$



Kendall's tau for binomial marginals and a Gauss copula (solid line), Frank copula (short-dashed line), Gumbel copula (dotted line), and Fréchet copula (long-dashed line).

References

M. Denuit and P. Lambert, 2005
Constraints on concordance measures in bivariate discrete data

W. Hoeffding, 1940
Maßstabinvariante Korrelationstheorie für diskontinuiereliche Verteilungen

A. Marshall, 1996
Copulas, marginals and joint distributions

J. Nešlehová, 2004
Dependence of non-continous random variables

Some Theory Behind: Dependence Structure in the Non-Continous Case

In the continuous case: $\mathcal{C}$ is the cdf. of the transformed vector $(F_1(X), F_2(Y))$.
Idea: In the non-continuous case, use a different transformation of the marginals:

$$\psi(x,u) := \mathrm{P}[X < x] + u\mathrm{P}[X = x] = F(x-) + u\Delta F(x)$$

Result: for a random vector (U, V) with uniform marginals which is independent of (X, Y), $(\psi(X,U), \psi(Y,V))$ has uniform marginals and the corresponding unique copula is a possible copula of (X, Y).
Moreover: different dependence structure of (U, V) leads to different possible copulas of (X, Y).