Elaborato per il corso di Complex Networks Analysis

Dario Comanducci

Master in Data Science and Statistical Learning
Università degli Studi di Firenze

28 Marzo 2025

Ricostruzione di reti ∼ *Max-Ent*

Pallavi Baral and José Pedro Fique.
Estimation of bilateral exposures: A copula approach.
In 1st annual CIRANO workshop on networks in trade and finance, 2012.

Generazione di reti ∼ *Configuration Model*

Mathias Raschke, Markus Schläpfer, and Konstantinos Trantopoulos.
Copula-based modeling of degree-correlated networks.
Journal of Statistical Mechanics: Theory and Experiment, 2014.

Distribuzioni congiunte & copule

- Ogni distribuzione congiunta descrive implicitamente sia il comportamento marginale delle singole variabili aleatorie sia la loro struttura di dipendenza
- Al di fuori del framework gaussiano, le marginali e la correlazione non specificano completamente la distribuzione congiunta
- Le copule forniscono un modo per isolare la descrizione della struttura di dipendenza, separandola dalle distribuzioni marginali

Definizione

Una copula è una funzione di distribuzione congiunta con distribuzioni marginali uniformi standard.

Per una copula in d dimensioni, viene impiegata la notazione

$$C(u_1, \dots, u_d) \equiv P(U_1 \leq u_1, \dots, U_d \leq u_d)$$

Il teorema di Sklar

Trasformazioni rilevanti

Sia X una v.a. continua (*), con *funzione di ripartizione* $F_X(x) = P(X \leq x)$ invertibile (**); valgono le seguenti proprietà:

- $F_X(X) \sim \mathcal{U}(0,1) \quad (*)$
- $U \sim \mathcal{U}(0,1) \Rightarrow$ la v.a. $F_X^{-1}(U)$ ha come funzione di ripartizione $F_X(x) \quad (**)$

Il teorema di Sklar

Trasformazioni rilevanti

Sia X una v.a. continua (*), con *funzione di ripartizione* $F_X(x) = P(X \leq x)$ invertibile (**); valgono le seguenti proprietà:

- $F_X(X) \sim \mathcal{U}(0,1) \quad (*)$
- $U \sim \mathcal{U}(0,1) \Rightarrow$ la v.a. $F_X^{-1}(U)$ ha come funzione di ripartizione $F_X(x) \quad (**)$

Sklar

Sia $F(x_i, \ldots, x_d)$ una funzione di ripartizione congiunta con ripartizioni marginali $F_1(x) \ldots F_d(x)$: allora esiste una copula $C : [0,1]^d \to [0,1]$ tale che

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \tag{1}$$

Viceversa, se C è una copula e $F_1(x) \ldots F_d(x)$ sono funzioni di ripartizione univariate, allora la funzione F in Eq. (1) è una funzione di ripartizione congiunta con marginali $F_1(x) \ldots F_d(x)$. C è unica per $F_i(x)$ continue.

Alcuni tipi di copule

Caso bivariato ($d = 2$)

- copula gaussiana $C(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$
  Φ_ρ distr. congiunta di 2 v.a. normali standard con indice di correlazione $\rho \in (-1,1)$; $\Phi \equiv$ cdf $\mathcal{N}(0,1)$
- copula t $C_{\nu,\rho} = t_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v))$
  $t_{\nu,\rho}$ distr. t congiunta di 2 v.a. con ν gradi di libertà e indice di correlazione $\rho \in (-1,1)$; t_ν è la cdf della distribuzione t standard univariata
- copula di Gumbel $C_\vartheta(u,v) = \exp\left(-\left[(-\ln u)^\vartheta + (-\ln v)^\vartheta\right]^{\frac{1}{\vartheta}}\right), \quad \vartheta \in [1, \infty)$
- copula di Clayton $C_\vartheta(u,v) = \left(u^{-\vartheta} + v^{-\vartheta} - 1\right)^{-1/\vartheta}, \quad \vartheta \in (0, \infty)$

Alcuni tipi di copule

Caso bivariato ($d = 2$)

- copula gaussiana $C(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$
  Φ_ρ distr. congiunta di 2 v.a. normali standard con indice di correlazione $\rho \in (-1,1)$; $\Phi \equiv$ cdf $\mathcal{N}(0,1)$
- copula t $C_{\nu,\rho} = t_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v))$
  $t_{\nu,\rho}$ distr. t congiunta di 2 v.a. con ν gradi di libertà e indice di correlazione $\rho \in (-1,1)$; t_ν è la cdf della distribuzione t standard univariata
- copula di Gumbel $C_\vartheta(u,v) = \exp\left(-\left[(-\ln u)^\vartheta + (-\ln v)^\vartheta\right]^{\frac{1}{\vartheta}}\right), \quad \vartheta \in [1,\infty)$
- copula di Clayton $C_\vartheta(u,v) = \left(u^{-\vartheta} + v^{-\vartheta} - 1\right)^{-1/\vartheta}, \quad \vartheta \in (0,\infty)$

Alcuni tipi di copule

Caso bivariato ($d = 2$)

- copula gaussiana $C(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$
  Φ_ρ distr. congiunta di 2 v.a. normali standard con indice di correlazione $\rho \in (-1,1)$; $\Phi \equiv$ cdf $\mathcal{N}(0,1)$
- copula t $C_{\nu,\rho} = t_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v))$
  $t_{\nu,\rho}$ distr. t congiunta di 2 v.a. con ν gradi di libertà e indice di correlazione $\rho \in (-1,1)$; t_ν è la cdf della distribuzione t standard univariata
- copula di Gumbel $C_\vartheta(u,v) = \exp\left(-\left[(-\ln u)^\vartheta + (-\ln v)^\vartheta\right]^{\frac{1}{\vartheta}}\right), \quad \vartheta \in [1,\infty)$
- copula di Clayton $C_\vartheta(u,v) = \left(u^{-\vartheta} + v^{-\vartheta} - 1\right)^{-1/\vartheta}, \quad \vartheta \in (0,\infty)$

Alcuni tipi di copule

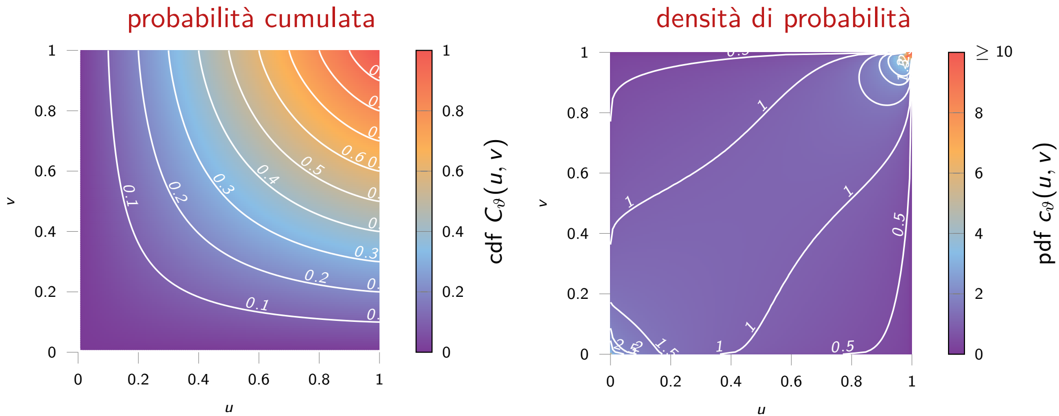
Caso bivariato ($d = 2$)

- copula gaussiana $C(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$
  Φ_ρ distr. congiunta di 2 v.a. normali standard con indice di correlazione $\rho \in (-1,1)$; $\Phi \equiv$ cdf $\mathcal{N}(0,1)$
- copula t $C_{\nu,\rho} = t_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v))$
  $t_{\nu,\rho}$ distr. t congiunta di 2 v.a. con ν gradi di libertà e indice di correlazione $\rho \in (-1,1)$; t_ν è la cdf della distribuzione t standard univariata
- copula di Gumbel $C_\vartheta(u,v) = \exp\left(-\left[(-\ln u)^\vartheta + (-\ln v)^\vartheta\right]^{\frac{1}{\vartheta}}\right), \quad \vartheta \in [1, \infty)$
- copula di Clayton $C_\vartheta(u,v) = \left(u^{-\vartheta} + v^{-\vartheta} - 1\right)^{-1/\vartheta}, \quad \vartheta \in (0, \infty)$

Nelle copule di Gumbel e Clayton ϑ misura, al suo crescere, il grado di dipendenza tra le due variabili. La copula Gumbel modella eventi con dipendenza forte nelle code superiori; quella di Clayton è adatta per dipendenza forte nelle code inferiori

Copula di Gumbel

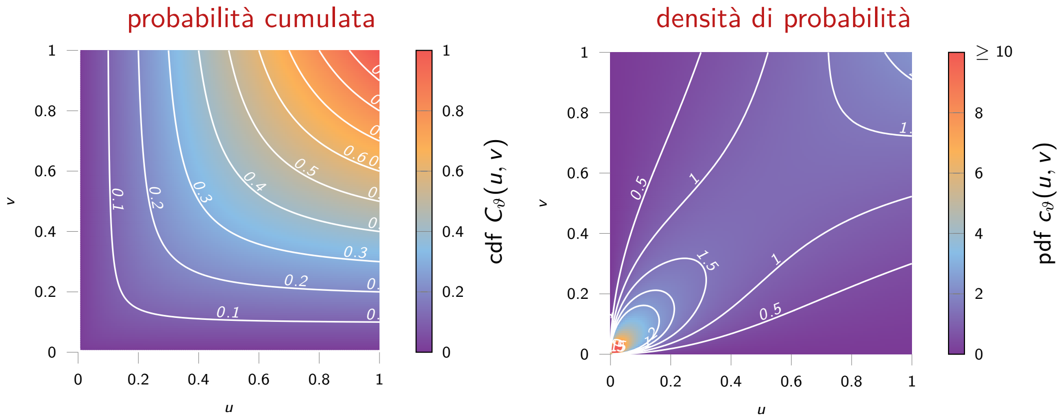
Un esempio: $\vartheta = 5/4 = 1.25$



C'è una maggior tendenza per Y a manifestare un evento estremo nella coda superiore quando lo fa anche X

Copula di Clayton

Un esempio: $\vartheta = 5/4 = 1.25$



C'è una maggior tendenza per Y a manifestare un evento estremo nella coda inferiore quando lo fa anche X

Stima di una copula tramite massima verosimiglianza

Dato un campione di N osservazioni su d v.a. $X_1 \ldots X_d$ con marginali $F_1(X_1) \ldots F_d(X_d)$

$$\left\{ \mathbf{x}^{(n)} \in \mathbb{R}^d \right\}_{n=1}^{N}$$

sia C_ϑ una copula avente $\boldsymbol{\vartheta}$ come vettore dei parametri da stimare su $\mathbf{x}^{(1)} \ldots \mathbf{x}^{(N)}$

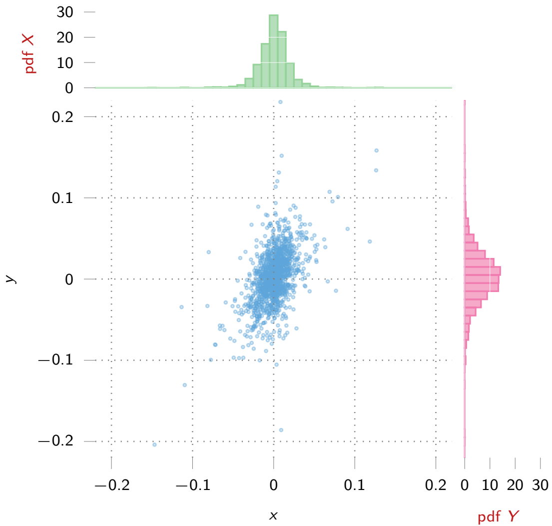
La stima di massima verosimiglianza MLE è ottenuta massimizzando rispetto a $\boldsymbol{\vartheta}$ la log-Likelihood

$$\sum_{n=1}^{N} \ln c_\vartheta \left(F_1\left(x_1^{(n)}\right), \ldots, F_d\left(x_d^{(n)}\right) \right)$$

dove c_ϑ è la densità di probabilità della copula

Un esempio

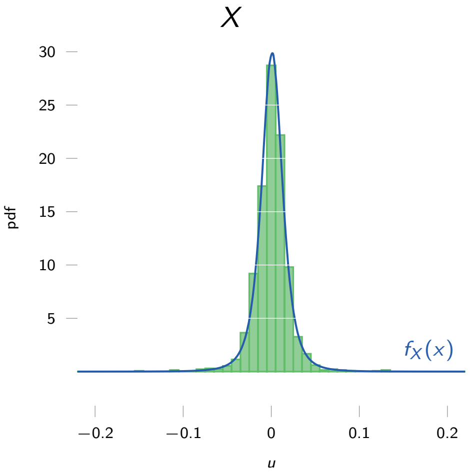
$P(X, Y)$?

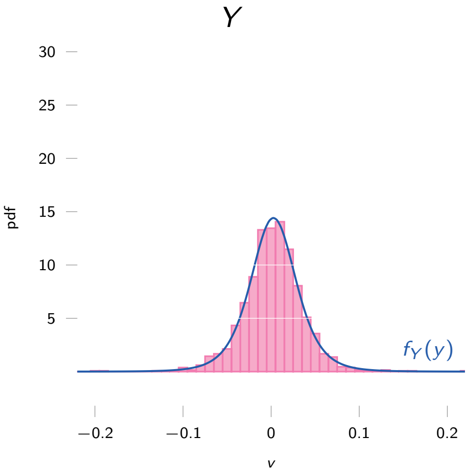


Complex Networks

Un esempio

$X, Y \sim t_\nu(\mu, \sigma^2)$; $\qquad f_t(x) = \dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)\sigma} \cdot \left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$



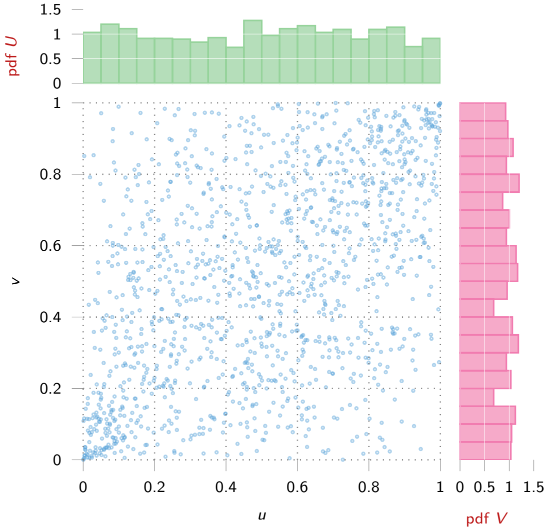


μ: 0.0009; σ: 0.0123; ν: 3.1775

μ: 0.0026; σ: 0.0264; ν: 5.0992

Un esempio

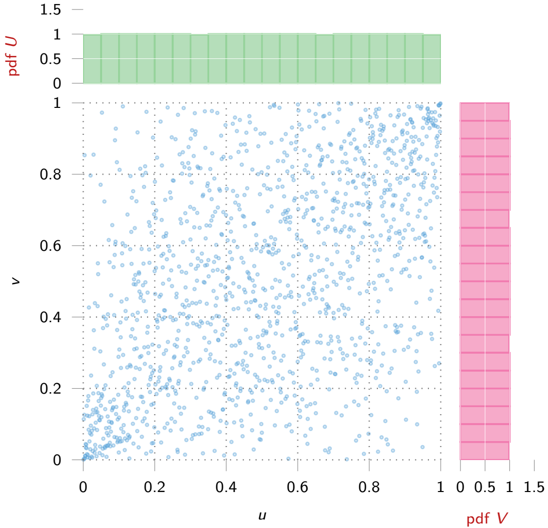
$U = F_X(X), V = F_Y(Y) \sim \mathcal{U}(0,1)$



Complex Networks

Un esempio

$U = \hat{F}_X(X), V = \hat{F}_Y(Y) \sim \mathcal{U}(0,1); \qquad \hat{F}(x) = \sum_{i=1}^{n} \mathbf{1}(x_i \leq x)/(n+1)$

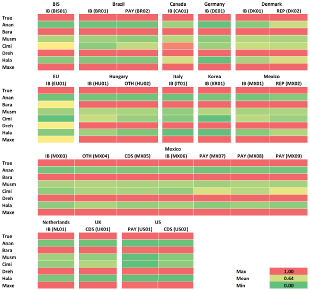


Complex Networks

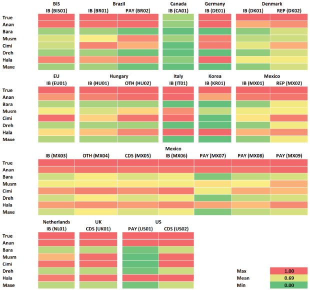
La Horse Race

Estimation of bilateral exposures: A copula approach

density



true links



true non-links

- Max-Ent ridistribuisce i pesi in maniera *omogenea* su tutti gli archi
- Idea: usare le copule per generare una matrice basata sulla dipendenza stimata, ottenendo quindi stime bilaterali più informate

Stima della copula

Estimation of bilateral exposures: A copula approach

$$
\begin{array}{c|cccc|c}
 & & & & & \sum_j x_{ij} \\
 & x_{11} & x_{12} & \cdots & x_{1N} & x_{1\bullet} \\
 & x_{21} & x_{22} & \cdots & x_{2N} & x_{2\bullet} \\
 & \vdots & \vdots & \ddots & \vdots & \vdots \\
 & x_{N1} & x_{N2} & \cdots & x_{NN} & x_{N\bullet} \\
\hline
\sum_i x_{ij} & x_{\bullet 1} & x_{\bullet 2} & \cdots & x_{\bullet N} &
\end{array}
$$

$u_n = \hat{F}_1(x_{\bullet n})$; $v_n = \hat{F}_2(x_{n\bullet})$ $\hat{F}_1(x), \hat{F}_2(x)$: funzioni di ripartizione empiriche

Gumbel

$$C_\vartheta(u_n, v_n) = \exp\left(-\left[(-\ln u_n)^\vartheta + (-\ln v_n)^\vartheta\right]^{1/\vartheta}\right)$$

$$\hat{\vartheta} = \arg\max_\vartheta \ln L(\vartheta | \{x_{\bullet n}, x_{n\bullet}\}_{n=1}^N) = \arg\max_\vartheta \sum_{n=1}^N \ln c_\vartheta(u_n, v_n)$$

Ricostruzione di $\mathtt{X}$

Estimation of bilateral exposures: A copula approach

Sketch dell'algoritmo

- Creazione della *matrice di probabilità* $\mathtt{P}$: $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$
- Conversione di $\mathtt{P}$ nella *mastrice stocastica* $\mathtt{S}$ con normalizzazione per riga di $\mathtt{P}$: $s_{ij} = \frac{p_{ij}}{\sum_{n=1}^{N} p_{in}}$
- First guess per $\mathtt{X}$: $\hat{x}_{ij}^{(0)} = s_{ij} \cdot x_{i\bullet}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$
- Recupero dei valori marginali lungo le colonne $\sum_i x_{ij}^{(0)} = x_{\bullet j}$ applicando a $\mathtt{X}^{(0)}$ l'algoritmo RAS fino a convergenza

Sketch dell'algoritmo

- Creazione della *matrice di probabilità* $\mathbb{P}$: $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$
- Conversione di $\mathbb{P}$ nella *mastrice stocastica* $\mathbb{S}$ con normalizzazione per riga di $\mathbb{P}$: $s_{ij} = \frac{p_{ij}}{\sum_{n=1}^{N} p_{in}}$
- First guess per $\mathbb{X}$: $\hat{x}_{ij}^{(0)} = s_{ij} \cdot x_{i\bullet}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$
- Recupero dei valori marginali lungo le colonne $\sum_i x_{ij}^{(0)} = x_{\bullet j}$ applicando a $\mathbb{X}^{(0)}$ l'algoritmo RAS fino a convergenza

Sketch dell'algoritmo

- Creazione della *matrice di probabilità* P: $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$
- Conversione di P nella *mastrice stocastica* S con normalizzazione per riga di P: $s_{ij} = \frac{p_{ij}}{\sum_{n=1}^{N} p_{in}}$
- First guess per X: $\hat{x}_{ij}^{(0)} = s_{ij} \cdot x_{i\bullet}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$
- Recupero dei valori marginali lungo le colonne $\sum_i x_{ij}^{(0)} = x_{\bullet j}$ applicando a $X^{(0)}$ l'algoritmo RAS fino a convergenza

Sketch dell'algoritmo

- Creazione della *matrice di probabilità* $\mathtt{P}$: $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$
- Conversione di $\mathtt{P}$ nella *mastrice stocastica* $\mathtt{S}$ con normalizzazione per riga di $\mathtt{P}$: $s_{ij} = \frac{p_{ij}}{\sum_{n=1}^{N} p_{in}}$
- First guess per $\mathtt{X}$: $\hat{x}_{ij}^{(0)} = s_{ij} \cdot x_{i\bullet}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$
- Recupero dei valori marginali lungo le colonne $\sum_i x_{ij}^{(0)} = x_{\bullet j}$ applicando a $\mathtt{X}^{(0)}$ l'algoritmo RAS fino a convergenza

Sketch dell'algoritmo

- Creazione della *matrice di probabilità* P: $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$
- Conversione di P nella *mastrice stocastica* S con normalizzazione per riga di P: $s_{ij} = \frac{p_{ij}}{\sum_{n=1}^{N} p_{in}}$
- First guess per X: $\hat{x}_{ij}^{(0)} = s_{ij} \cdot x_{i\bullet}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$
- Recupero dei valori marginali lungo le colonne $\sum_i x_{ij}^{(0)} = x_{\bullet j}$ applicando a $X^{(0)}$ l'algoritmo RAS fino a convergenza

L'algoritmo RAS

Input $\mathtt{A} \in \mathbb{R}^{m \times n} = [a_{ij}]$ a valori *non negativi*; $\mathbf{z} \in \mathbb{R}^m$ e $\mathbf{w} \in \mathbb{R}^n$ vettori *positivi*

Sketch dell'algoritmo

- Creazione della *matrice di probabilità* $\mathtt{P}$: $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$
- Conversione di $\mathtt{P}$ nella *mastrice stocastica* $\mathtt{S}$ con normalizzazione per riga di $\mathtt{P}$: $s_{ij} = \frac{p_{ij}}{\sum_{n=1}^{N} p_{in}}$
- First guess per $\mathtt{X}$: $\hat{x}_{ij}^{(0)} = s_{ij} \cdot x_{i\bullet}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$
- Recupero dei valori marginali lungo le colonne $\sum_i x_{ij}^{(0)} = x_{\bullet j}$ applicando a $\mathtt{X}^{(0)}$ l'algoritmo RAS fino a convergenza

L'algoritmo RAS

Input $\mathtt{A} \in \mathbb{R}^{m \times n} = [a_{ij}]$ a valori *non negativi*; $\mathbf{z} \in \mathbb{R}^m$ e $\mathbf{w} \in \mathbb{R}^n$ vettori *positivi*

Output $\mathtt{B} \in \mathbb{R}^{m \times n} = [b_{ij}] : \mathtt{B} = \mathrm{diag}(\mathbf{x})\mathtt{A}\mathrm{diag}(\mathbf{y})$ e $\sum_{j=1}^{n} b_{ij} = z_i; \ \sum_{i=1}^{m} b_{ij} = w_j$ valgano per qualche $\mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$ vettori non negativi.

Sketch dell'algoritmo

- Creazione della *matrice di probabilità* $\mathtt{P}$: $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$
- Conversione di $\mathtt{P}$ nella *mastrice stocastica* $\mathtt{S}$ con normalizzazione per riga di $\mathtt{P}$: $s_{ij} = \frac{p_{ij}}{\sum_{n=1}^{N} p_{in}}$
- First guess per $\mathtt{X}$: $\hat{x}_{ij}^{(0)} = s_{ij} \cdot x_{i\bullet}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$
- Recupero dei valori marginali lungo le colonne $\sum_i x_{ij}^{(0)} = x_{\bullet j}$ applicando a $\mathtt{X}^{(0)}$ l'algoritmo RAS fino a convergenza

L'algoritmo RAS

Input $\mathtt{A} \in \mathbb{R}^{m \times n} = [a_{ij}]$ a valori *non negativi*; $\mathbf{z} \in \mathbb{R}^m$ e $\mathbf{w} \in \mathbb{R}^n$ vettori *positivi*

Output $\mathtt{B} \in \mathbb{R}^{m \times n} = [b_{ij}] : \mathtt{B} = \mathrm{diag}(\mathbf{x})\mathtt{A}\mathrm{diag}(\mathbf{y})$ e $\sum_{j=1}^{n} b_{ij} = z_i$; $\sum_{i=1}^{m} b_{ij} = w_i$ valgano per qualche $\mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$ vettori non negativi.

Sketch: Ripetere fino a convergenza

- $r_i \leftarrow z_i \big/ \sum_j a_{ij}$ $(i = 2 \ldots m$, con $r_1 \leftarrow 1)$
- $\mathtt{A} \leftarrow \mathrm{diag}(\mathbf{r})\, \mathtt{A}$
- $s_j \leftarrow w_j \big/ \sum_i a_{ij}$ per $j = 1 \ldots m$
- $\mathtt{A} \leftarrow \mathtt{A}\, \mathrm{diag}(\mathbf{s})$

Critiche

Estimation of bilateral exposures: A copula approach

- Perché $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$ e non $p_{ij} = c_{\hat{\vartheta}}(u_i, v_j)$?

$$c_\vartheta(u, v) = \frac{C_\vartheta(u, v)}{u\,v}\,(t_\vartheta(u)+t_\vartheta(v))^{-2+2/\vartheta} \cdot (\ln u \ln v)^{\vartheta-1}(1+(\vartheta-1)(t_\vartheta(u)+t_\vartheta(v))^{-1/\vartheta})$$

$$t_\vartheta(w) = (-\ln w)^\vartheta$$

- Perché normalizzare $\mathbb{S}$ per riga, invece che $\mathbb{S} \leftarrow \mathbb{P}/\sum_{ij} p_{ij}$?
- Perché inizializzare $\mathbb{X}^{(0)}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$, invece che $\mathbb{X}^{(0)} \leftarrow \mathbb{S} \sum_i x_{i\bullet}$?
- Perché usare solo Gumbel, invece di testare un pool di copule e selezionare quella col miglior AIC o BIC?

- Perché $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$ e non $p_{ij} = c_{\hat{\vartheta}}(u_i, v_j)$?

$$c_\vartheta(u,v) = \frac{C_\vartheta(u,v)}{u\,v}(t_\vartheta(u)+t_\vartheta(v))^{-2+2/\vartheta} \cdot (\ln u \ln v)^{\vartheta-1}(1+(\vartheta-1)(t_\vartheta(u)+t_\vartheta(v))^{-1/\vartheta})$$
$$t_\vartheta(w) = (-\ln w)^\vartheta$$

- Perché normalizzare $\mathtt{S}$ per riga, invece che $\mathtt{S} \leftarrow \mathtt{P}/\sum_{ij} p_{ij}$?
- Perché inizializzare $\mathtt{X}^{(0)}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$, invece che $\mathtt{X}^{(0)} \leftarrow \mathtt{S}\sum_i x_{i\bullet}$?
- Perché usare solo Gumbel, invece di testare un pool di copule e selezionare quella col miglior AIC o BIC?

Critiche

Estimation of bilateral exposures: A copula approach

- Perché $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$ e non $p_{ij} = c_{\hat{\vartheta}}(u_i, v_j)$?

$$c_\vartheta(u,v) = \frac{C_\vartheta(u,v)}{u\,v}\,(t_\vartheta(u)+t_\vartheta(v))^{-2+2/\vartheta}\cdot(\ln u \ln v)^{\vartheta-1}(1+(\vartheta-1)(t_\vartheta(u)+t_\vartheta(v))^{-1/\vartheta})$$

$$t_\vartheta(w) = (-\ln w)^\vartheta$$

- Perché normalizzare $\mathtt{S}$ per riga, invece che $\mathtt{S} \leftarrow \mathtt{P}/\sum_{ij} p_{ij}$?
- Perché inizializzare $\mathtt{X}^{(0)}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$, invece che $\mathtt{X}^{(0)} \leftarrow \mathtt{S}\sum_i x_{i\bullet}$?
- Perché usare solo Gumbel, invece di testare un pool di copule e selezionare quella col miglior AIC o BIC?

- Perché $p_{ij} = C_{\hat{\vartheta}}(u_i, v_j)$ e non $p_{ij} = c_{\hat{\vartheta}}(u_i, v_j)$?

$$c_\vartheta(u,v) = \frac{C_\vartheta(u,v)}{u\,v}\,(t_\vartheta(u)+t_\vartheta(v))^{-2+2/\vartheta} \cdot (\ln u \ln v)^{\vartheta-1}(1+(\vartheta-1)(t_\vartheta(u)+t_\vartheta(v))^{-1/\vartheta})$$

$$t_\vartheta(w) = (-\ln w)^\vartheta$$

- Perché normalizzare $\mathtt{S}$ per riga, invece che $\mathtt{S} \leftarrow \mathtt{P}/\sum_{ij} p_{ij}$?
- Perché inizializzare $\mathtt{X}^{(0)}$ t.c. $\sum_j x_{ij}^{(0)} = x_{i\bullet}$, invece che $\mathtt{X}^{(0)} \leftarrow \mathtt{S}\sum_i x_{i\bullet}$?
- Perché usare solo Gumbel, invece di testare un pool di copule e selezionare quella col miglior AIC o BIC?

Dipendenze di grado

Copula-based modeling of degree-correlated networks

Assortatività I nodi con un alto grado tendono a connettersi con altri nodi che hanno anch'essi un alto grado. Allo stesso modo, i nodi con un basso grado tendono a connettersi con altri nodi a basso grado.

Disassortatività In questo scenario, i nodi con un alto grado tendono a connettersi con nodi che hanno un basso grado, e viceversa.

Il fattore di Newman r è definito come il coefficiente di correlazione di Pearson dei gradi dei due vertici alle estremità di un arco scelto a caso nella rete:

- $r > 0$ indica una tendenza assortativa
- $r < 0$ indica una tendenza disassortativa
- $r \approx 0$ suggerisce che non vi sono preferenze nella connessione basata sul grado

Idea: utilizzare le copule per catturare la struttura di dipendenza tra i gradi dei nodi adiacenti e generare ensemble di reti che rispecchino tale dipendenza

Modello per le dipendenze di grado

Copula-based modeling of degree-correlated networks

- Consideriamo la distribuzione del numero di connessioni alle due estremità di un arco, incluso l'arco stesso: ad ogni arco è associata cioè una coppia di v.a. $H{=}(I,J)$ con distribuzione $P(i,j) \equiv P(I=i, J=j)$ indicante la probabilità che un arco selezionato a caso insista su una coppia di nodi con grado i e j.
- Nel caso di una rete non direzionata, la distribuzione $P(i,j)$ è simmetrica e possiamo ricavare la distribuzione marginale ai capi degli archi come

$$P_e(k) = \sum_i P(i,k)$$

- Poiché ogni nodo di grado k ha probabilità di grado $P_\kappa(k)$, la probabilità di selezionare un arco connesso ad un nodo di grado k è proporzionale a $kP_\kappa(k)$
- Normalizzando si ottiene quindi la relazione (con $\overline{k}$ grado medio della rete)

$$P_e(k) = \frac{kP_\kappa(k)}{\sum_k kP_\kappa(k)} = \frac{kP_\kappa(k)}{\overline{k}} \qquad \Longleftrightarrow \qquad P_\kappa(k) = \frac{\overline{k}}{k} P_e(k)$$

Modello per le dipendenze di grado

Copula-based modeling of degree-correlated networks

- Consideriamo la distribuzione del numero di connessioni alle due estremità di un arco, incluso l'arco stesso: ad ogni arco è associata cioè una coppia di v.a. $H{=}(I{,}J)$ con distribuzione $P(i,j) \equiv P(I=i, J=j)$ indicante la probabilità che un arco selezionato a caso insista su una coppia di nodi con grado i e j.
- Nel caso di una rete non direzionata, la distribuzione $P(i,j)$ è simmetrica e possiamo ricavare la distribuzione marginale ai capi degli archi come

$$P_e(k) = \sum_i P(i,k)$$

- Poiché ogni nodo di grado k ha probabilità di grado $P_\kappa(k)$, la probabilità di selezionare un arco connesso ad un nodo di grado k è proporzionale a $kP_\kappa(k)$
- Normalizzando si ottiene quindi la relazione (con $\overline{k}$ grado medio della rete)

$$P_e(k) = \frac{kP_\kappa(k)}{\sum_k kP_\kappa(k)} = \frac{kP_\kappa(k)}{\overline{k}} \quad \Longleftrightarrow \quad P_\kappa(k) = \frac{\overline{k}}{k} P_e(k)$$

Modello per le dipendenze di grado

Copula-based modeling of degree-correlated networks

- Consideriamo la distribuzione del numero di connessioni alle due estremità di un arco, incluso l'arco stesso: ad ogni arco è associata cioè una coppia di v.a. $H{=}(I,J)$ con distribuzione $P(i,j) \equiv P(I=i, J=j)$ indicante la probabilità che un arco selezionato a caso insista su una coppia di nodi con grado i e j.
- Nel caso di una rete non direzionata, la distribuzione $P(i,j)$ è simmetrica e possiamo ricavare la distribuzione marginale ai capi degli archi come

$$P_e(k) = \sum_i P(i,k)$$

- Poiché ogni nodo di grado k ha probabilità di grado $P_\kappa(k)$, la probabilità di selezionare un arco connesso ad un nodo di grado k è proporzionale a $kP_\kappa(k)$
- Normalizzando si ottiene quindi la relazione (con $\overline{k}$ grado medio della rete)

$$P_e(k) = \frac{kP_\kappa(k)}{\sum_k kP_\kappa(k)} = \frac{kP_\kappa(k)}{\overline{k}} \qquad \Longleftrightarrow \qquad P_\kappa(k) = \frac{\overline{k}}{k} P_e(k)$$

Modello per le dipendenze di grado

Copula-based modeling of degree-correlated networks

- Consideriamo la distribuzione del numero di connessioni alle due estremità di un arco, incluso l'arco stesso: ad ogni arco è associata cioè una coppia di v.a. $H{=}(I,J)$ con distribuzione $P(i,j) \equiv P(I=i, J=j)$ indicante la probabilità che un arco selezionato a caso insista su una coppia di nodi con grado i e j.
- Nel caso di una rete non direzionata, la distribuzione $P(i,j)$ è simmetrica e possiamo ricavare la distribuzione marginale ai capi degli archi come

$$P_e(k) = \sum_i P(i,k)$$

- Poiché ogni nodo di grado k ha probabilità di grado $P_\kappa(k)$, la probabilità di selezionare un arco connesso ad un nodo di grado k è proporzionale a $kP_\kappa(k)$
- Normalizzando si ottiene quindi la relazione (con $\overline{k}$ grado medio della rete)

$$P_e(k) = \frac{kP_\kappa(k)}{\sum_k kP_\kappa(k)} = \frac{kP_\kappa(k)}{\overline{k}} \qquad \Longleftrightarrow \qquad P_\kappa(k) = \frac{\overline{k}}{k} P_e(k)$$

$P(i,j)$ tramite copule: procedura I

Copula-based modeling of degree-correlated networks

Nota 1: $P(u_1 \leq U < u_2, v_1 \leq V < v_2) = C(u_1, v_1) + C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1)$

Nota 2: il numero di connessioni all'estremo di un arco è compreso in $[h_{\min}, h_{\max}]$

Procedura I

- $P(i,j)$ è definita da una cdf a priori $F(x)$ *continua* secondo $P(i,j) = C(F(i), F(j)) - C(F(i-1), F(j)) + C(F(i-1), F(j-1)) - C(F(i), F(j-1))$
- La matrice viene poi normalizzata rispetto a $T = \sum_{i,j} P(i,j)$ in modo che sia a somma 1 su $[h_{\min}, h_{\max}] \times [h_{\min}, h_{\max}]$

$$P(i,j) \leftarrow P(i,j)/T$$

$$P_e(k) = \sum_{i=h_{\min}}^{h_{\max}} P(i,k)$$

Nota 1: $P(u_1 \leq U < u_2, v_1 \leq V < v_2) = C(u_1, v_1) + C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1)$

Nota 2: il numero di connessioni all'estremo di un arco è compreso in $[h_{\min}, h_{\max}]$

Procedura II

- In questo caso $P_e(h)$ è fornita in ingresso
- Costruita la cdf "discreta" di $P_e(h)$ come $G(h) = \sum_{j=h_{\min}}^{h} P_e(j)$, si ha $P(i,j) = C(G(i), G(j)) - C(G(i-1), G(j)) + C(G(i-1), G(j-1)) - C(G(i), G(j-1))$
- $P(i,j)$ può poi essere troncata a $h_{\max}$ come nella Procedura I.

Procedura III

$P_e(h)$ è fornita in ingresso ma *obbligata a priori* ad essere troncata in $h_{\max}$.

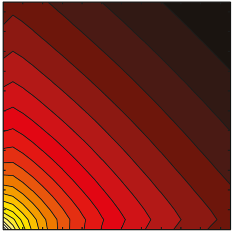
$P(i,j)$ tramite copule: procedure a confronto

Copula-based modeling of degree-correlated networks

Nota 1: copula di Gumbel con $\vartheta = 2$

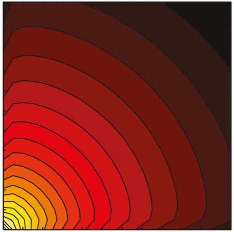
Nota 2: Il numero di connessioni all'estremo di un arco varia in $[h_{\text{min}} = 2, h_{\text{max}} = 25]$

P(i,j) Proc. I



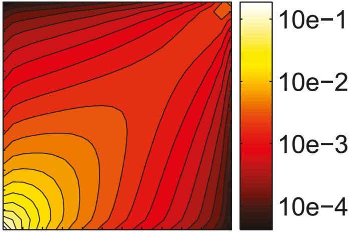
Distribuzione di Pareto
$F(h) = 1 - (h-1)^{-\gamma}$
$(\gamma = 0.7)$

P(i,j) Proc. II



Distribuzione di Zipf
$P_e(h) = \frac{h^{-\gamma}}{\sum_{h_{\text{min}}}^{\infty} h^{-\gamma}}$
$(\gamma = 2)$

P(i,j) Proc. III



Distribuzione di Zipf troncata a $h_{\text{max}} = 25$
$P_e(h) = \frac{h^{-\gamma}}{\sum_{h_{\text{min}}}^{h_{\text{max}}} h^{-\gamma}}$

Benchmark: rete di interazione proteina-proteina del lievito: $n = 1846$ nodi (le proteine), $E = 2203$ archi (0 self-loop). La rete ha grado massimo 56 e mostra disassortatività

Verifica: $P_\vartheta(i,j)$ tramite copula C_ϑ vs $\hat{P}(i,j)$ empirica

$$\hat{P}(i,j) = \begin{cases} m(i,j)/2E & (i \neq j) \\ m(i,j)/E & (i = j) \end{cases} \qquad (m(i,j) \ \# \text{ di coppie } (i,j) \text{ rilevate})$$

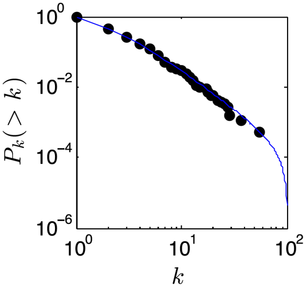
- best-fit per la distribuzione di grado $P_\kappa(k) \quad \Rightarrow \quad P_e(h) = P_\kappa(h)h/\overline{k}$
- best-fit su varie copule C_ϑ massimizzando la log-Likelihood

$$\sum_{h=1}^{h_{\max}} \sum_{h'=1}^{h_{\max}} m(h,h') \ln P_\vartheta(h,h')$$

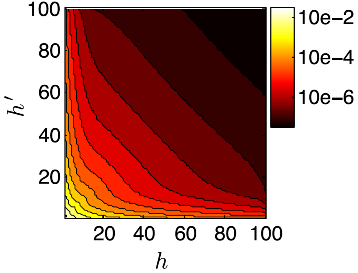
Verifica sperimentale: risultati

Copula-based modeling of degree-correlated networks

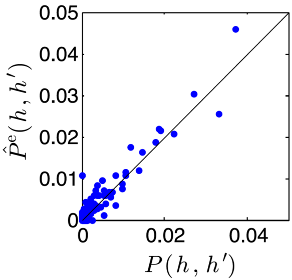
Stima $P_\kappa(k)$



Stima $P_\vartheta(i,j)$



$P_\vartheta(i,j)$ vs $\hat{P}(i,j)$



Distribuzione generalizzata di Pareto

$C(u,v) = \sum_i \lambda_i C_i(u,v)$
$i \in \{\text{Gauss, Clayton}\}$
$\sum_i \lambda_i = 1, \quad (\lambda_i > 0)$

Dato un generico nodo j, di grado k_j, da connettere, la matrice $P(i,j)$ interviene nella *probabilità* $P_i(i|k_j)$ *di assegnare il nodo* j *ad un altro nodo* i *(con grado* k_i*)*:

$$P_i(i|k_j) = r_{k_i} \frac{P(k_j|k_i)}{\sum_{m=1}^{n} P(k_j|k_m)}, \qquad \text{con } P(k|k') = P(k,k')/P_e(k')$$

dove r_{k_i} rappresenta il numero di archi sconnessi (stub) rimanenti del nodo i.

Generazione di una rete data $P(i,j)$

Copula-based modeling of degree-correlated networks

Dato un generico nodo j, di grado k_j, da connettere, la matrice $P(i,j)$ interviene nella *probabilità $P_i(i|k_j)$ di assegnare il nodo j ad un altro nodo i (con grado k_i)*:

$$P_i(i|k_j) = r_{k_i} \frac{P(k_j|k_i)}{\sum_{m=1}^{n} P(k_j|k_m)}, \qquad \text{con } P(k|k') = P(k,k')/P_e(k')$$

dove r_{k_i} rappresenta il numero di archi sconnessi (stub) rimanenti del nodo i.

1. Generazione casuale di n realizzazioni di K, tratte dalla distribuzione di probabilità di grado $P_\kappa(k)$, t. c. $\sum_{m=1}^{n} k_m$ sia pari.
2. Selezione casuale di un nodo con almeno uno stub rimanente e grado k'.
3. Assegnazione, secondo la probabilità di assegnamento $P_i(i|k')$, del nodo selezionato a un nodo i con grado k_i, con almeno uno stub rimanente e che non sia ancora connesso al nodo selezionato
4. I due stub selezionati vengono collegati per formare l'arco
5. Se sono presenti nodi con stub rimanenti, tornare al passaggio 2

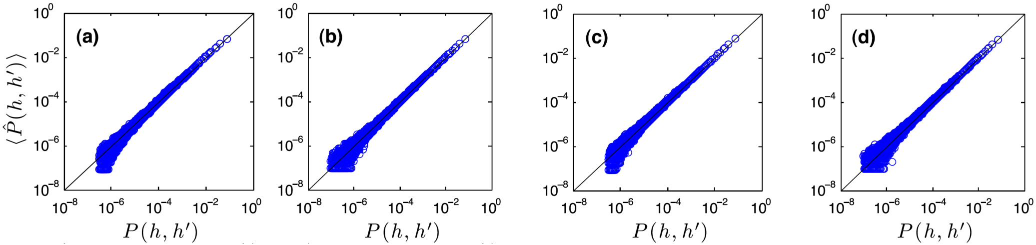
Generazione di una rete data $P(i,j)$: verifica sperimentale

Copula-based modeling of degree-correlated networks

Setup simulazioni

- Distribuzione di grado $P_\kappa(k) = h^{-1.5}$ $\Rightarrow$ $P_e(h) = h^{-1.5} / \sum_{h_{\min}}^{h_{\max}} h^{-1.5}$
- Copula $C(u,v) = uv(1 - a(1-u)(1-v))$
- $n = 5k$ nodi; $h_{\min} = 1$, $h_{\max} = 100$; 1000 realizzazioni

$P_\vartheta(i,j)$ vs $\langle \hat{P}(i,j) \rangle$



$a = -0.5$, reti connesse e assortative

$a = 0.5$, reti connesse e disassortative

$a = 0.5$, reti assortative

$a = -0.5$, reti disassortative

Conclusioni

- Finora le copule non sono state impiegate molto nella teoria delle reti, ma la loro capacità di modellare le dipendenze tra v.a. nelle code le rende degli strumenti matematici interessanti da usare in questo settore
- Ad esempio il loro impiego ha il vantaggio rispetto al Configuration Model di controllare esplicitamente le correlazioni tra i gradi dei nodi
- Purtroppo in uno dei pochi lavori che ha fatto utilizzo delle copule (Baral & Fique), esse non non sono state impiegate correttamente, per cui i risultati finora prodotti ed analizzati non sono da ritenere effettivamente validi
- In ogni caso il lavoro di Baral e Fique rientra negli *early works* sulla ricostruzione delle reti, ed anche la sua corretta formulazione probabilmente non potrebbe competere con lo stato dell'arte, dal momento che resta comunque un approccio molto simile a Max-Ent
- Ciò nonostante, sarebbe lo stesso interessante confrontare le prestazioni di Max-Ent con la rivisitazione corretta dell'approccio basato su copule

Riferimenti bibliografici

Copule

- B. Choroś, R. Ibragimov, E. Permiakova. Copula estimation. In P. Jaworski, F. Durante, W. Karl Härdle, and T. Rychlik, editors, *Copula Theory and Its Applications*, 2010. Springer Berlin Heidelberg.
- U. Schepsmeier, J. Stöber. Derivatives and fisher information of bivariate copulas. *Statistical Papers*, 55(2):525–542, 2014.

Ricostruzione di reti

- Kartik Anand, K. et al. The missing links: A global study on uncovering financial network structures from partial data. *Journal of Financial Stability*, 35:107–119, 2018.
- A. Bachem, B. Korte. On the RAS-algorithm. *Computing*, 23:189–198, 1979.
- **P. Baral, J. P. Fique.** Estimation of bilateral exposures: A copula approach. In *1st annual CIRANO workshop on networks in trade and finance*, 2012.
- C. Upper, A. Worms. Estimating bilateral exposures in the german interbank market: Is there a danger of contagion? *European Economic Review*, 48(4):827–849, 2004.
- H.-C. Xu, Z.-Y. Wang, F. Jawadi, W.-X. Zhou. Reconstruction of international energy trade networks with given marginal data: A comparative analysis. *Chaos, Solitons & Fractals*, 167:113031, 2023.

Generazione di reti

- M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701, 2002.
- M. Raschke, M. Schläpfer, R. Nibali. Measuring degree-degree association in networks. *Phys. Rev. E*, 82:037102, 2010.
- **M. Raschke, M. Schläpfer, K. Trantopoulos.** Copula-based modeling of degree-correlated networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(2):P02019, 2014.
- S. Weber, M. Porto. Generation of arbitrarily two-pointcorrelated random networks. *Phys. Rev. E*, 76:046111, 2007.