Econophysics - written exam 26 June 2012 (with solutions)

Docent: Dr. Diego Garlaschelli

You should return these pages completed. Please write your answers to all questions by filling the dotted spaces and by ticking the boxes. Motivate each of your answers in the additional sheets you will hand in. Answers are ignored if there is no explanation in the extra sheets. Number all sheets and all answers using the numbers given to the questions in these pages. Please write clearly and in English. On every sheet you hand in, please write your name and student ID. Don't forget to do the same on these pages. Good luck!

NAME: . . . .

SURNAME: . . . .

STUDENT ID: . . . .

PROBLEM 1

Consider a very simple market with only 3 stocks, and denote the corresponding time series of log-returns as $x_1(t)$, $x_2(t)$, $x_3(t)$ $(t = 1, \dots T)$. Moreover, as a simple Index of the market, consider the quantity

$$I(t) \equiv x_1(t) + x_2(t) + x_3(t) \qquad \forall t$$

The time variances of the 3 stocks and of the Index are measured, over the entire period $[1, T]$, as

$$Var[x_1] = 1.2 \qquad Var[x_2] = 2.3 \qquad Var[x_3] = 0.7 \qquad Var[I] = 4.8$$

Question 1.1

On the basis of the above measurements, what can be concluded about the correlation among the 3 stocks?

☐ The stocks are all mutually uncorrelated;

☐ At least 1 pair of stocks is correlated;

☐ At least 2 pairs of stocks are correlated;

☐ All the 3 pairs of stocks are correlated;

☐ Nothing can be concluded.

Question 1.2

Given the logical relationship between statistical correlation and probabilistic dependence, what can be concluded about the (in)dependence among the 3 stocks?

☐ All the 3 pairs of stocks are dependent;

☐ At most 1 pair of stocks is independent;

☐ At most 2 pairs of stocks are independent;

☐ All the 3 pairs of stocks are independent;

☐ Nothing can be concluded.

Question 1.3

The second moments of the log-returns of stocks 1, 3 and of the Index are measured as

$$\overline{x_1^2} = 1.36 \qquad \overline{x_3^2} = 0.95 \qquad \overline{I^2} = 4.84$$

Find all the average log-returns $\overline{x_1}$, $\overline{x_2}$, $\overline{x_3}$, and $\overline{I}$ knowing that $\overline{x_1} > 0$, $\overline{x_3} > 0$, and $\overline{I} < 0$.

$\overline{x_1} = \dots\dots\dots$ $\qquad$ $\overline{x_2} = \dots\dots\dots$ $\qquad$ $\overline{x_3} = \dots\dots\dots$ $\qquad$ $\overline{I} = \dots\dots\dots$

Question 1.4

Knowing that

$$\overline{x_1 x_2} = 0.06 \qquad \overline{x_1 x_3} = 0.2$$

find the 3×3 covariance matrix among the three stocks.

$$C \equiv \begin{pmatrix} \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots \end{pmatrix}$$

Question 1.5

Find the 3×3 correlation matrix among the three stocks.

$$\rho \equiv \begin{pmatrix} \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots \end{pmatrix}$$

Question 1.6

Draw the (correlation-based) Minimum Spanning Tree connecting the three stocks.

PROBLEM 1 - SOLUTION

Recall that the *time variance* of a time series $x_i(t)$ is defined as

$$Var[x_i] \equiv \overline{x_i^2} - \overline{x_i}^2 \tag{1}$$

where $\overline{x_i} \equiv \sum_{t=1}^{T} x_i(t)/T$ is the *first moment* of $x_i(t)$ and $\overline{x_i^2} \equiv \sum_{t=1}^{T} x_i^2(t)/T$ is the *second moment.* Also recall that the *time covariance* between two time series $x_i(t)$ and $x_j(t)$ is defined as

$$Cov[x_i, x_j] \equiv \overline{x_i x_j} - \overline{x_i} \cdot \overline{x_j} \tag{2}$$

where $\overline{x_i x_j} \equiv \sum_{t=1}^{T} x_i(t)x_j(t)/T$. Finally, recall that the variance of a sum of N random variables $x_1, \ldots x_N$ is the sum of the covariances between all possible pairs of the N variables:

$$Var\left[\sum_{i=1}^{N} x_i\right] = \sum_{i=1}^{N}\sum_{j=1}^{N} Cov[x_i, x_j] = \sum_{i=1}^{N} Var[x_i] + 2\sum_{i=1}^{N}\sum_{j<i} Cov[x_i, x_j]$$

Answer to question 1.1

In this particular case,

$$\begin{aligned} Var\,[I] &= Var\,[x_1 + x_2 + x_3] \\ &= Var[x_1] + Var[x_2] + Var[x_3] + 2(Cov[x_1, x_2] + Cov[x_1, x_3] + Cov[x_2, x_3]) \end{aligned} \tag{3}$$

and the measured values imply that

$$\begin{aligned} Cov[x_1, x_2] + Cov[x_1, x_3] + Cov[x_2, x_3] &= \frac{Var\,[I] - Var[x_1] - Var[x_2] - Var[x_3]}{2} \\ &= \frac{4.8 - 1.2 - 2.3 - 0.7}{2} = 0.3 \end{aligned} \tag{4}$$

This implies that at least one of the covariances must be different from zero, i.e. *at least one pair of stocks is correlated.*

Answer to question 1.2

Recall that if two random variables x and y are independent (i.e. their joint probability $P(x, y)$ is the product $P_x(x) \cdot P_y(y)$ of their two marginal probabilities) they are also uncorrelated (i.e. their covariance and correlation are zero), while the inverse is in general not true (dependent variables can be uncorrelated). This implies that if two random variables are correlated (nonzero covariance and nonzero correlation) they are necessarily not independent. Since in this case at least 1 pair of stocks are correlated, we can conclude that at least 1 pair of stocks are dependent, i.e. *at most 2 pairs of stocks are independent.*

Answer to question 1.3

From eq.(1) it follows that

$$\overline{x_i} = \pm\sqrt{\overline{x_i^2} - Var[x_i]}$$

Selecting the signs as given in the text, we have

$$\overline{x_1} = +\sqrt{\overline{x_1^2} - Var[x_1]} = +\sqrt{1.36 - 1.2} = 0.4$$

$$\overline{x_3} = +\sqrt{\overline{x_3^2} - Var[x_3]} = +\sqrt{0.95 - 0.7} = 0.5$$

$$\overline{I} = -\sqrt{\overline{I^2} - Var[I]} = -\sqrt{4.84 - 4.8} = -0.2$$

Then, since $\overline{I} = \overline{x_1} + \overline{x_2} + \overline{x_3}$, we have

$$\overline{x_2} = \overline{I} - \overline{x_1} + \overline{x_3} = -0.2 - 0.4 - 0.5 = -1.1$$

Answer to question 1.4

From eq.(2) we have

$$Cov[x_1, x_2] = \overline{x_1 x_2} - \overline{x_1} \cdot \overline{x_2} = 0.06 - 0.4 \cdot (-1.1) = +0.5$$

$$Cov[x_1, x_3] = \overline{x_1 x_3} - \overline{x_1} \cdot \overline{x_3} = 0.2 - 0.4 \cdot 0.5 = 0$$

The missing covariance $Cov[x_2, x_3]$ can be found from eq.(4):

$$Cov[x_2, x_3] = 0.3 - Cov[x_1, x_2] - Cov[x_1, x_3] = -0.2$$

Therefore the 3×3 covariance matrix is

$$C \equiv \begin{pmatrix} Cov[x_1,x_1] & Cov[x_1,x_2] & Cov[x_1,x_3] \\ Cov[x_2,x_1] & Cov[x_2,x_2] & Cov[x_2,x_3] \\ Cov[x_3,x_1] & Cov[x_3,x_2] & Cov[x_3,x_3] \end{pmatrix} = \begin{pmatrix} Var[x_1] & Cov[x_1,x_2] & Cov[x_1,x_3] \\ Cov[x_1,x_2] & Var[x_2] & Cov[x_2,x_3] \\ Cov[x_1,x_3] & Cov[x_2,x_3] & Var[x_3] \end{pmatrix}$$
$$= \begin{pmatrix} 1.2 & 0.5 & 0 \\ 0.5 & 2.3 & -0.2 \\ 0 & -0.2 & 0.7 \end{pmatrix}$$

where we have used the symmetry property $Cov[x_i, x_j] = Cov[x_j, x_i]$ and the fact that $Cov[x_i, x_i] = Var[x_i]$.

Answer to question 1.5

Recall that the *correlation coefficient* between two random variables x_i and x_j is defined as

$$\rho_{i,j} \equiv \frac{Cov[x_i, x_j]}{\sqrt{Var[x_i]Var[x_j]}}$$

Therefore the 3×3 correlation matrix can be obtained from the covariance matrix and reads

$$\rho \equiv \begin{pmatrix} \rho_{1,1} & \rho_{1,2} & \rho_{1,3} \\ \rho_{2,1} & \rho_{2,2} & \rho_{2,3} \\ \rho_{3,1} & \rho_{3,2} & \rho_{3,3} \end{pmatrix} = \begin{pmatrix} 1 & \rho_{1,2} & \rho_{1,3} \\ \rho_{1,2} & 1 & \rho_{2,3} \\ \rho_{1,3} & \rho_{2,3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.3 & 0 \\ 0.3 & 1 & -0.16 \\ 0 & -0.16 & 1 \end{pmatrix}$$

where we have used the facts that $\rho_{i,j} = \rho_{j,i}$ and $\rho_{i,i} = 1$.

Answer to question 1.6

The correlation-based Minimum Spanning Tree (MST) connecting the three stocks must contain 2 links (not 1 otherwise it wouldn't be spanning, and not 3 otherwise it wouldn't be a tree). Those 2 links correspond to the 2 largest correlation coefficients. So in this case the MST contains the links (1,2) and (1,3). Note that, curiously, one of the two pairs of stocks is actually uncorrelated ($\rho_{1,3} = 0$). Still, the MST prescription requires to include it.

PROBLEM 2

In a board-interlock network, N firms are connected by L links. Each link indicates that there are common directors sitting in the boards of the two connected firms. It is proposed that the observed structure of the network can be reproduced by a *fitness model* where each pair of firms i and j is connected with probability

$$p_{ij} = x_i + x_j$$

where the *fitness* $x_i \in [x_{min}, x_{max}]$ is an intrinsic property of firm i, assumed to be proportional to the number of directors in its board. This model predicts that the expected degree of firm i is

$$k_i = \sum_{j \neq i} p_{ij} \approx \sum_{j=1}^{N} p_{ij}$$

Question 2.1

If $\overline{x} \equiv \sum_{i=1}^{N} x_i/N$ is the average fitness in the network, write the expected degree of firm i as a function of N, x_i and $\overline{x}$ using the approximation above.

$$k_i = \ldots\ldots\ldots$$

Question 2.2

Write the average expected degree $\overline{k} \equiv \sum_{i=1}^{N} k_i/N$ of the network as a function of N and $\overline{x}$. Then, using the fact that $\overline{k} = 2L/N$, write $\overline{x}$ as a function of N and L alone.

$$\overline{x} = \ldots\ldots\ldots$$

Question 2.3

Write the minimum and maximum possible values $x_{min} = 0$ and x_{max} (assuming that these extremes are the same for all x_i's) required in order to ensure that p_{ij} is a probability, under the model considered. Then write the minimum and maximum values k_{min} and k_{max} for all degrees in the network predicted by the model, as a function of N and L alone. *(Note: these extreme values are not necessarily realized, they represent lower and upper bounds for each vertex in the network.)*

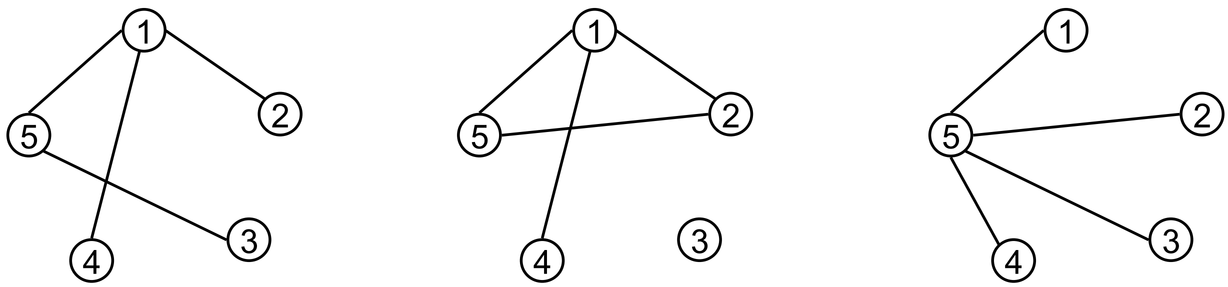
$$x_{min} = \ldots\ldots\ldots \qquad x_{max} = \ldots\ldots\ldots$$

$$k_{min} = \ldots\ldots\ldots \qquad k_{max} = \ldots\ldots\ldots$$

Question 2.4

Use the last result to determine which of the following networks is consistent with the proposed model. *(Strictly speaking, since k_{min} and k_{max} are expected quantities, they can be violated in individual realizations; however they are the most likely bounds, so the question refers to the most likely situation for each network.)*

☐ Network 1 ☐ Network 2 ☐ Network 3



Network 1 Network 2 Network 3

Question 2.5

Rewrite k_i as a function of x_i, L, and N alone. Invert this relation to find x_i as a function of k_i, L, and N alone. Then, for the network consistent with the model, write the values of the fitness x_i for all vertices. *(Again, since k_i is an expected quantity, it might differ from its realized value in the network; so strictly speaking the question refers to the most likely values of x_i.)*

$k_i = \dots\dots\dots$ $x_i = \dots\dots\dots$

$x_1 = \dots\dots\dots$ $x_2 = \dots\dots\dots$ $x_3 = \dots\dots\dots$ $x_4 = \dots\dots\dots$ $x_5 = \dots\dots\dots$

Question 2.6

For the network consistent with the model, find the ratio r_1 between the number of directors of the least connected firm and the number of directors of the most connected firm, as predicted by the model. Then write the ratio r_2 between the number of directors of the most connected firm and the total number of directors in all boards of the network (counting multiple times the shared directors). Note that you are obtaining this information using only the knowledge of the topology of the network!

$r_1 = \dots\dots\dots$ $r_2 = \dots\dots\dots$

PROBLEM 2 - SOLUTION

Answer to question 2.1

The expected degree of firm i is

$$k_i \approx \sum_{j=1}^{N} p_{ij} = \sum_{j=1}^{N}(x_i + x_j) = N(x_i + \overline{x}) \tag{5}$$

Answer to question 2.2

The average expected degree in the network is

$$\overline{k} \equiv \frac{\sum_{i=1}^{N} k_i}{N} = \frac{\sum_{i=1}^{N} N(x_i + \overline{x})}{N} = 2N\overline{x} \tag{6}$$

Equating the above expression with $\overline{k} = 2L/N$, we find

$$\overline{x} = \frac{\overline{k}}{2N} = \frac{L}{N^2} \tag{7}$$

Answer to question 2.3

In order to ensure that $p_{ij} = x_i + x_j$ is a probability, we need to enforce $0 \leq p_{ij} \leq 1\ \forall i$. Since the extreme values for the fitness are assumed to be the same for all vertices, this is only guaranteed if the minimum and maximum possible fitness values are

$$x_{min} = 0 \qquad x_{max} = 1/2$$

From eq.(5), it then follows that the minimum and maximum values for all degrees in the network, as predicted by the model, are

$$k_{min} = N\overline{x} = \frac{L}{N} \qquad k_{max} = N\overline{x} + \frac{N}{2} = \frac{L}{N} + \frac{N}{2}$$

Answer to question 2.4

Since in all the three networks $N = 5$ and $L = 4$, the minimum and maximum values for the degree predicted by the model are always

$$k_{min} = \frac{4}{5} = 0.8 \qquad k_{max} = \frac{4}{5} + \frac{5}{2} = 3.3$$

Only in the first network the degrees of all vertices are consistent with the above predicted bounds, therefore only the first network is consistent with the proposed model (in the most likely situation).

Answer to question 2.5

Using eq.(7), we can rewrite eq.(5) as

$$k_i = Nx_i + \frac{L}{N}$$

which is a function of x_i, L, and N alone. We then invert this relation to find

$$x_i = \frac{k_i}{N} - \frac{L}{N^2}$$

which is a function of k_i, L, and N alone. For the only network consistent with the model, the values of the fitness x_i are then

$$x_1 = \frac{11}{25} \qquad x_2 = x_3 = x_4 = \frac{1}{25} \qquad x_5 = \frac{6}{25}$$

Answer to question 2.6

Since in the model it is assumed that the fitness is proportional to the number of directors in the board of a firm, the ratio r_1 between the number of directors of the least connected firm (which is any one of the vertices $2, 3, 4$) and the number of directors of the most connected firm (which is vertex 1) is simply

$$r_1 = \frac{x_2}{x_1} = \frac{x_3}{x_1} = \frac{x_4}{x_1} = \frac{1}{11}$$

(note that, even if k_1 is only 3 times larger than k_2, the board of firm 1 is 11 times larger than that of firm 2). Similarly, the ratio r_2 between the number of directors of the most connected firm (vertex 1) and the total number of directors (counting multiple times the shared ones) in all boards of the network is

$$r_2 = \frac{x_1}{x_1 + x_2 + x_3 + x_4 + x_5} = \frac{11}{11 + 1 + 1 + 1 + 6} = \frac{11}{20}$$