STAT0017
SELECTED TOPICS IN STATISTICS

Topic 2: "Dependence modelling in finance using copulas"
In-Course Assessment 2 (ICA2)

Hand-in Date: Wednesday 1st May 2019
Hand-in Time: 4pm

Submission rules

The ICA2 has two sections: *Section A* is for group work and group submission; *Section B* is for individual work and individual submission. Every student needs to hand in their solutions to *Section B* individually. Every group of students needs to submit solutions to *Section A*, clearly indicating the student ID of all group members. A student can only participate in one group. The same ICA2 deadline holds for both *Section A* and *Section B*. *Section A* and *Section B* have the same weight for the overall mark. See below for more information on group submissions.

Declaration

I am aware of the UCL Statistical Science Departments regulations on plagiarism for assessed coursework. I have read the guidelines in the student handbook and understand what constitutes plagiarism.
I hereby affirm that the work I am submitting for *Section B* of this in-course assessment is entirely my own, and that the work for *Section A* submitted in my name is entirely produced by the indicated group members.

Student name: **Student ID number**:

Date: **Signature:**

TURN OVER

General submission rules and information

- What you hand in for this assessment should be your own work (see rules on group submission below). It is to be handed in by you to the Statistics Department Office by Wednesday 1st May 2019, 4 pm.
- Write your **student ID** number on a cover sheet. To allow anonymous marking, provide your **student ID** number at the top of each of the answer sheets, but not your name.
- The printed solutions for both sections *Section A* and *Section B* must contain no more than **20 A4 pages**, including graphs, tables, pictures and R code. Solutions should have reasonable margins and a font size no less than 10pt. Submissions exceeding the maximum length limit will be normally ignored in terms of marking. **This is a generous limit, which you shouldn't aim to reach: taking up more space will not gain more marks.**
- Your submitted work can be composed in any way that you find convenient (R markdown, LaTeX, Word). Handwriting can only be used for *Section B*.
- Late submission will incur a penalty unless there are extenuating circumstances supported by appropriate documentation. Penalties are set out in the Departmental Student Handbooks.
- Non-submission of in-course assessment may mean that your overall examination mark is recorded as non-complete, i.e. you might not obtain a pass for the course.
- Any plagiarism will normally result in zero marks for all students involved, and may also mean that your overall examination mark is recorded as non-complete. Guidelines as to what constitutes plagiarism may be found in the Departmental Student Handbooks.
- A feedback sheet will be returned to you. The course lecturer will keep your original submission, which the external examiner may wish to see, so make sure that you keep a copy of your work. You will receive a provisional grade – grades are provisional until confirmed by the Statistics Examiners Meeting in June 2019.

Rules for group submission of *Section A*

- A single set of solutions can be handed in by groups of **3 students**. If, for some reason, you wish to form a group of **2 people**, please inform the course lecturer as soon as possible.
- **Students alone are responsible for forming groups.** The course lecturer will neither decide about who works together, nor help with the organisation of such groups. The student ID number of each group member will be requested in due course.

CONTINUED

- **All members of a group will get the same mark for *Section A*** of this ICA2, except in exceptional circumstances (e.g. a member of a group did not contribute fairly to the project).
- Working together and discussing solutions *within groups* is **fine for *Section A* only** and the usual plagiarism regulations do not apply to this. However, they do apply to plagiarism of work of other groups or other sources. Do not discuss your work with students that are not members of your group. Do not show your work to students that are not members of your group.
- For *Section B*, plagiarism regulations apply to your individual work. **Do not** discuss your work on *Section B* with other students, not even those that are members of your *Section A* group. **Do not** show your work on *Section B* to other students, not even to those who are members of your *Section A* group.
- All group members must be indicated (by their **student ID** number) on all pages of the submission.

Section A

You have been hired as an analyst at an investment bank, and as part of your portfolio risk assessment you use the Value-at-Risk (VaR) method. The VaR estimates the maximum potential loss of a portfolio over a specified period of time with a particular confidence level, typically 95% or 99% level. You have been provided with the historical data on the following six financial stock indices:

1. `ftse100`: FTSE100 is a share index of the 100 companies listed on the London Stock Exchange with the highest market capitalisation.
2. `sp500`: S&P500 is an American stock market index based on the market capitalizations of 500 large companies.
3. `sse`: The Shanghai Stock Exchange (SSE) Composite Index is a stock market index made up of all the A-shares and B-shares that are traded at the Shanghai Stock Exchange.
4. `dax30`: DAX is a blue chip stock market index consisting of the 30 major German companies trading on the Frankfurt Stock Exchange.
5. `nikkei225`: Nikkei 225 is a price-weighted stock market index that includes the top 225 blue-chip companies listed on the Tokyo Stock Exchange.
6. `cac40`: CAC 40 is a stock market index that represents a capitalization weighted measure of the the performance of the 40 largest and most actively traded shares listed on Euronext Paris.

These are weekly data for the period February 25, 1999 to February 28, 2019, providing 1045 observations in total for each time series.

(a) Select three stock indices of your choice, and estimate the 99% and 95% Value-at-Risk of a portfolio log-returns consisting of these stock indices over the next week. You are required to compute the Value-at-Risk using the Monte Carlo simulation approach based on vine copula theory. When simulating asset returns, you may assume that all marginal models for asset returns are standard normal. That is, you may apply the inverse CDF of a standard normal distribution to simulated standard uniform random variables. [25]

Marks will not only be given for final (numerical) answers but also for:

- *the accuracy and clarity of the reasoning.*
- *correctly carrying out analyses, including relevant checks (e.g. distributional checks, autocorrelation, ARCH effects, etc.).*

- *explanation why particular marginal and copula models were selected.*
- *reasonable and accurate interpretation of estimation and simulation results.*
- *use of relevant descriptive statistics.*
- *use of adequately sized and captioned relevant figures and tables.*
- *producing well structured report, good presentation, spelling and grammar.*
- *clarity and readability of computer code.*

(b) Using all three stock indices from part (a) and the chosen vine copula model, carry out validation by simulation of the model. That is, investigate whether the simulated bivariate paired data have similar features to the observed standardised paired data. In particular, you are expected to comment on the pair of variables that was not explicitly modelled in the vine copula. (**Note**: You only need to consider data on $[0, 1]$). [5]

(c) In this part you are required to write the R code for the likelihood function of the bivariate Gaussian copula model. Select two stock indices of your choice, estimate the Gaussian copula parameter based on this likelihood function, and compare the resulting estimate with the parameter value estimated using function `BiCopEst`. (**Hint**: You may want to modify `normal_loglikelihood` function and use `BiCopPDF` function.) [10]

(d) Provide a literature review on applications of copula theory to compute the Value-at-Risk. The review should not exceed the limit of one page, and should include the following:

- definition of the VaR and its use.
- whether copula-based multivariate models perform well in the VaR estimation.
- the effects of misspecified marginals and copulas on the VaR estimation.
- time-varying copula models for the VaR estimation [10]

Marks will be given for:

- *clear insight into the subject area.*
- *a succinct survey of the relevant literature.*
- *discussion of results provided in the relevant studies.*
- *ability to select most important points from an analysis.*
- *accurate references.*

Section B

(a) Using the relation of Archimedean generators to Laplace-Stieltjes transforms of distribution functions on the positive real line $\mathbb{R}^+$, derive the inverse generator φ^{-1} of the Clayton copula. (**Hint**: You need to use the $Gamma(\frac{1}{\theta}, 1)$ distribution). [5]

(b) Using the inverse generator φ^{-1} derived in (a), construct the Clayton copula. (*You can provide the form of the Clayton copula in which parameter θ is restricted to* $(0, \infty)$). [5]

(c) [**BSc students only**] Derive the density of the Clayton copula. [5]

[**MSc/MSci students only**] Show that the Fréchet lower bound is not a copula for $d = 3$. (**Hint**: Compute a C-volume V_C on $[0.5, 1]^3$ and show that it is negative.) [5]

(d) Obtain the conditional distribution of the Clayton copula. [5]

(e) Using results from (d) outline an algorithm to simulate a pair of observations (u_1, u_2) from the Clayton copula using the conditional distribution approach. Hence, using this algorithm and the parameter of the Clayton copula $\theta = 3$, simulate manually a pair of numerical values rounded to 3 d.p. without using the `CDVine` or any other package. Also provide the 2 independent uniform random variables used in the simulation. [5]

(f) Consider the random vector $(Y_1, Y_2)'$ with the joint distribution function $F_\theta(y_1, y_2)$, and marginal distributions $F_1(y_1)$ and $F_2(y_2)$:

$$F_\theta(y_1, y_2) = \left(1 + e^{-y_1} + e^{-y_2} + (1-\theta)e^{-y_1-y_2}\right)^{-1}$$

where $y_1,\ y_2 \in \overline{\mathbb{R}} = [-\infty, \infty]$, $\theta \in [-1, 1]$.

i. Find the marginal distributions of Y_1 and Y_2. [2]

ii. Using the Inversion Method, construct the copula based on the joint CDF of (Y_1, Y_2), and show that it is given by:

$$C_\theta(u_1, u_2) = \frac{u_1 u_2}{1 - \theta(1-u_1)(1-u_2)}$$

. [3]

(g) Consider the following 6×6 R-Vine Matrix:

$$\begin{pmatrix} 1 & & & & & \\ 6 & 4 & & & & \\ 4 & 6 & 2 & & & \\ 5 & 5 & 6 & 3 & & \\ 3 & 2 & 5 & 6 & 6 & \\ 2 & 3 & 3 & 5 & 5 & 5 \end{pmatrix}$$

i. Provide a graphical representation of the R-vine tree sequence stored in the above R-Vine Matrix. [6]

ii. State whether the resulting vine structure is a C- or D-vine, or neither of these. (You need to justify your answer in order to get full marks.) [3]

iii. Indicate how many different possible C- and D-vine pair-copula decompositions there are. [4]

iv. Write down the 6-dimensional density $f(\mathbf{x})$ corresponding to a vine structure stored in the above 6×6 R-Vine Matrix. [7]

END OF PAPER