Homework 7, MATH 5261

Zongyi Liu

Wed, Oct 29, 2025

Github Repository Directory: `https://github.com/zongyiliu/STAT_5261/tree/main/Homework_7`

1 Question 1

Let C be a d-dimensional copula. Show that for any $(u_1, u_2, \ldots, u_d) \in [0,1]^d$

$$\max\left(\sum_{i=1}^{d} u_i - d + 1, 0\right) \leq C\left(u_1, u_2, \ldots, u_d\right) \leq \min\left(u_1, u_2, \ldots, u_d\right)$$

Answer

Let C be a d-dimensional copula. For any $\mathbf{u} = (u_1, u_2, \ldots, u_d) \in [0,1]^d$, the following bounds hold:

$$\max\left(\sum_{i=1}^{d} u_i - d + 1, 0\right) \leq C\left(u_1, u_2, \ldots, u_d\right) \leq \min\left(u_1, u_2, \ldots, u_d\right)$$

This is also known as Fréchet-Hoeffding bounds.

Let $U_1, U_2, \ldots, U_d$ be random variables uniformly distributed on $[0,1]$ such that their joint CDF is $C(\mathbf{u}) = P(U_1 \leq u_1, \ldots, U_d \leq u_d)$.

Part 1: the upper bound

We show $C(\mathbf{u}) \leq \min(u_1, u_2, \ldots, u_d)$ here. By the definition of a joint probability:

$$C(\mathbf{u}) = P\left(\bigcap_{i=1}^{d}\{U_i \leq u_i\}\right)$$

Since the intersection of events is always a subset of any individual event, for a fixed index $j \in \{1, \ldots, d\}$:

$$\bigcap_{i=1}^{d}\{U_i \leq u_i\} \subseteq \{U_j \leq u_j\}$$

Taking probabilities, we have:

$$P\left(\bigcap_{i=1}^{d}\{U_i \leq u_i\}\right) \leq P(U_j \leq u_j)$$

By the marginal property of the copula, $P(U_j \leq u_j) = u_j$. Thus, $C(\mathbf{u}) \leq u_j$ for all $j = 1, \ldots, d$. Therefore, $C(u_1, u_2, \ldots, u_d) \leq \min(u_1, u_2, \ldots, u_d)$.

Part 2: the lower bound

We show $C(\mathbf{u}) \geq \max\left(\sum_{i=1}^{d} u_i - d + 1, 0\right)$ here. Firstly, we use the relationship between the probability of the intersection and the union of complementary events:

$$C(\mathbf{u}) = P\left(\bigcap_{i=1}^{d}\{U_i \leq u_i\}\right) = 1 - P\left(\left(\bigcap_{i=1}^{d}\{U_i \leq u_i\}\right)^c\right)$$

By De Morgan's Laws, the complement is the union of the complements:

$$C(\mathbf{u}) = 1 - P\left(\bigcup_{i=1}^{d}\{U_i > u_i\}\right)$$

Now, we apply Boole's Inequality (Union Bound) to the union of the complementary events $\{U_i > u_i\}$:

$$P\left(\bigcup_{i=1}^{d}\{U_i > u_i\}\right) \leq \sum_{i=1}^{d}P(U_i > u_i)$$

The marginal probability is $P(U_i > u_i) = 1 - P(U_i \leq u_i) = 1 - u_i$. Substituting this into the inequality:

$$P\left(\bigcup_{i=1}^{d}\{U_i > u_i\}\right) \leq \sum_{i=1}^{d}(1 - u_i) = d - \sum_{i=1}^{d}u_i$$

Finally, substituting this back into the expression for $C(\mathbf{u})$:

$$C(\mathbf{u}) \geq 1 - \left(d - \sum_{i=1}^{d}u_i\right) = \sum_{i=1}^{d}u_i - d + 1$$

Since $C(\mathbf{u})$ is a probability (a value in $[0, 1]$), it must be non-negative. Therefore, the tightest lower bound must be the maximum of 0 and the result from Step A:

$$C(u_1, u_2, \ldots, u_d) \geq \max\left(\sum_{i=1}^{d}u_i - d + 1, 0\right)$$

Combining Part 1 and Part 2 we get the the Fréchet-Hoeffding bounds; in this case for copula, we have:

$$\max\left(\sum_{i=1}^{d}u_i - d + 1, \, 0\right) \ \leq \ C(u_1, \ldots, u_d) \ \leq \ \min(u_1, \ldots, u_d).$$

2   Question 2

$\mathbf{X} = (X_1, X_2)^T$ have a bivariate normal distribution with mean:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{and covariance} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

Find $c_{\mathbf{X}}(u_1, u_2)$, the pdf of the $C_{\mathbf{X}}(u_1, u_2)$, the copula corresponding to $\mathbf{X}$.

Answer

Let $\mathbf{X} = (X_1, X_2)^T$ follow a bivariate normal distribution $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$ with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{and covariance} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

The correlation parameter is ρ.

The density of the copula $C_{\mathbf{X}}$ is the ratio of the joint PDF $f_{\mathbf{X}}(x_1, x_2)$ to the product of the marginal PDFs $f_{X_1}(x_1)f_{X_2}(x_2)$:

$$c_{\mathbf{X}}(u_1, u_2) = \frac{f_{\mathbf{X}}(x_1, x_2)}{f_{X_1}(x_1)f_{X_2}(x_2)}$$

where the variables are linked by the standard normal quantile function Φ^{-1}:

$$z_1 = \frac{x_1 - \mu_1}{\sigma_1} = \Phi^{-1}(u_1) \quad \text{and} \quad z_2 = \frac{x_2 - \mu_2}{\sigma_2} = \Phi^{-1}(u_2)$$

Let $\phi_\rho(z_1, z_2)$ be the PDF of the standard bivariate normal distribution (with zero mean, unit variance, and correlation ρ), and $\phi(z)$ be the standard univariate normal PDF.

$$\text{Joint PDF: } f_{\mathbf{X}}(x_1, x_2) = \frac{1}{\sigma_1\sigma_2}\phi_\rho(z_1, z_2)$$

$$\text{Marginal PDFs: } f_{X_1}(x_1) = \frac{1}{\sigma_1}\phi(z_1) \quad \text{and} \quad f_{X_2}(x_2) = \frac{1}{\sigma_2}\phi(z_2)$$

Substituting these into the copula density formula:

$$c_{\mathbf{X}}(u_1, u_2) = \frac{\frac{1}{\sigma_1\sigma_2}\phi_\rho(z_1, z_2)}{\left(\frac{1}{\sigma_1}\phi(z_1)\right)\left(\frac{1}{\sigma_2}\phi(z_2)\right)} = \frac{\phi_\rho(z_1, z_2)}{\phi(z_1)\phi(z_2)}$$

Using the explicit formulas for ϕ_ρ and ϕ:

$$\phi_\rho(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\left[z_1^2 - 2\rho z_1 z_2 + z_2^2\right]\right)$$

$$\phi(z_1)\phi(z_2) = \left(\frac{1}{\sqrt{2\pi}}e^{-z_1^2/2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-z_2^2/2}\right) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(z_1^2 + z_2^2)\right)$$

Taking the ratio:

$$c_{\mathbf{X}}(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2) + \frac{1}{2}(z_1^2 + z_2^2)\right)$$

$$= \frac{1}{\sqrt{1-\rho^2}}\exp\left(\frac{1}{2(1-\rho^2)}\left[-(z_1^2 - 2\rho z_1 z_2 + z_2^2) + (1-\rho^2)(z_1^2 + z_2^2)\right]\right)$$

$$= \frac{1}{\sqrt{1-\rho^2}}\exp\left(\frac{1}{2(1-\rho^2)}\left[-z_1^2 + 2\rho z_1 z_2 - z_2^2 + z_1^2 - \rho^2 z_1^2 + z_2^2 - \rho^2 z_2^2\right]\right)$$

$$= \frac{1}{\sqrt{1-\rho^2}}\exp\left(\frac{1}{2(1-\rho^2)}\left[2\rho z_1 z_2 - \rho^2(z_1^2 + z_2^2)\right]\right)$$

The density of the Gaussian Copula is:

$$c_{\mathbf{X}}(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}}\exp\left(\frac{2\rho z_1 z_2 - \rho^2(z_1^2 + z_2^2)}{2(1-\rho^2)}\right)$$

where $z_i = \Phi^{-1}(u_i)$ for $i = 1, 2$.

3

3   Question 3

3.1   Problem 1

Consider the R code above.

(a) What type of copula model has been sampled? Give the copula family, the correlation matrix, and any other parameters that specify the copula.

(b) What is the sample size?

Answer
Part a
The code given for this question in book is as:

```
1    library(copula)
2    cop_t_dim3 = tCopula(dim = 3, param = c(-0.6,0.75,0),dispstr = "un", df = 1)
3    set.seed(5640)
4    rand_t_cop = rCopula(n = 500, copula = cop_t_dim3)
5    pairs(rand_t_cop)
6    cor(rand_t_cop)
```

- Copula family: The function tCopula is used, so the model is a Student's t Copula (or t-Copula).

- Dimension: `dim = 3`, so the copula is 3-dimensional.

- Correlation matrix ($\mathbf{P}$): The parameter vector is `param = c(-0.6,0.75,0)`, which specifies the three unique correlation coefficients $(\rho_{12}, \rho_{13}, \rho_{23})$ for a 3-dimensional copula. The correlation matrix $\mathbf{P}$ is:

$$\mathbf{P} = \begin{pmatrix} 1 & -0.6 & 0.75 \\ -0.6 & 1 & 0 \\ 0.75 & 0 & 1 \end{pmatrix}$$

- Other parameters: The degrees of freedom is df $= 1$.

Part b
The function `rCopula` is called with `n = 500`, thus the sample size is 500.

3.2 Problem 2

Examine the scatterplot matrix (generated by line 6) and answer the questions below. Include the scatterplot matrix with your answer.

(a) Components 2 and 3 are uncorrelated. Do they appear independent? Why or why not?

(b) Do you see signs of tail dependence? If so, where?

(c) What are the effects of dependence upon the plots?

(d) The nonzero correlations in the copula do not have the same values as the corresponding sample correlations. Do you think this is just due to random variation or is something else going on? If there is another cause besides random variation, what might that be? To help answer this question, you can get confidence intervals for the Pearson correlation: For example,

```
cor.test(rand_t_cop[,1], rand_t_cop[,3])
```

will give a confidence interval (95 percent by default) for the correlation (Pearson by default) between components 1 and 3. Does this confidence interval include 0.75?
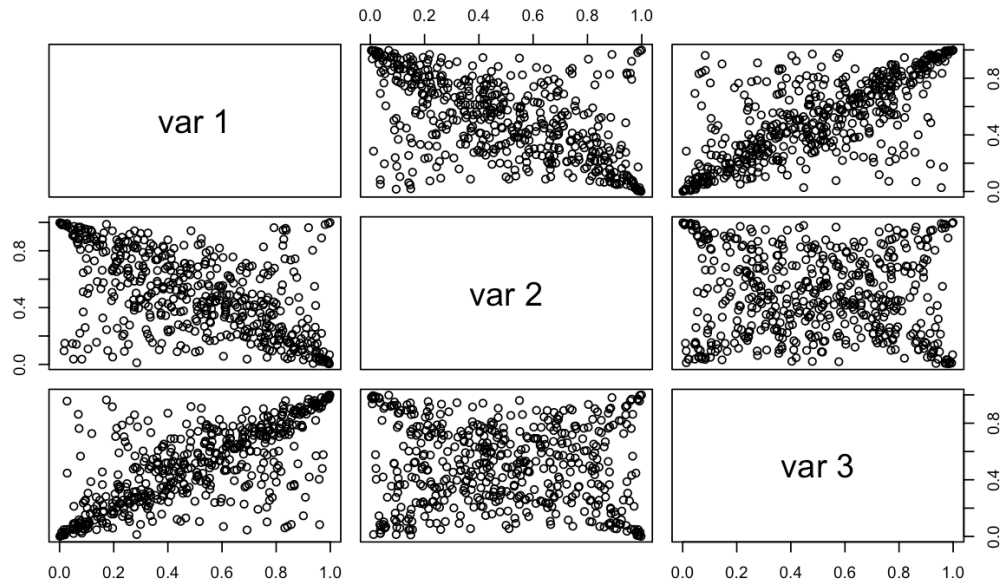
Answer
Part a

Figure 1: Scatterplot Matrix

From the plot we can see that components 2 and 3 (var 2 vs. var 3 plot) do not appear independent. Although the theoretical linear correlation parameter is $\rho_{23} = 0$, the scatterplot exhibits clear tail dependence (clustering of points in the corners, particularly the top-right and bottom-left, as well as the top-left and bottom-right). Independence would require the points to be uniformly scattered across the unit square. This non-independent structure is expected because the function is a t-Copula with a very low degrees of freedom (df $= 1$), which induces dependence in the tails even when $\rho = 0$.

Moreover, from the plot we can see that component 1 and 2 are negatively correlated, and component 1 and 3 are positive correlated.

Part b

Yes, strong signs of tail dependence are evident in all three pairs of components.

- var 1 and var 3 ($\rho_{13} = 0.75$): Strong clustering in the bottom-left $(0,0)$ and top-right $(1,1)$ corners (positive tail dependence).

5

- var 1 and var 2 ($\rho_{12} = -0.6$): Strong clustering in the top-left $(0, 1)$ and bottom-right $(1, 0)$ corners (negative tail dependence).

- var 2 and var 3 ($\rho_{23} = 0$): Although $\rho = 0$, there is visible clustering in the corners, indicating symmetric tail dependence without a linear trend, characteristic of a low df t-Copula.

Part c

- Effect 1 (Linear trend): The non-zero correlation parameters ($\rho_{12} = -0.6$ and $\rho_{13} = 0.75$) determine the primary linear association or "squeeze" of the cloud of points along the diagonal/anti-diagonal.

- Effect 2 (Tail dependency): The low degrees of freedom (df $= 1$) forces a higher density of points into the corners $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$, indicating that extreme values in one component are highly likely to be paired with extreme values in the other. We can see this in both negative (var 1 and var 2) and positive (var 2 and var 3) cases.

- Effect 3 (Non-uniformity): Compared to an independent copula (where points would be uniformly scattered), the dependence causes the points to be highly concentrated in specific regions dictated by ρ and df.

Part d

The Pearson correlation matrix is printed as below, whereas the correlation we entered was `param = c(-0.6,0.75,0)`; they are different.

```
              [,1]          [,2]          [,3]
[1,]   1.0000000 -0.54999514   0.70707296
[2,]  -0.5499951  1.00000000  -0.06538499
[3,]   0.7070730 -0.06538499   1.00000000
```

Besides random variation, there might be strong nonlinear effects which caused the problem; for t-corpula with small degree of freedom, it has strong tail dependence (as in this case). In other words, the parameters of corpula only measures linear related parameters, but here there is a strong nonlinear trend.

The confidence interval generated by the codes give are shown below:

```
Pearson's product-moment correlation

data:  rand_t_cop[, 1] and rand_t_cop[, 3]
t = 22.314, df = 498, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
0.6603249 0.7483624
sample estimates:
cor
0.707073
```

As we can see the confidence interval is $(0.6603249, 0.7483624)$, which does not include 0.75. Becasue of the nonlinear transform when we build copula, the relationships between variables would change, thus causing the change of correlation parameter.

4   Question 4

Suppose that there two bonds A and B that may default and let T_A and T_B denote their respective default times. Suppose that $T_A \sim$ Exponential $(\lambda_A = 0.10)$ and $T_B \sim$ Exponential $(\lambda_B = 0.07)$. Assume that the risk free rate is 0 and that $(F_A(T_A), F_B(T_B))$ is distributed according to Joe's copula:

$$C(u, v) = 1 - \left[(1-u)^2 + (1-v)^2 - (1-u)^2(1-v)^2\right]^{1/2}$$

(a) Consider an insurance with a payoff of $2 million if both bonds default before the end of the first year and $1 million if both survive the first year but they default before the end of the second year. Find the fair price of this insurance.

(b) Consider an insurance with a payoff of $2 million if both bonds default before the end of the first year and $1 million if both survive the first year but A or B defaults before the end of the second year. Find the fair price of this insurance.

Answer
<u>Part a</u>
We are given two default times:

$$T_A \sim \text{Exponential}(\lambda_A = 0.10), \qquad T_B \sim \text{Exponential}(\lambda_B = 0.07),$$

and Joe's copula

$$C(u, v) = 1 - \sqrt{(1-u)^2 + (1-v)^2 - (1-u)^2(1-v)^2}.$$

Let the survival probabilities be:

$$s_A(t) = e^{-0.1t}, \qquad s_B(t) = e^{-0.07t}.$$

Then the joint survival copula is:

$$\overline{C}(s_A, s_B) = s_A + s_B - \sqrt{s_A^2 + s_B^2 - (s_A s_B)^2}.$$

Also, the joint CDF is:

$$H(t_A, t_B) = \mathbb{P}(T_A \le t_A, T_B \le t_B) = 1 - s_A(t_A) - s_B(t_B) + \overline{C}(s_A(t_A), s_B(t_B)).$$

$$s_A(1) = e^{-0.10} = 0.904837, \quad s_B(1) = e^{-0.07} = 0.932394,$$
$$s_A(2) = e^{-0.20} = 0.818731, \quad s_B(2) = e^{-0.14} = 0.869358.$$
$$\overline{C}(s_A(1), s_B(1)) = 0.8491428483,$$
$$\overline{C}(s_A(2), s_B(2)) = 0.7291902477,$$
$$H(1, 1) = 0.0119116104,$$
$$H(2, 2) - H(1, 2) - H(2, 1) + H(1, 1) = 0.0088558816.$$

Payoff:
$$2M \text{ if both default by 1 year,} \quad 1M \text{ if both survive 1 year but default by 2 years.}$$

Thus the fair price is

$$P_a = 2{,}000{,}000\, H(1, 1) + 1{,}000{,}000\, [H(2, 2) - H(1, 2) - H(2, 1) + H(1, 1)]$$
$$= 2{,}000{,}000(0.0119116104) + 1{,}000{,}000(0.0088558816)$$
$$= \$32{,}679.10.$$

<u>Part b</u>
Payoff:

$$2M \text{ if both default by 1 year,} \quad 1M \text{ if both survive 1 year but at least one defaults by 2 years.}$$

$$P_b = 2{,}000{,}000\, H(1,1) + 1{,}000{,}000\left[\mathbb{P}(T_A > 1, T_B > 1) - \mathbb{P}(T_A > 2, T_B > 2)\right]$$
$$= 2{,}000{,}000\, H(1,1) + 1{,}000{,}000\left[\overline{C}(s_A(1), s_B(1)) - \overline{C}(s_A(2), s_B(2))\right]$$
$$= 2{,}000{,}000(0.0119116104) + 1{,}000{,}000(0.1199526006)$$
$$= \$143{,}775.82.$$