



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

Treball final de grau

DOBLE GRAU MATEMÀTIQUES - ADMINISTRACIÓ I DIRECCIÓ
D'EMPRESSES

*Facultat de Matemàtiques - Facultat d'Economia i Empresa
Universitat de Barcelona*

AVALUACIÓ DE LA DEPENDÈNCIA I RISC EXTREM ENTRE INDICADORS D'INSOLVÈNCIA: UNA APLICACIÓ DE LES CÒPULES EN L'ANÀLISI FINANCER

Autora: Maria Aiguadé Carulla

Directors: Dr. Jordi Martí Pidelaserra - Dra. Maria Carmen Florit Selma

Realitzat a: Departament d'Empresa - Departament de Matemàtiques i Informàtica

Barcelona, Gener 2026

Abstract

This work studies how the risk of corporate insolvency evolves before and during the crisis caused by COVID-19, with the aim of analyzing both the behavior of extreme risk and the relationship between different financial ratios that indicate insolvency. The analysis is carried out on a set of companies at risk of insolvency, combining non-parametric dependence measures, bivariate copula adjustment and extreme risk measures such as Value at Risk (VaR) and Expected Shortfall (ES).

Thus, the work combines tools from statistics and risk theory allowing us to study insolvency risk from a perspective complementary to that of the classic methods used in the literature.

The tools described are applied to a data sample configured by about 1300 companies and their respective annual accounts during the years 2017 - 2023. The analyzed period allows to compare the behavior of the indicators before and during the crisis caused by COVID-19. The data have been extracted from the SABI database (*Sistema de Análisis de Balances Ibéricos*).

Overall, the study shows the usefulness of combining dependency analysis with extreme risk measures to better understand the behavior of insolvency risk in crisis environments, as well as the importance of taking into account business variety when interpreting the results.

Resum

Aquest treball estudia com evoluciona el risc d'insolvència empresarial abans i durant la crisi provocada pel COVID-19, amb l'objectiu d'analitzar tant el comportament del risc extrem com la relació entre diferents ràtios financeres indicadores d'insolvència. L'anàlisi es realitza sobre un conjunt d'empreses en situació de risc d'insolvència, combinant mesures de dependència no paramètriques, ajust de còpules bivariants i mesures de risc extrem com el Value at Risk (VaR) i l'Expected Shortfall (ES).

Així doncs, el treball combina eines d'estadística i de teoria del risc que permeten estudiar el risc d'insolvència des d'una perspectiva complementària a la dels mètodes clàssics utilitzats en la literatura.

Les eines descrites s'apliquen a una mostra de dades configurada per unes 1300 empreses i els seus comptes anuals respectius durant els anys 2017 - 2023. El període analitzat permet comparar el comportament dels indicadors abans i durant la crisi provocada pel COVID-19. Les dades han sigut extretes de la base de dades SABI (*Sistema de Análisis de Balances Ibéricos*).

En conjunt, el treball mostra la utilitat de combinar l'anàlisi de dependència amb mesures de risc extrem per entendre millor el comportament del risc d'insolvència en entorns de crisi, així com la importància de tenir en compte la varietat empresarial a l'hora d'interpretar els resultats.

Agraïments

Primer de tot, vull agrair el constant suport rebut per part dels meus tutors durant la redacció d'aquest Treball de Fi de Grau. La seva guia i ajuda han estat crucials per completar aquest document i per tant, m'agradaria compartir amb ells el resultat.

Finalment, vull agrair el suport incondicional per part de la meva germana i els meus pares, així com totes les oportunitats que m'han brindat, que han estat determinants per arribar fins a la realització d'aquest document.

Índex

1	Introducció	1
1.1	Context i motivació	1
1.2	Objectius generals i específics	1
1.3	Abast i limitacions	2
1.4	Estructura del document	2
2	Marc Teòric	3
2.1	Introducció a l'anàlisi d'estats financers	3
2.1.1	Definició i tipologies d'insolvència	3
2.1.2	Descripció formal de les ràtios	4
2.2	Concepte de risc: VaR i Expected Shortfall	8
2.2.1	Definició i propietats	8
2.2.2	Limitacions	9
2.3	Teoria de dependència i còpules	10
2.3.1	Antecedents històrics	10
2.3.2	Definició de còpula i Teorema de Sklar	11
2.3.3	Famílies de còpules	15
2.4	Mesures de dependència	18
2.4.1	Correlació lineal versus Correlació de Rang	18
2.4.2	Tau de Kendall	20
2.4.3	Rho de Spearman	20
2.4.4	Propietats de la correlació de rang	21
2.4.5	Dependència de cues: coeficient superior i inferior	21
3	Dades i processament	23
3.1	Fonts de dades i criteris de selecció d'empreses concursades	23
3.2	Neteja, tractament de valors extrems i normalització	24
3.3	Construcció de “bucket” CNAE X Comunitat Autònoma	25
3.4	Càlcul de les ràtios financeres	26
3.5	Segmentació temporal (pre, COVID, post)	27
4	Metodologia Empírica	28

4.1	Modelització de la distribució marginal de cada ràtio	28
4.1.1	Transformació a escala uniforme (pseudo-observacions)	28
4.2	Estimació de paràmetres de còpules per MLE	29
4.3	Selecció de models: AIC, BIC i tests de bondat d'ajust	30
4.4	Ajust paramètric (Gaussiana, t-Student i Arximedianes)	31
4.4.1	Algoritmes de simulació	31
4.4.2	Comparació de models de còpula	34
4.5	Estimació de mesures de dependència	34
4.5.1	Kendall's τ i Spearman's ρ	34
4.5.2	Dependència de cua	35
4.6	Estimació de mesures de risc: VaR i ES	35
4.6.1	Transformació prèvia de ràtios per ús en VaR/ES	35
4.6.2	Càlcul del VaR i ES empírics	37
4.7	Implementació computacional	37
4.7.1	Estructura del codi	37
5	Resultats	40
5.1	Dependència empírica entre ratis financers	40
5.2	Ajust de còpules: selecció del model de dependència	41
5.3	Validació del models: dependència observada vs modelada	43
5.4	Mesures de risc: VaR i Expected Shortfall	44
5.5	Dependència de cua i risc extrem conjunt	45
5.6	Identificació de ràtios tractors	46
6	Conclusions i Línies futures	47
	Bibliografia	49
	Apèndix	51

Capítol 1

Introducció

1.1 Context i motivació

La crisi derivada del COVID-19 va suposar una alteració en el comportament financer de les empreses a escala mundial. En aquest context de crisi financera, van ser moltes les empreses que van resultar afectades i les quals van entrar en períodes d'insolvència. A Espanya, d'acord amb l'INE (Instituto Nacional de Estadística) [1], es va registrar un increment aproximat del 9% en els procediments concursals durant la segona meitat del 2020 tot i els esforços del govern, que va decretar una moratòria concursal el març del 2020 amb l'objectiu de reduir la mortalitat empresarial. Aquestes mesures, tot i reduir temporalment la pressió sobre les empreses, van generar una acumulació d'empreses vulnerables que van acabar en concurs entre 2021 – 2023.

Aquest context evidencia la necessitat de comprendre millor com el deteriorament financer afecta l'estructura empresarial i, en particular, com es reflecteix en els principals indicadors comptables i financers, com podrien ser les ràtios financeres: **ROA, ROE, Marge Brut, Palanquejament...**, que són eines habituals per avaluar la situació econòmica i financera d'una empresa. Tanmateix, més enllà de l'anàlisi individual de cada indicador, resulta rellevant estudiar com es relacionen entre ells i com aquestes relacions es poden veure alterades en períodes de crisi.

És en aquest punt on sorgeix la motivació d'aquest treball. En contextos d'estrès financer, les relacions entre ràtios poden deixar de ser estables o lineals, i les mesures de correlació tradicionals poden resultar insuficients, donat que no són capaces de capturar dependències no lineals ni comportaments asimètrics en les cues de les distribucions. Per aquest motiu, l'ús de **models de còpules** resulta una eina adequada per analitzar l'estructura de dependència entre ràtios financeres, ja que permet estudiar tant la dependència mitjana com la seva evolució en situacions extremes. Així doncs, aquest treball es proposa analitzar com la crisi del COVID-19 ha afectat tant al risc extrem com a les relacions de dependència entre diferents ràtios financeres.

1.2 Objectius generals i específics

Ens trobem amb dos objectius principals. D'una banda, estudiar l'evolució de les ràtios financeres associades a la solvència, la liquiditat, la rendibilitat i l'estructura financera durant el període 2017-2023, posant atenció al risc extrem. D'altra banda, analitzar com varia la dependència entre aquestes ràtios en contextos normals i de crisi, mitjançant l'ús de models de còpules bivariants, diferenciant entre els períodes pre-COVID, COVID i post-COVID.

Al llarg del treball es plantegen diversos objectius específics. En primer lloc s'avaluarà la intensitat de la dependència entre ràtios mitjançant mesures no paramètriques com la τ de Kendall, ρ de Spearman i els **coeficients de dependència de cues**. En segon lloc, és essencial determinar quina d'entre les famílies de còpules estudiades: **Gaussiana, t-Student o Arquimedianes**,

és capaç d'ajustar-se millor a les dades reals, ja que cada família permet capturar diferents tipus de dependència. En tercer lloc, es vol quantificar l'impacte de la crisi mitjançant mesures de risc individuals, com són el **Value at risk** i **Expected Shortfall**, i s'analitza com aquests indicadors evolucionen al llarg dels diferents períodes considerats. Finalment, es vol investigar si hi ha algunes ràtios que tendeixen a propagar el risc cap a altres variables, mitjançant la probabilitat condicionada, sense buscar establir relacions causals estrictes.

1.3 Abast i limitacions

Les dades seleccionades per portar a terme l'estudi es centren exclusivament en empreses que van declarar concurs entre els anys 2020 - 2023, període en que va succeir la crisi del COVID-19. L'elecció es basa en el fet que l'objectiu principal és analitzar com es deterioren les mesures d'insolvència en un període d'estrès extrem i quines relacions entre les mesures es veuen alterades. Així doncs, estudiar empreses insolvents permet observar directament l'evolució que aquestes empreses pateixen al canviar d'estat.

La principal limitació ha estat la falta o irregularitat d'informació comptable completa durant el període COVID i post-COVID. Aquest fet es dona per la naturalesa de les empreses estudiades. Moltes empreses insolvents, degut a que la situació de concurs no els permet seguir amb l'activitat habitual, presenten retards, inconsistències o directament deixen de presentar comptes. A més, les empreses que s'extingeixen durant el procediment concursal deixen de generar informació. Aquests fets, redueixen la mida de la mostra i poden produir biaixos.

Aquestes restriccions afecten a la representativitat global, especialment en algunes Comunitats Autònomes on el nombre d'observacions és reduït. Tot i així, l'estudi no perd rellevància perquè permet caracteritzar patrons de deteriorament financer en períodes d'estrès.

1.4 Estructura del document

El document es troba dividit en dos seccions: Una part teòrica i una part pràctica.

La part teòrica presenta definicions de conceptes claus per entendre en profunditat els elements econòmics i les eines matemàtiques que s'utilitzaran per realitzar el treball. Primer, s'introdueixen les principals ràtios financeres relacionades amb la solvència i la rendibilitat, amb les seves definicions i interpretació econòmica. A continuació, es descriuen les mesures de risc **Value at Risk** i **Expected Shortfall**, així com la seva rellevància en entorns de fallida. Finalment es presenta la teoria de còpules, on s'introdueixen les famílies de còpules, les seves principals propietats i les mesures de dependència que s'aplicaran per treballar la relació entre variables financeres.

La part pràctica introdueix el tractament i construcció de la base de dades, la definició dels "buckets" per Comunitat Autònoma i la metodologia aplicada per estimar marginals, dependències i mesures de risc. Un cop queda definida la base de dades i els elements que la configuren, s'apliquen els conceptes exposats durant la part teòrica mitjançant un programa que es troba inclòs en l'Apèndix, per tal de que qualsevol interessat pugui complementar aquesta secció.

El document finalitza amb les conclusions obtingudes a partir dels resultats treballats en les dos seccions prèvies així com amb una reflexió de possibles línies futures per continuar el treball.

Capítol 2

Marc Teòric

2.1 Introducció a l'anàlisi d'estats financers

Podríem descriure la comptabilitat com una disciplina que registra, classifica i analitza les operacions efectuades per una entitat. La informació que recull aquesta disciplina permet confeccionar els anomenats estats financers, documents normalitzats que reflecteixen la situació econòmica i el rendiment de l'empresa i que permeten dur a terme un anàlisi financer rigorós.

En aquest treball, les eines utilitzades deriven principalment de la informació continguda en dos d'aquests estats: el **balanç de situació** i el **compte de pèrdues i guanys**.

De manera col·loquial, el balanç de situació pot entendre's com una "fotografia" de l'empresa en una data concreta. Està format per dues parts. En la primera hi trobem l'**actiu**, que recull els bens, drets i recursos controlats econòmicament per l'empresa i dels quals s'espera obtenir beneficis futurs. En la segona, hi figura el **passiu i patrimoni net**, que agrupa, d'una banda, les obligacions actuals derivades d'esdeveniments passats i, de l'altra, els recursos aportats pels propietaris i els resultats acumulats.

En quant al compte de **pèrdues i guanys**, tal i com el nom indica, recull tant els **ingressos** (conseqüència d'un increment en els recursos de l'empresa), com les **despeses** (generats degut a una disminució dels recursos de l'empresa). Així doncs, el propòsit del compte de pèrdues i guanys és determinar si una entitat ha estat exitosa durant l'any, és a dir, si l'empresa ha generat benefici o pèrdua.

En síntesi, la informació proporcionada pels estats financers és l'eina fonamental per avaluar la rendibilitat, la solvència i, en definitiva, la continuïtat d'una entitat. Aquest marc és el punt de partida per configurar les ràtios financeres que permeten detectar situacions de risc d'insolvència.

El lector que estigui interessat en profunditzar sobre els conceptes exposats en aquesta secció pot referir-se a [14] o [15].

2.1.1 Definició i tipologies d'insolvència

Quan es parla de solvència, és imprescindible introduir dos conceptes fonamentals. **La solvència** fa referència a la capacitat d'una entitat per atendre les seves obligacions de pagament en els terminis establerts, mentre que **la liquiditat** indica el volum de caixa i actius líquids disponibles. Sovint ambdós termes es fan servir de manera conjunta, però no són equivalents: una empresa pot ser solvent però no disposar de liquiditat immediata, i la manca de liquiditat pot conduir a tensions financeres fins i tot en entitats amb estructura patrimonial aparentment sòlida.

D'acord amb el que es descriu a [11] i [12], **la insolvència** es defineix com una situació jurídica en la que es troba aquella persona física o empresa quan no pot fer front al pagament de les seves obligacions. Aquesta pot ser **actual**, quan el deutor ja ha deixat d'atendre les seves obligacions,

o **imminent**, quan el deutor preveu que no podrà complir regular i puntualment amb les seves obligacions.

Dins del concepte general d'insolvència, es distingeixen habitualment dues tipologies principals:

- **Insolvència de flux de caixa o efectiu.** Coneguda també com insolvència de liquiditat, la qual es refereix a una situació temporal i específica en la que la persona o entitat no pot fer front a les seves obligacions financeres a la data pactada, però si compta amb els recursos suficients per complir-les en un termini de temps curt. En aquest cas, el balanç mostra un actiu total superior al passiu, però els actius no són prou líquids o tarden a convertir-se en efectiu. El problema, per tant, no és patrimonial sinó de gestió de circulat i de calendari de cobraments i pagaments.
- **Insolvència de balanç o patrimonial.** Es produeix quan el valor dels passius supera el de l'actiu, de manera que l'empresa no pot satisfer les seves obligacions ni mitjançant els ingressos futurs ni liquidant els seus actius. En aquest escenari, les mesures habituals passen per reestructuracions de deute, ampliacions de capital o obtenció de finançament addicional. Si aquestes actuacions no prosperen, la situació pot conduir a una liquidació definitiva.

Adicionalment, segons [13] una empresa es considera en **risc d'insolvència** quan la taxa de retorn del capital invertit és persistentment baixa, per tant, la insolvència es pot relacionar amb la **rendibilitat**, que reflecteix la capacitat de generar beneficis i, en conseqüència, de mantenir la continuïtat del negoci. Així doncs, una rendibilitat insuficient sostinguda en el temps deteriora el patrimoni net i pot desencadenar problemes tant de liquiditat com de capacitat.

En la secció següent s'introduiran diversos instruments d'anàlisi que permeten avaluar com les diferents partides del balanç s'associen amb les tipologies d'insolvència descrites.

2.1.2 Descripció formal de les ràtios

Preveure amb total certesa que una empresa arribarà a una situació d'insolvència és pràcticament impossible, ja que les entitats estan exposades contínuament a incerteses pròpies de l'entorn econòmic. Tanmateix, tal com assenyala la literatura [10], és possible identificar amb antelació patrons i tendències en les ràtios financeres que anticipen un risc d'insolvència més elevat en el futur.

Al llarg del temps s'han desenvolupat múltiples tècniques i models per mesurar el risc financer, cadascun amb objectius i limitacions específiques. Tot i això, totes comparteixen un element comú: l'anàlisi de la informació comptable per transformar-la en indicadors que permetin extreure conclusions sobre la situació econòmica d'una empresa. En aquest treball optem per utilitzar les ràtios financeres més habituals, ja que l'objectiu no és construir un model predictiu de fallida, sinó estudiar com evolucionen aquestes ràtios i com varia la dependència entre elles quan una empresa entra en un període de deteriorament financer que pot acabar en insolvència.

D'acord amb [10], l'anàlisi financer tradicional s'articula en tres dimensions principals:

1. **Anàlisi de liquiditat.** Mesura la capacitat de l'empresa per fer front als compromisos a curt termini, i per tant està vinculat al risc d'insolvència de flux de caixa.
2. **Anàlisi de la solvència.** Avalua si l'empresa pot satisfer les seves obligacions a llarg termini. Està relacionat amb l'estructura de capital i el pes relatiu del deute.
3. **Anàlisi de la rendibilitat.** Examina la capacitat de generar beneficis tant a partir dels actius com del capital aportat pels accionistes.

Seguint [24], en l'anàlisi de la liquiditat ens interessa especialment el concepte de **puntualitat**, que reflecteix la capacitat de l'empresa per complir amb els pagaments en el moment requerit. En canvi, en l'anàlisi de solvència s'estudia la **capacitat**, entesa com la possibilitat de satisfer les obligacions globals del balanç, més enllà del curt termini.

A partir d'aquests dos conceptes (capacitat i puntualitat) podem agrupar les empreses en quatre situacions:

- Empreses amb alta capacitat i alta puntualitat (situació teòricament òptima).
- Empreses amb alta capacitat però baixa puntualitat (p.e. sectors on els cobraments són lents).
- Empreses amb alta puntualitat però baixa capacitat (p.e. entitats financeres amb elevada rotació de passius i actius líquids).
- Empreses amb baixa capacitat i puntualitat (empreses insolvents).

L'anàlisi d'aquests aspectes es farà sobre un conjunt reduït de ràtios seleccionades amb l'objectiu de minimitzar redundàncies i garantir que cada indicador aporta informació addicional rellevant.

Donat que el concepte de solvència s'associa a una situació en un moment determinat, i no a una evolució en el temps, les dades comparades en les ràtios de capacitat i puntualitat fan referència al mateix exercici. Aquestes mateixes ràtios utilitzen els valors comptables extrets dels estats financers de l'empresa en un moment específic en el temps, al final de l'any. En quant a les ràtios de rendibilitat, aquests capturen una evolució al llarg del temps, per aquest motiu, les dades utilitzades en aquestes ràtios fan referència a valors comptables a final d'any però corresponents a exercicis diferents.

Ens referim a [14] o [15] per conèixer la naturalesa d'aquests comptes.

Ràtios de Capacitat. Aquestes ràtios informen sobre si l'empresa disposa de suficients recursos (propis o líquids) per fer front al conjunt de les seves obligacions. Són crucials per detectar insolvència patrimonial.

- $$\frac{\text{Actiu Corrent} - \text{Existències}}{\text{Actiu Total}}$$

Mesura el pes dels actius més líquids (tresoreria i comptes a cobrar) sobre el total d'actius. El fet de ser els actius més líquids significa que són els que es converteixen més fàcilment en efectiu.

L'ús d'aquesta ràtio es justifica per la idea que com més actius líquids tingui una empresa, més capaç serà de complir amb les seves obligacions financeres. Excloem les existències perquè varien molt segons sector i poden trigar a convertir-se en efectiu, i per tant podria ser distorsionador en alguns casos.

- $$\frac{\text{Actiu Corrent}}{\text{Passiu Corrent}}$$

Avalua la capacitat de l'empresa de fer front al conjunt del passiu a curt termini mitjançant els actius que es convertiran en efectiu durant l'any següent. No s'utilitza l'actiu total perquè, seguint el principi comptable d'empresa en funcionament, els actius no corrents no s'acostumen a liquidar.

Si l'actiu corrent és inferior al passiu corrent, significa que l'empresa no pot cobrir els seus pagaments immediats amb els recursos que obtindrà en un any. És un indicador clau de solvència a curt termini i un dels primers senyals d'alerta quan una empresa comença a experimentar dificultats.

- Patrimoni Net
Passiu

Descriu l'estructura de finançament de l'empresa, comparant recursos propis i deute. Una major proporció de patrimoni net indica més solidesa financera i menor dependència del finançament extern.

Les empreses molt endeutades pateixen més en contextos adversos perquè d'entre d'altres punts, tenen càrregues financeres més elevades, tenen menys flexibilitat per buscar finançament i són més vulnerables a increments en els tipus d'interès.

És una ràtio fonamental per detectar insolvència de balanç o patrimonial.

Ràtios de Puntualitat. Aquestes ràtios identifiquen problemes en el cicle de caixa, clau per detectar insolvència de flux de tresoreria.

- *Període mitjà de pagament* = $\frac{\text{Compte de Proveïdors}}{\text{Compres}}$

En calcular aquesta ràtio en termes de dies, el resultat mostra quant de temps (quants dies) triga l'empresa a pagar els seus creditors. Si el PMP és excessivament llarg, pot indicar tensions de liquiditat: l'empresa ajorna pagaments perquè no té prou efectiu, si és molt curt, pot indicar falta de negociació amb proveïdors o necessitat de pagar per avançat.

És un indicador essencial per veure si l'empresa manté la caixa prolongant pagaments, senyal habitual en empreses en risc de fallida.

- *Període mitjà de cobrament* = $\frac{\text{Compte de Clients}}{\text{Vendes}}$

En calcular aquesta ràtio en termes de dies, el resultat mostra quant de temps (quants dies) triguen els deutors de l'empresa a pagar.

Si el PMC és molt alt, l'empresa finança involuntàriament els seus clients, cosa que redueix la liquiditat.

La diferència entre ambdós indicadors de puntualitat, $PMP - PMC$ reflecteix si l'empresa rep efectiu abans d'haver de realitzar els seus pagaments (situació favorable) o si, al contrari, afronta sortides de caixa abans d'ingressar fons (situació de risc).

Si és positiu, significa que l'empresa posseeix prou efectiu per pagar als seus creditors. Si la diferència és negativa, l'empresa s'enfronta al pagament dels seus creditors sense haver rebut prèviament efectiu dels seus deutors, que origina un risc alt d'exhauriment de tresoreria.

Així doncs, un valor persistentment negatiu podria ser un índex d'insolvència operativa.

Ràtios de Rendibilitat. Un dels indicadors més ben considerats per tal de mesurar la rendibilitat és el **Return On Investment (ROI)**, que compara els ingressos generats, amb el capital invertit. Així doncs, aquesta mesura relaciona el resultat obtingut d'una inversió de capital específica.

De moment, comptem amb la següent definició $ROI = \frac{\text{Retorn}}{\text{Inversió}}$ per tant, cal especificar quina inversió i quin retorn es volen estudiar. De cara al nostre propòsit ens interessa estudiar els següents conceptes

1. Utilitzant **Actius Totals** com a variable d'Inversió:

Al comparar els ingressos generats a final d'any, és a dir, el resultat d'explotació, amb els actius totals que posseeix l'empresa a principis d'any, estem mesurant el rendiment generat pels actius, denominat com **Return On Assets (ROA)**. Així doncs, la fórmula anterior

prendria la següent forma:

$$ROI_{\text{Actius Totals}} = ROA = \frac{\text{Resultat d'Explotació}_{31/12/X+1}}{\text{Actiu Total}_{31/12/X}} \quad (2.1.1)$$

2. Utilitzant el **Patrimoni Total** com a variable d'Inversió:

Al comparar benefici net amb el patrimoni total que posseeix l'empresa a principis d'any, estem mesurant la rendibilitat per als accionistes, denominada **Return On Equity (ROE)**. Així doncs, la fórmula de l'inici prendria la següent forma:

$$ROI_{\text{Patrimoni Total}} = ROE = \frac{\text{Benefici Net}_{31/12/X+1}}{\text{Patrimoni Net}_{31/12/X}} \quad (2.1.2)$$

Aquestes dues ràtios, es poden desglossar de la següent manera

$$ROA = \text{Marge Brut} * \text{Rotació} = \frac{\text{Resultat d'Explotació}_{31/12/X+1}}{\text{Vendes}_{31/12/X+1}} * \frac{\text{Vendes}_{31/12/X+1}}{\text{Actiu Total}_{31/12/X}} \quad (2.1.3)$$

De manera informal, el Marge Brut mesura euros de resultat d'explotació generats per euro de vendes. D'altra banda, la rotació indica euros de resultat d'explotació generats per euro de vendes.

Un ROA baix prolongat indica actius improductius o inversions inadequades, el que resulta en que l'empresa no generi caixa suficient per sostenir-se. Per tant és un indicador de deteriorament de la capacitat operativa.

$$ROE = \text{Marge Net} * \text{Rotació} * \text{Palanquejament} =$$

$$= \frac{\text{Benefici Net}_{31/12/X+1}}{\text{Vendes}_{31/12/X}} \frac{\text{Vendes}_{31/12/X+1}}{\text{Actiu Total}_{31/12/X}} \frac{\text{Actiu Total}_{31/12/X}}{\text{Patrimoni Net}_{31/12/X}} \quad (2.1.4)$$

Informalment, el Marge Net mesura el percentatge de benefici net obtingut per cada euro de vendes. D'altra banda, el palanquejament mesura euros invertits per cada euro de patrimoni net. Així doncs, el palanquejament també és un indicador del nivell d'endeutament i del potencial efecte amplificador del deute:

- >1 . L'efecte palanquejament és positiu, és a dir, el deute augmenta la rendibilitat.
- $=1$. L'efecte palanquejament és neutral, és a dir, el deute no afecta a la rendibilitat.
- <1 . L'efecte palanquejament és negatiu, és a dir, el deute disminueix la rendibilitat.

Un ROE baix significa que l'empresa no compensa el risc assumit pels accionistes. Si aquest és persistentment negatiu deteriora el patrimoni net, fet que pot provocar insolvència patrimonial. És un indicador molt utilitzat per distingir empreses viables d'empreses que acabaran entrant en processos pre-concursals.

Així com en seccions prèvies, el lector interessat pot expandir en les idees dels conceptes descrits anteriorment referint-se a [14] o [15].

2.2 Concepte de risc: VaR i Expected Shortfall

Les mesures de risc que s'introdueixen a continuació, el **Value at Risk** (VaR) i l'**Expected Shortfall** (ES), s'utilitzen habitualment per quantificar el risc associat a variables aleatòries que representen pèrdues. Tot i que en aquest treball no s'analitzen directament pèrdues financeres, aquestes mesures continuen essent útils per caracteritzar els valors més desfavorables de les ràtios estudiades.

La seva presentació en aquest apartat segueix l'enfocament estàndard de [9], tot i que en el marc metodològic posterior s'explicarà com es transformen les ràtios perquè l'aplicació del VaR i l'ES sigui coherent i permeti extreure conclusions significatives sobre el deteriorament financer de les empreses durant el període COVID.

Així, encara que el context d'aplicació difereixi del tradicional, les propietats d'aquestes mesures continuen aportant informació rellevant sobre els extrems de la distribució de les ràtios i sobre la seva evolució en situacions adverses.

En situacions de fallida empresarial o d'estrès financer, les distribucions de les ràtios comptables i dels fluxos econòmics mostren comportaments extrems, asimètrics i molt allunyats de la normalitat. En aquest context, el Value at Risk (VaR) i l'Expected Shortfall (ES) esdevenen eines útils per analitzar i quantificar els riscos associats a aquests valors extrems.

2.2.1 Definició i propietats

Degut a la importància del quantils en contextos d'anàlisi de variables extremes, donarem la seva definició exacte.

Definició 2.2.1. (*Inversa generalitzada i funció quantil*)

- Donada una funció creixent $T : \mathbb{R} \rightarrow \mathbb{R}$, la inversa generalitzada de T es defineix com $T^{\leftarrow}(y) := \inf\{x \in \mathbb{R} : T(x) \geq y\}$, on utilitzem la convenció de que l'ínfim d'un conjunt buit és ∞ .
- Donada una funció de distribució F , anomenem a la inversa generalitzada $F^{\leftarrow}$ funció quantil de F . Per a $\alpha \in (0, 1)$ el α -quantil de F ve donat per

$$q_{\alpha}(F) := F^{\leftarrow}(\alpha) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}. \quad (2.2.1)$$

Si F és contínua i estrictament creixent, aleshores tenim que $q_{\alpha}(F) = F^{-1}(\alpha)$, on F^{-1} és la inversa ordinària de F .

Definició 2.2.2. (*Value at Risk*) Donat un cert nivell de confiança $\alpha \in (0, 1)$, el VaR d'una cartera amb pèrdua L al nivell de confiança α ve donat pel nombre més petit l tal que la probabilitat de que la pèrdua L superi l no sigui més gran que $1 - \alpha$. Formalment,

$$VaR_{\alpha} = VaR_{\alpha}(L) = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\} \quad (2.2.2)$$

En termes probabilístics, el VaR és simplement un quantil de la distribució de pèrdues.

Exemple 2.2.3. Per a $\alpha = 0.95$, si el valor del VaR fos 2 (en un context de pèrdues), indicaria que la probabilitat de que les pèrdues siguin majors a 2 és del 5%.

Definició 2.2.4. (*Expected Shortfall*) Per a una pèrdua L amb $E(|L|) < \infty$ i funció de distribució F_L , l'ES a un nivell de confiança $\alpha \in (0, 1)$ es defineix com

$$ES_{\alpha} = \frac{1}{1 - \alpha} \int_{\alpha}^1 q_u(F_L) du, \quad (2.2.3)$$

on $q_u(F_L) = F_L^{\leftarrow}(u)$ és la funció quantil de F_L .

Nota: La condició $E(|L|) < \infty$ assegura que la integral a (2.2.3) es troba ben definida.

Passem a enunciar un Lema que ens servirà més endavant per calcular l'ES de manera pràctica.

Lema 2.2.5. *Per una pèrdua integrable L amb una funció de distribució contínua F_L i per a qualsevol $\alpha \in (0, 1)$ tenim*

$$Es_\alpha = \frac{E(L; L \geq q_\alpha(L))}{1 - \alpha} = E(L|L \geq VaR_\alpha). \quad (2.2.4)$$

On hem utilitzat la notació $E(X; A) := E(XI_A)$ per a un genèric integrable d'una variable aleatòria X i un conjunt genèric $A \in \mathcal{F}$.

La prova d'aquest lema es pot trobar a [9].

Exemple 2.2.6. Per $\alpha = 0.95$ Si el valor de l'ES fos 4 (en un context de pèrdues), indicaria que en el 5% pitjor dels casos, la pèrdua mitjana és igual a 4.

Així doncs, les mesures de risc descrites anteriorment, resulten rellevants en aquest estudi degut a que

- Capturen els pitjors escenaris, no el comportament mitjà.
- Ajuden a detectar umbrals de risc (valors a partir dels quals una ràtio entra en zona crítica) i deteriorament progressiu (com les cues es fan més pesades en moments de crisi).
- L'ES és una mesura coherent i robusta en presència de cues pesades.

2.2.2 Limitacions

Tot i la seva popularitat en l'àmbit de la gestió del risc, tant el **Value at Risk** com l'**Expected Shortfall** presenten diverses limitacions que cal tenir en compte a l'hora d'utilitzar-los en aplicacions pràctiques o en contextos que s'allunyen del seu ús tradicional.

En primer lloc, el VaR és altament sensible al model utilitzat per estimar la distribució de pèrdues. Aquesta dependència del model implica que si la distribució escollida és o bé inadequada o bé els paràmetres s'estimen amb dades incompletes o bé les dades són poc representatives, el resultat del VaR pot no ser realista. Aquest problema, conegut com a **model risk**, és especialment rellevant en entorns amb estructures de dades canviants o en períodes d'alta incertesa.

En segon lloc, el VaR no té en compte la liquiditat dels actius ni el cost temporal de desfer posicions. En situacions de tensió financera, els preus poden reaccionar lentament o ser poc representatius, fet que provoca que el risc real sigui superior al que estima el VaR en condicions de mercat "normals". Aquest problema és especialment greu en episodis de volatilitat.

En tercer lloc, el VaR depèn de l'horitzó temporal escollit, és a dir, un VaR calculat en un dia no és necessàriament extrapolable a períodes més llargs, especialment quan els rendiments no són independents o quan existeix dependència de cues.

Finalment, una de les limitacions més conegudes del VaR és que no és una mesura subadditiua. Això significa que el VaR d'una cartera pot ser superior a la suma dels VaR individuals dels seus components, contràriament al que seria desitjable en una bona mesura de risc. La manca de subadditivitat implica que el VaR pot jutjar injustament la diversificació, tot i que la diversificació és un principi fonamental de la gestió del risc.

L'Expected Shortfall, tot i complir amb el principi de subadditivitat i proporcionar una informació més completa sobre les pèrdues extremes, comparteix algunes de les limitacions anteriors:

continua depenent del model, del període temporal escollit i pot ser difícil d'estimar de manera robusta en mostres petites o en distribucions molt asimètriques.

En conjunt, aquestes deficiències fan necessari utilitzar el VaR i l'ES amb precaució i complementar-los amb altres eines de gestió del risc, especialment en contextos d'incertesa elevada o quan l'estructura de dependència entre variables pot canviar significativament, com és el cas en períodes d'estrès financer.

2.3 Teoria de dependència i còpules

Nota: Al llarg d'aquesta secció representarem els vectors multivariants de la següent manera: $Y=(Y_1, \dots, Y_d)$

2.3.1 Antecedents històrics

L'ús de còpules en el món de les finances es va popularitzar a partir dels anys 2000. La seva motivació principal prové de l'evidència empírica que mostra que la distribució multi-normal és insuficient per modelar adequadament els rendiments dels actius financers. Això es deu, sobretot, a dos factors:

- Les distribucions marginals empíriques sovint presenten asimetries i cues pesades, allunyant-se clarament de la normalitat.
- La normal multivariant no capta els co-moviments extrems, és a dir, la tendència dels actius a moure's conjuntament en situacions d'estrès financer.

En altres paraules, l'estructura de dependència real observada als mercats financers és diferent de l'assumida pel model gaussià, especialment durant crisis. Les còpules, en canvi, permeten separar el modelatge de les distribucions marginals de la dependència entre variables, i això ofereix una flexibilitat molt superior. Per aquest motiu, s'han convertit en una eina útil i relativament senzilla d'implementar per simular escenaris realistes i construir models de risc més adequats.

Segons el que s'exposa a [22], les aplicacions més exitoses de les còpules s'han desenvolupat en àrees com el risc de mercat, risc de crèdit i risc operacional. En risc de mercat, s'utilitzen per modelar la dependència entre diferents factors de risc, com ara rendiments d'accions, tipus d'interès o divises, aconseguint una representació més precisa que la que oferiria la simple correlació lineal.

En l'àmbit del crèdit, les còpules han estat especialment útils perquè aquests actius presenten distribucions no normals, asimetries importants i dependència de cua. Això ha permès modelar de manera més realista la dependència entre impagaments.

Cal destacar també les aplicacions en assignació d'actius i gestió de carteres de crèdit, on s'ha observat que la composició òptima de la cartera pot dependre fortament del tipus de còpula utilitzada (vegeu [22] per referències addicionals).

Finalment, les funcions de còpula també han tingut un paper rellevant en la gestió de riscos sistèmics, ajudant a estimar com es transmet el risc entre institucions financeres i a identificar entitats sistèmicament importants (vegeu [22]).

En conjunt, les còpules han aportat millores clares en múltiples camps, especialment allà on la dependència no lineal i els co-moviments extrems són determinants. Tot i això, en l'àmbit específic del risc d'insolvència empresarial, el seu ús ha estat relativament limitat. Això pot

deure's a la dificultat d'obtenir dades prou homogènies, al predomini d'altres mètodes com pot ser l'Altman Z-score o a la preferència d'indicadors comptables més simples.

Per aquest motiu, aquest treball pretén explorar si les còpules poden ser una eina adequada per identificar patrons de dependència entre les ràtios financeres que la literatura assenyala com a indicadors d'insolvència, i així avaluar-ne la utilitat en aquest context concret.

2.3.2 Definició de còpula i Teorema de Sklar

Per tal d'introduir el problema de modelar dependència i el tipus de còpules que s'utilitzaran per atacar el problema, cal presentar una serie d'elements bàsics i conceptes sobre **còpules i dependència**.

D'acord amb [2], les còpules ens permeten tenir una comprensió més profunda de la dependència. Normalment, disposem de dades i informació relativa a les distribucions marginals de certes variables. La idea darrere de les còpules, es "ajuntar" les distribucions marginals i modelar la distribució conjunta per tal d'aïllar la descripció de la estructura de dependència.

Definició 2.3.1. Una còpula, denotada per $C(\mathbf{u})$, és una funció de distribució multivariant per a variables arbitràries amb marginals uniformes $(0,1)$, és a dir,

$$C(\mathbf{u}) = P(U_1 \leq u_1, \dots, U_d \leq u_d) \quad (2.3.1)$$

Utilitzem la notació $C(\mathbf{u}) = C(u_1, \dots, u_d)$ per a les funcions de distribució multivariants que són còpules. Per tant, C és una funció de la forma $C : [0, 1]^d \rightarrow [0, 1]$. Aquesta funció ha de seguir les següents tres propietats:

1. $C(u_1, \dots, u_d) = 0$ si $u_i = 0$ per a qualsevol i .
2. $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i \ \forall i \in \{1, \dots, d\}, u_i \in [0, 1]$.
3. $\forall (a_1, \dots, a_d), (b_1, \dots, b_d) \in [0, 1]^d$ amb $a_i \leq b_i$ tenim

$$\sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1+\dots+i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0,$$

on $u_{j1} = a_j$ i $u_{j2} = b_j \ \forall j \in \{1, \dots, d\}$.

Cal ressaltar que la segona propietat fa referència al requeriment de que les distribucions marginals siguin uniformes. Per altra banda, la desigualtat de la tercera propietat assegura que si el vector arbitrari $(U_1, \dots, U_d)'$ té funció de distribució C , aleshores $P(a_1 \leq U_1 \leq b_1, \dots, a_d \leq U_d \leq b_d) \geq 0$.

Aquestes tres propietats caracteritzen les còpules, és a dir, si una funció les compleix, aleshores és una còpula.

Per tal de treballar amb còpules, cal estar familiaritzats amb les següents proposicions elementals relacionades amb la probabilitat.

Proposició 2.3.2. (Transformació quantil) Si F és una funció de distribució i $F^{\leftarrow}$ la seva inversa generalitzada, si $U \sim \text{Uniforme}(0,1)$, aleshores

$$P(F^{\leftarrow}(U) \leq x) = F(x) \quad (2.3.2)$$

Prova: $P(F^{\leftarrow}(U) \leq x) = P(U \leq F(x)) = F(x)$.

Nota: Definim inversa generalitzada tal i com s'exposa en la Definició 2.2.1.

Proposició 2.3.3. (Transformació de probabilitat) Si X té una funció de distribució contínua F , aleshores

$$F(X) \sim \text{Uniforme}(0, 1) \quad (2.3.3)$$

Prova: $P(F(X) \leq u) = P(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u$.

Les darreres dues proposicions són clau en la simulació, donat que si podem generar una variable uniforme U i calcular la inversa d'una funció de distribució F , aleshores podem crear una mostra a partir d'aquesta funció de distribució. Si considerem conjuntament les dues proposicions, podem **transformar riscos amb una funció de distribució contínua** particular per tenir qualsevol altra distribució contínua.

Exemple 2.3.4. Sigui X una variable que segueix la distribució d'una normal estàndard, aleshores $\phi(X)$ és uniforme segons la *Proposició 2.3.3*, i com que la funció quantil d'una funció de distribució exponencial estàndard F és $F^{-1}(u) = -\ln(1 - u)$, la variable transformada $Y = -\ln(1 - \phi(X))$ segueix una distribució exponencial estàndard segons la *Proposició 2.3.2*. Passem a comprovar-ho

$$P(Y \leq y) = P(-\ln(1 - \phi(X)) \leq y) = P(1 - \phi(X) \geq e^{-y}) = P(\phi(X) \leq 1 - e^{-1}) = 1 - e^{-1}.$$

Teorema 2.3.5. (Teorema de Sklar) Sigui F una funció de distribució conjunta amb marginals $F_1, \dots, F_d$. Aleshores, existeix una còpula $C : [0, 1]^d \rightarrow [0, 1]$ tal que, per a tot $x_1, \dots, x_d$ en $\mathbb{R} = (-\infty, \infty)$:

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (2.3.4)$$

Si les marginals són contínues, aleshores C és única. Altrament, C està únicament determinada per $F_1(\mathbb{R}) \times F_2(\mathbb{R}) \times \dots \times F_d(\mathbb{R})$, on $F_i(\mathbb{R})$ denota l'abast de F_i . A l'inrevés, si C és una còpula i $F_1, \dots, F_d$ són funcions de distribució univariants, aleshores la funció F definida en (2.3.4) és una funció de distribució conjunta amb marginals $F_1, \dots, F_d$.

Prova: Passem a demostrar l'existència i unicitat d'una còpula en el cas que $F_1, \dots, F_d$ siguin contínues i l'afirmació inversa en la seva forma general.

Sigui $\mathbf{X}$ un vector aleatori amb funció de distribució F i marginals contínues $F_1, \dots, F_d$. Per a $i = 1, \dots, d$, establim $U_i = F_i(X_i)$. D'acord amb la *Proposició 2.3.3*, $U_i \sim U(0, 1)$, i $F_i^{\leftarrow}(U_i) = X_i$, gairebé amb seguretat (donat que si X és una variable aleatòria amb funció de distribució F , aleshores $P(F^{\leftarrow} \circ F(X) = X) = 1$). Sigui C la funció de distribució de $(U_1, \dots, U_d)$, que és una còpula per la *Definició 2.3.1*. Per a qualsevol $x_1, \dots, x_d$ a $\mathbb{R} = [-\infty, \infty]$, inferim, utilitzant que si F és contínua per la dreta, aleshores $F \circ F^{\leftarrow}(x) \geq x$, que

$$\begin{aligned} F(x_1, \dots, x_d) &= P(X_1 \leq x_1, \dots, X_d \leq x_d) = P(F_1^{\leftarrow}(U_1) \leq x_1, \dots, F_d^{\leftarrow}(U_d) \leq x_d) = \\ &= P(U_1 \leq F_1(x_1), \dots, U_d \leq F_d(x_d)) = C(F_1(x_1), \dots, F_d(x_d)), \end{aligned}$$

així doncs, obtenim la identitat (2.3.4).

Ara, suposem per reducció a l'absurd, que existeixen dues còpules C i C' que satisfan (2.3.4), és a dir, $F(\mathbf{X}) = C(F_1(x_1), \dots, F_d(x_d))$ i $F(\mathbf{X}) = C'(F_1(x_1), \dots, F_d(x_d))$. Si avaluem la identitat (2.3.4) amb els arguments $x_i = F_i^{\leftarrow}(u_i)$, $0 \leq u_i \leq 1$, $i = 1, \dots, d$, i utilitzem que si F és contínua aleshores $F \circ F^{\leftarrow}(x) = x$, obtenim

$$F(x_1, \dots, x_d) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)) = C(F_1(F_1^{\leftarrow}(u_1)), \dots, F_d(F_d^{\leftarrow}(u_d))) = C(u_1, \dots, u_d),$$

que dona una representació explícita de C en termes de F i les seves marginals. Fent el mateix amb C' , obtenim $F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)) = C'(u_1, \dots, u_d)$. Així doncs, $C \equiv C'$ i per tant tenim unicitat.

Per a l'afirmació inversa, suposem que C és una còpula i que $F_1, \dots, F_d$ són funcions de distribució univariants arbitràries. Construïm un vector aleatori amb funció de distribució (2.3.4) prenent $\mathbf{U}$ com a qualsevol vector aleatori amb funció de distribució C i establint $\mathbf{X} = (F_1^{\leftarrow}(U_1), \dots, F_d^{\leftarrow}(U_d))$. Aleshores podem seguir exactament la mateixa seqüència d'equacions utilitzades prèviament per veure que $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$ i establir que la funció de distribució de $\mathbf{X}$ satisfà (2.3.4).

Aquest darrer Teorema és fonamental en l'estudi de les funcions de distribució multivariants donat que dona els següents dos resultats:

1. Totes les funcions de distribució multivariants contenen còpules.
2. Les còpules es poden utilitzar conjuntament amb funcions de distribució univariants per construir funcions de distribució multivariants.

Altrament, les fórmules (2.3.4) i $C(u_1, \dots, u_d) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d))$ són fonamentals a l'hora de tractar còpules. La primera mostra com es formen les distribucions conjuntes F ajuntant distribucions marginals amb còpules C i com les còpules expressen la dependència en una escala de quantils, donat que el valor $C(u_1, \dots, u_d)$ fa referència a la probabilitat conjunta que $\forall i, X_i$ es trobi per sota del seu quantil u_i . La segona, mostra com s'extreuen les còpules de funcions de distribució multivariants amb marginals contínues.

Definició 2.3.6. (Còpula de F) Si un vector arbitrari X té funció de distribució F amb distribucions marginals contínues $F_1, \dots, F_d$, aleshores la còpula de F (o X) és la funció de distribució C de $(F_1(X_1), \dots, F_d(X_d))$.

Una propietat útil de la còpula d'una distribució és la seva **Invariància** sota transformacions estrictament creixents de les marginals. D'acord amb **Teorema de Sklar** i aquesta propietat **d'invariància**, podem interpretar la còpula d'una distribució com una manera molt natural de representar l'estructura de dependència de la distribució (clarament en el cas de marginals contínues).

Proposició 2.3.7. Sigui $(X_1, \dots, X_d)$ un vector aleatori amb marginals contínues i còpula C i sigui $T_1, \dots, T_d$ funcions estrictament creixents. Aleshores, $(T_1(X_1), \dots, T_d(X_d))$ també té per còpula C .

El lector interessat pot trobar la prova d'aquesta proposició a [?].

Per últim, tanquem aquesta secció ressaltant la importància dels *Límits de Fréchet* per còpules, els quals resulten tenir interpretacions de dependència importants que es discuteixen més endavant.

Teorema 2.3.8. Per a cada còpula $C(u_1, \dots, u_d)$, tenim les següents acotacions

$$\max\left(\sum_{i=1}^d u_i + 1 - d, 0\right) \leq C(\mathbf{u}) \leq \min(u_1, \dots, u_d) \quad (2.3.5)$$

Prova cota superior:

$$C(\mathbf{u}) = P(U_1 \leq u_1, \dots, U_d \leq u_d) = P(\cap_{i=1}^d (U_i \leq u_i)) \leq P(U_i \leq u_i) = u_i,$$

per a qualsevol $i = 1, \dots, d$. Així,

$$C(\mathbf{u}) \leq \min\{u_1, \dots, u_d\}$$

Prova cota inferior:

$$C(\mathbf{u}) = P(\cap_{i=1}^d (U_i \leq u_i)) = 1 - P((\cap_{i=1}^d (U_i \leq u_i))^c) = 1 - P(\cup_{i=1}^d (U_i > u_i)) \geq$$

$$1 - \sum_{i=1}^d P(U_i > u_i) = 1 - \sum_{i=1}^d (1 - u_i) = \sum_{i=1}^d u_i + 1 - d.$$

Així,

$$C(\mathbf{u}) \geq \max(\sum_{i=1}^d u_i + 1 - d, 0)$$

Nota: La notació $W(u_1, \dots, u_d)$ es refereix a la cota inferior, mentre que $M(u_1, \dots, u_d)$ a la cota superior.

Passem a donar dos exemples que il·lustren el procés de construcció d'una còpula a partir del *Teorema de Sklar* en les dues direccions.

Exemple 2.3.9. Càlcul d'una còpula a partir d'una distribució conjunta.

Considerem el vector aleatori (X, Y) amb densitat conjunta:

$$f_{X,Y}(x, y) = \begin{cases} 2, & 0 \leq x \leq y \leq 1, \\ 0, & \text{altrament.} \end{cases}$$

Passem a calcular la distribució conjunta. Recordem la definició de distribució conjunta,

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) dv du.$$

Com que la densitat només és positiva quan $0 \leq u \leq v \leq 1$, només calculem la integral sobre aquesta regió.

Cas 1: Si $0 \leq x \leq y \leq 1$, l'esdeveniment $\{X \leq x, Y \leq y\}$ correspon a la regió $0 \leq u \leq x$, $u \leq v \leq y$. Per tant,

$$F_{X,Y}(x, y) = \int_0^x \int_u^y 2 dv du = 2yx - x^2.$$

Cas 2: Si $0 \leq y < x \leq 1$, la condició $X \leq x$ no restringeix res, donat que dins del suport, la condició $X \leq Y \leq y \leq x$ sempre es dona. Per tan, aquest esdeveniment correspon a la regió $0 \leq u \leq v \leq y$ i

$$F_{X,Y}(x, y) = \int_0^y \int_0^v 2 du dv = y^2.$$

Així doncs, obtenim que la funció de distribució conjunta és

$$F(x, y) = \begin{cases} 0, & x < 0 \text{ o } y < 0, \\ 2xy - x^2, & 0 \leq x \leq y \leq 1, \\ y^2, & 0 \leq y < x \leq 1, \\ 1, & x \geq 1, y \geq 1. \end{cases}$$

Passem a calcular les marginals

$$F_X(x) = \lim_{y \rightarrow \infty} F(x, y) = 2x - x^2, 0 \leq x \leq 1,$$

$$F_Y(y) = \lim_{x \rightarrow \infty} F(x, y) = y^2, 0 \leq y \leq 1.$$

Cal remarcar que les funcions són contínues, per tant la còpula serà contínua i única. Calculem les inverses: $F_Y^{-1}(v) = \sqrt{v}$ i $F_X^{-1}(u) = 1 - \sqrt{1-u}$. Finalment, apliquem el *Teorema de Sklar* i obtenim la següent còpula

$$C(u, v) = F(F_X^{-1}(u), F_Y^{-1}(v)) = F(1 - \sqrt{1-u}, \sqrt{v}) = \begin{cases} 2(1 - \sqrt{1-u})\sqrt{v} - (1 - \sqrt{1-u})^2, & \text{si } 1 - \sqrt{1-u} \leq \sqrt{v}, \\ v, & \text{si } \sqrt{v} < 1 - \sqrt{1-u} \end{cases}$$

Exemple 2.3.10. Construcció d'una distribució conjunta a partir d'una còpula i distribucions marginals.

Definim les següents marginals polinòmiques

$$\begin{aligned} X &\sim U(0, 1) : F_X(x) = x, 0 \leq x \leq 1, \\ Y &\sim U(0, 1) : F_Y(y) = y^2, 0 \leq y \leq 1. \end{aligned}$$

Seleccionem una de les còpules que presentarem més endavant anomenada **Còpula de Clayton**, que té la forma següent

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, u, v \in (0, 1],$$

triem per exemple $\theta = 2$ (estem buscant que la còpula no sigui independent, per tant, podem triar qualsevol valor de $\theta \neq 0$). Ara, pel *Teorema de Sklar*, tenim

$$F_{X,Y}(x, y) = C_2(x, y^2) = (x^{-2} + y^4 - 1)^{-1/2}, 0 < x \leq 1, 0 < y \leq 1.$$

2.3.3 Famílies de còpules

Durant aquesta secció i d'acord amb el que s'exposa a [2] i [3], es donarà un nombre d'exemples de les còpules que s'utilitzaran en el marc pràctic. En primer lloc però, es donaran alguns conceptes sobre la dependència per a una millor comprensió del que s'exposa al llarg de la secció.

Dependència Perfecta

Definició 2.3.11. (Comonotonicitat) Les variables aleatòries $X_1, \dots, X_d$ s'anomenen comonotòniques si admeten com a còpula l'acotació de Fréchet superior, és dir, $M(u_1, \dots, u_d) = \min(u_1, \dots, u_d)$.

Proposició 2.3.12. X_1 i X_2 són comonotòniques si i només si

$$(X_1, \dots, X_d) \stackrel{d}{=} (v_1(Z), \dots, v_d(Z)) \quad (2.3.6)$$

per a alguna variable aleatòria Z amb $v_1, \dots, v_d$ funcions creixents.

El lector interessat pot profunditzar en aquesta i altres propietats d'aquest tipus de còpules a [6].

Definició 2.3.13. (Contramonotonicitat) Les variables aleatòries X_1 i X_2 són contramonotòniques si tenen com a còpula l'acotació de Fréchet inferior. És a dir, $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$

Proposició 2.3.14. X_1 i X_2 són contramonotòniques si i només si

$$(X_1, X_2) \stackrel{d}{=} (v_1(Z), v_2(Z)) \quad (2.3.7)$$

per a alguna variable aleatòria Z amb v_1 creixent i v_2 decreixent, o vice versa.

Nota: El concepte de contramonotonia no es generalitza a dimensions superiors. El límit inferior de Fréchet $W(u_1, \dots, u_d)$ no és en si mateix una còpula per a $d > 2$, ja que no és una funció de distribució pròpia i no satisfà la tercera propietats de les còpules. Es pot profunditzar i validar aquesta resultat a partir de [7].

INDEPENDENT

L'esmenada **Còpula Independent**, forma part de la categoria **Còpules Fundamentals**, que inclou aquelles còpules que representen una estructura de dependència especial. Dit això, la **Còpula Independent** ve donada per

$$\prod(u_1, \dots, u_d) = \prod_{i=1}^d u_i \quad (2.3.8)$$

Queda clar a partir del **Teorema de Sklar** i la fórmula (2.3.4) que les variables aleatòries amb funcions de distribució contínues són independents si i només si la seva estructura de dependència és la descrita per la fórmula (2.3.8).

GAUSSIANA

L'anomenada **Còpula Gaussiana** forma part de la categoria **Còpules Implícites**. Aquesta tipologia de còpules s'extreuen de distribucions multivariants molt conegudes amb funcions de distribució marginals contínues utilitzant el **Teorema de Sklar**, però no necessàriament posseïxen expressions simples.

Sigui $\mathbf{Y} \sim N_d(\mu, \Sigma)$ un vector aleatori normal multivariant, on $\mu = E[\mathbf{Y}] = (E[Y_1], \dots, E[Y_d])$ fa referència al vector mitjana i $\Sigma_{i,j} = Cov[Y_i, Y_j]$ per $1 \leq i \leq d$ i $1 \leq j \leq d$ a la matriu de covariància. La còpula associada a la variable $\mathbf{Y}$ és l'anomenada **Còpula Gaussiana**.

Donat que l'operació d'estandarditzar les marginals equival a aplicar una sèrie de transformacions estrictament creixents ($\varphi : Y_i \rightarrow \frac{Y_i - \mu_i}{\sigma_i}$, on σ_i fa referència a la desviació estàndard de Y_i), la *Proposició 2.3.7* implica que la còpula de $\mathbf{Y}$ és exactament la mateixa que la còpula de $\mathbf{X} \sim N_d(\mathbf{0}, P)$, on $P = \rho(\Sigma)$ és la matriu de correlació de $\mathbf{Y}$. Ara, per la *Definició 2.3.6*, aquesta còpula ve donada per

$$C_P^{Ga}(\mathbf{u}) = P(\phi(X_1) \leq u_1, \dots, \phi(X_d) \leq u_d) = \phi_P(\phi^{-1}(u_1), \dots, \phi^{-1}(u_d)) \quad (2.3.9)$$

on ϕ fa referència a la funció de distribució d'una normal estàndard univariant i ϕ_P fa referència a la funció de distribució conjunta de $\mathbf{X}$.

La còpula Gaussiana no té una forma simple tancada, però pot ser expressada com una integral sobre la densitat de $\mathbf{X}$. Donarem el cas de dos dimensions per $|\rho| < 1$, utilitzant (2.3.9)

$$C_\rho^{Ga}(u_1, u_2) = \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left(\frac{-(s_1^2 - 2\rho s_1 s_2 + s_2^2)}{2(1-\rho^2)}\right) ds_1 ds_2 \quad (2.3.10)$$

Com a curiositat, es pot observar com per $P = I_d$, la Còpula Gaussiana és igual a la **Còpula Independent**. Per $P = J_d$, on J_d és una matriu $d \times d$ que consisteix completament d'uns, obtenim la **Còpula de Comonotonicitat**, la qual s'obté a partir de variables aleatòries i

funcions de distribució que són perfectament positivament dependents. Finalment, per $d = 2$ i $\rho = -1$, obtenim la **Còpula de Contramonotònica**, obtinguda a partir de variables aleatòries que tenen funcions de distribució perfectament negativament dependents.

Així doncs, la **Còpula Gaussiana** es pot entendre com una estructura de dependència que interpola entre perfecta positiva i negativa dependència amb ρ representant la força d'aquesta dependència.

T-STUDENT

Així com la Còpula Gaussiana, l'anomenada **Còpula t-Student** també es troba dins de la categoria de **Còpules Implícites**.

Segui $\mathbf{X} \sim t_d(\nu, \mathbf{0}, P)$ un vector aleatori t multivariant, on $\nu > 0$ són els graus de llibertat i P és la matriu de correlació. Aleshores, la seva còpula és l'anomenada **Còpula t-Student**.

Ara, per la *Definició 2.3.6*, aquesta còpula ve donada per

$$C_{c,P}^t(\mathbf{u}) = t_{\nu,P}(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)) \quad (2.3.11)$$

on t_{ν} és la funció de distribució d'una distribució t univariant estàndard amb ν graus de llibertat i $t_{\nu,P}$ és la funció de distribució conjunta del vector aleatori $\mathbf{X}$.

Així com la Còpula Gaussiana, si $P = J_d$, aleshores obtenim la **Còpula de Comonotonicitat**. Tot i això, en el cas $P = I_d$, a diferència de la Còpula Gaussiana, no obtenim la Còpula Independent (assumint $\nu < \infty$), donat que variables aleatòries multivariants t-distribuïdes no correlacionades, no són independents.

El lector interessat pot profunditzar en aquesta última afirmació a [5].

ARXIMEDIANES (Clayton, Gumbel, Frank)

Les següents notes s'han extret de [2] i [3]. Les còpules presentades en aquesta secció, es troben dins de la categoria de **Còpules Explícites**, que inclou aquelles còpules que tenen expressions simples de forma tancada i segueixen construccions matemàtiques que se sap que donen còpules. Concretament, estudiarem les **Còpules Arximedianes Bivariants**.

Nota: En aquesta secció, només es donaran descripcions d'algunes de les còpules bivariants amb construccions més senzilles, donat que en el marc pràctic seran les que s'utilitzaran.

Tot seguit donarem una sèrie de Teoremes i Definicions d'acord amb [3] per tal de caracteritzar les **Còpules Arximedianes**

Teorema 2.3.15. (Còpula Arximediana bivariant) *Segui $\psi : [0, \infty) \rightarrow [0, 1]$ una funció decreixent i contínua que satisfà les condicions $\psi(0) = 1$ i $\lim_{t \rightarrow \infty} \psi(t) = 0$. Aleshores*

$$C(u_1, u_2) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2)) \quad (2.3.12)$$

és una còpula si i només si ψ és convexa

Definició 2.3.16. (Generador de Còpules Arximedianes) *Una funció decreixent, contínua i convexa $\psi : [0, \infty) \rightarrow [0, 1]$ que satisfà $\psi(0) = 1$ i $\lim_{t \rightarrow \infty} \psi(t) = 0$ és coneguda com a Generador de Còpules Arximedianes.*

La primera que estudiarem serà la **Còpula de Gumbell** que presenta la següent forma

$$C_{\theta}^{Gu}(u_1, u_2) = \exp(-((- \ln(u_1))^{\theta} + (- \ln(u_2))^{\theta})^{\frac{1}{\theta}}) \quad (2.3.13)$$

amb $1 \leq \theta < \infty$.

Observem que per $\theta = 1$, obtenim la **Còpula Independent** i que el límit de C_θ^{Gu} per $\theta \rightarrow \infty$ resulta en la **Còpula de Comonotonicitat** en dimensió 2.

Així doncs, la còpula descrita interpola entre independència i perfecte independència i el paràmetre θ representa la força de dita independència.

En segon lloc, presentem la **Còpula de Clayton**, que té la següent forma

$$C_\theta^{Cl}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta} \quad (2.3.14)$$

amb $0 \leq \theta < \infty$.

El cas $\theta = 0$ s'hauria d'interpretar com el límit de $\theta \rightarrow 0$, que representa la **Còpula Independent**. Per $\theta \rightarrow \infty$, ens acostem a la **Còpula de Comonotonicitat** en dimensió 2.

I finalment, tot i que hi ha moltes altres variants de les **Còpules Arximedians**, considerem com a últim exemple a estudiar la **Còpula de Frank**

$$C_\theta^{Fr}(u_1, u_2) = \frac{-1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right) \quad (2.3.15)$$

amb $\theta \in \mathbb{R}$.

Es pot verificar que, sempre que el paràmetre θ es trobi dins dels rangs que hem especificat a les definicions de les còpules, els tres exemples que hem descrit tenen la forma

$$C(u_1, u_2) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2)) \quad (2.3.16)$$

on $\psi : [0, \infty) \rightarrow [0, 1]$ és una funció que compleix amb les condicions establertes a la *Definició 2.3.16*.

Seguidament, es deixa una taula resumint els generadors, els rangs dels paràmetres i l'acotació inferior i superior de de les **Còpules Arximedians** definides.

Còpula	Generador $\psi(t)$	Rang del paràmetre	Inferior	Superior
C_θ^{Gu}	$\exp(-t^{1/\theta})$	$\theta \geq 1$	Π	M
C_θ^{Cl}	$\max((1 + \theta t)^{-1/\theta}, 0)$	$\theta \geq -1$	W	M
C_θ^{Fr}	$-\frac{1}{\theta} \ln(1 + (e^{-\theta} - 1)e^{-t})$	$\theta \in \mathbb{R}$	W	M

Al introduir la **Còpula de Clayton** prèviament, hem definit el seu paràmetre com a no negatiu. Tot i això, a la taula també s'ha definit el paràmetre per $-1 \leq \theta < 0$. Per adaptar-nos a aquest cas, el generador s'ha escrit de la següent manera $\psi(t) = \max((1 + \theta t)^{-1/\theta}, 0)$. Observem que $\psi(t)$ és estrictament decreixent a $[0, \frac{-1}{\theta}]$, però $\psi(t) = 0$ a $[\frac{-1}{\theta}, \infty)$. Per definir la inversa del generador únicament a zero, establim $\psi^{-1}(0) = \inf\{t : \psi(t) = 0\} = \frac{-1}{\theta}$.

2.4 Mesures de dependència

Aquesta secció pretén donar una sèrie d'eines de mesura de dependència tal i com es descriu a [2]. Aquestes eines funcionen com a mesures escalars de la "força de la dependència" per a un parell de variables aleatòries (X_1, X_2) .

2.4.1 Correlació lineal versus Correlació de Rang

En primer lloc, es presenta una visió general de les principals diferències entre la correlació lineal i la correlació de rang, amb l'objectiu d'entendre per què aquesta darrera és la més utilitzada en entorns de crisi i en estudis basats en còpules.

Correlació lineal

La correlació és protagonista en la teoria financera, però es important ressaltar que el concepte només és realment natural en el context dels models normals multivariants. Aquestes variables es descriuen mitjançant un vector de mitjanes i una matriu de covariància. Com que les mitjanes i les variàncies són característiques de les distribucions marginals, es pot considerar que les còpules d'aquestes distribucions depenen només de la matriu de correlació. Així doncs, la matriu de correlació té un paper paramètric natural en aquests models, que no té en models multivariants més generals. La nostra discussió es centrarà en les deficiències de la correlació lineal i els inconvenients subtils que presenta al allunyar-nos dels models normals multivariants.

Definim de manera informal la correlació $\rho(X_1, X_2)$ entre dues variables X_1 i X_2 com una mesura de dependència lineal que pren valors entre $[-1, 1]$. Si X_1 i X_2 són independents, aleshores $\rho(X_1, X_2) = 0$, però cal ressaltar que el contrari és fals, és a dir, la no correlació de X_1 i X_2 no implica en general la seva independència.

Exemple 2.4.1. Prenem $X_1 = Z \sim N(0, 1)$ i $X_2 = Z^2$. Aquestes dos variables aleatòries són clarament dependents però la seva correlació és 0.

Si $|\rho(X_1, X_2)| = 1$, aleshores equival a dir que X_1 i X_2 són perfectament linealment dependents, és a dir, $X_2 = \alpha + \beta X_1$ gairebé amb seguretat per a algun $\alpha \in \mathbb{R}$ i $\beta \neq 0$, amb $\beta > 0$ per dependència positiva i $\beta < 0$ per dependència negativa. A més a més, per $\beta_1, \beta_2 > 0$

$$\rho(\alpha_1 + \beta_1 X_1, \alpha_2 + \beta_2 X_2) = \rho(X_1, X_2),$$

és a dir, la correlació és invariant sota transformacions lineals estrictament creixents. Tanmateix, la correlació no és invariant sota transformacions no lineals estrictament creixents $T : \mathbb{R} \rightarrow \mathbb{R}$ (per exemple la transformació logarítmica). Per a dues variables aleatòries de valors reals, en general, tenim que $\rho(T(X_1), T(X_2)) \neq \rho(X_1, X_2)$.

Una altra observació òbvia, però important, és que la correlació només es defineix quan les variàncies de X_1 i X_2 són finites. En altres paraules, la correlació lineal no existeix per a distribucions on els segons moments no estan definits. Aquesta restricció a models de variància infinita fa que la correlació lineal no existeix per a distribucions de cua molt pesada on els segons moments no estan definits.

Correlació de Rang

Les correlacions de rang són mesures escalars simples de dependència que depenen només de la còpula d'una distribució bivariant i no de les distribucions marginals, a diferència de la correlació lineal, que depèn de les dues ja que utilitza els valors numèrics exactes de les variables.

Els estimadors empírics de les correlacions de rang es poden calcular simplement a partir dels rangs de les dades: només cal saber quin valor és més gran o més petit, i no els valors numèrics reals. Aquesta característica les fa especialment útils a la pràctica, ja que permeten calibrar còpules directament a partir de dades empíriques sense necessitat d'assumir o modelitzar les distribucions marginals.

A nivell teòric, les correlacions de rang tenen propietats més atractives que la correlació lineal, ja que són funcions directes de la còpula i, per tant, reflecteixen de manera més precisa la dependència pròpia de la distribució conjunta. Dins d'aquestes mesures, les dues més utilitzades són la **correlació de Kendall** i la **correlació de Spearman**, les quals seran discutides a continuació.

Per tal d'entendre al construcció de la τ -Kendall i la ρ de Spearman, cal, en primer lloc, donar la següent definició.

Definició 2.4.2. Dos punts a $\mathbb{R}^2$, denotats per (x_1, x_2) i $(\tilde{x}_1, \tilde{x}_2)$ s'anomenen concordants si $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$, i s'anomenen discordants si $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$

2.4.2 Tau de Kendall

Considerem un vector aleatori (X_1, X_2) i una còpia independent $(\tilde{X}_1, \tilde{X}_2)$ (és a dir, un segon vector amb la mateixa distribució, però independent del primer). Si X_2 tendeix a augmentar amb X_1 , aleshores esperem que la probabilitat de concordança sigui alta en relació amb la probabilitat de discordància. Si X_2 tendeix a disminuir amb l'augment de X_1 , aleshores esperem el contrari. Aquest fet motiva la correlació de rang de Kendall, que és simplement la probabilitat de concordança menys la probabilitat de discordància per a aquests parells.

$$\rho_\tau(X_1, X_2) = P((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0) - P((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0) \quad (2.4.1)$$

Seguidament, donem una versió més compacte d'aquesta última fórmula.

Definició 2.4.3. Per variables aleatòries X_1 i X_2 la τ de Kendall ve donada per

$$\rho_\tau(X_1, X_2) = E(\text{sign}((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2))) \quad (2.4.2)$$

on $(\tilde{X}_1, \tilde{X}_2)$ és una còpia independent de (X_1, X_2) i $\text{sign}(x) = I_{\{x>0\}} - I_{\{x<0\}}$.

A continuació, veurem que per variables aleatòries amb funcions de distribució contínues, la τ de Kendall només depèn de la còpula única C de (X_1, X_2) i donarem una fórmula explícita per a calcular-la.

Proposició 2.4.4. Suposem que X_1 i X_2 tenen distribucions marginals contínues i una còpula única. Aleshores

$$\rho_\tau(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 \quad (2.4.3)$$

La prova d'aquest resultat queda recollida a [8].

2.4.3 Rho de Spearman

Aquesta mesura també es pot definir en termes de la concordança i la discordància de dos vectors aleatoris bivariants.

Definició 2.4.5. Per a variables aleatòries X_1 i X_2 , la ρ de Spearman ve donada per

$$\rho_S = 3(P((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0) - P((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0)), \quad (2.4.4)$$

on $\tilde{X}_1$ i $\tilde{X}_2$ són variables aleatòries que satisfàn $\tilde{X}_1 \stackrel{d}{=} X_1$ i $\tilde{X}_2 \stackrel{d}{=} X_2$ i on (X_1, X_2) , $\tilde{X}_1$ i $\tilde{X}_2$ són totes independents.

No és immediat veure que aquesta definició doni una mesura de correlació dins de l'interval $[-1, 1]$. La següent proposició dona una representació alternativa d'aquesta mesura, que ho deixa clar per a variables aleatòries contínues.

Proposició 2.4.6. Si X_1 i X_2 tenen distribucions marginals contínues F_1 i F_2 , aleshores $\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2))$

En el cas de variables aleatòries bivariants, la formula de la ρ de Spearman en termes de la còpula C de (X_1, X_2) s'expressa de la següent manera.

Proposició 2.4.7. *Suposem que X_1 i X_2 tenen distribucions marginals contínues i una única còpula C . Aleshores*

$$\rho_S(X_1, X_2) = 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) du_1 du_2 \quad (2.4.5)$$

La prova d'aquests dos últims resultats queda recollida a [8].

2.4.4 Propietats de la correlació de rang

La τ de Kendall i la ρ de Spearman tenen moltes propietats en comú. Ambdues són mesures de dependència simètriques que prenen valors en l'interval $[-1, 1]$. Donen el valor zero per a variables aleatòries independents, tot i que una correlació de rang de 0 no implica necessàriament independència. Es pot demostrar que prenen el valor 1 si i només si X_1 i X_2 són comonòtones (veure [8]) i el valor -1 si i només si són contramonòtones (veure [8]). A més a més, són invariants sota transformacions estrictament creixents de X_1 i X_2 .

2.4.5 Dependència de cues: coeficient superior i inferior

Abans de començar amb aquesta nova secció, cal introduir un nou concepte d'acord amb [2]. Sigui $\mathbf{X}$ un vector aleatori amb funció de supervivència multivariant $\bar{F}$ i funcions marginals de supervivència $\bar{F}_1, \dots, \bar{F}_d$, és a dir $\bar{F}_i = 1 - F_i$. Tenim la següent identitat

$$\bar{F}(x_1, \dots, x_d) = \hat{C}(\bar{F}_1(x_1), \dots, \bar{F}_d(x_d)), \quad (2.4.6)$$

per a una còpula $\hat{C}$, que es coneix com a **còpula de supervivència**.

Una relació important entre una còpula i la seva còpula de supervivència en el cas bivariant és la següent

$$\hat{C}(1 - u_1, 1 - u_2) = 1 - u_1 - u_2 + C(u_1, u_2). \quad (2.4.7)$$

Igual que les correlacions de rang, els coeficients de dependència de la cua són mesures de dependència que depenen només de la còpula d'un parell de variables aleatòries X_1 i X_2 amb funcions de distribució marginals contínues. La motivació per examinar aquests coeficients és que proporcionen mesures sobre la força de la dependència a les cues d'una distribució bivariant.

En el cas de la dependència de la cua superior, observem la probabilitat que X_2 superi el seu q -quantil, quan X_1 supera el seu q -quantil, i després considerem el límit quan q tendeix a 1. Òbviament, els rols de X_1 i X_2 són intercanviables. Passem ara a donar unes definicions més formals.

Definició 2.4.8. *Sigui X_1 i X_2 variables aleatòries amb funcions de distribució F_1 i F_2 respectivament. El coeficient de dependència de la cua superior de X_1 i X_2 és*

$$\lambda_u := \lambda_u(X_1, X_2) = \lim_{q \rightarrow 1^-} P(X_2 > F_2^{\leftarrow}(q) | X_1 > F_1^{\leftarrow}(q)), \quad (2.4.8)$$

sempre que existeixi un límit $\lambda_u \in [0, 1]$. Si $\lambda_u \in (0, 1]$, aleshores X_1 i X_2 es diu que mostren dependència de cua superior o extrema dependència a la cua superior. Si $\lambda_u = 0$, aleshores es

diu que són asimptòticament independents a la cua superior. De manera anàloga, el coeficient de dependència de la cua inferior és

$$\lambda_l = \lambda_l(X_1, X_2) = \lim_{q \rightarrow 0^+} P(X_2 \leq F_2^{\leftarrow}(q) | X_1 \leq F_1^{\leftarrow}(q)), \quad (2.4.9)$$

sempre que existeixi un límit $\lambda_l \in [0, 1]$.

Si F_1 i F_2 son funcions de distribució contínues, aleshores s'obtenen expressions simples per λ_u i λ_l en termes de la única còpula C de la distribució bivariant. Utilitzant probabilitat condicional elemental i la igualtat (2.4.7), obtenim

$$\lambda_l = \lim_{q \rightarrow 0^+} \frac{P(X_2 \leq F_2^{\leftarrow}(q), X_1 \leq F_1^{\leftarrow}(q))}{P(X_1 \leq F_1^{\leftarrow}(q))} = \lim_{q \rightarrow 0^+} \frac{C(q, q)}{q} \quad (2.4.10)$$

Per la dependència de la cua superior tenim

$$\lambda_u = \lim_{q \rightarrow 1^-} \frac{P(X_2 > F_2^{\leftarrow}(q), X_1 > F_1^{\leftarrow}(q))}{P(X_1 > F_1^{\leftarrow}(q))} = \lim_{q \rightarrow 1^-} \frac{\hat{C}(1-q, 1-q)}{1-q} = \lim_{q \rightarrow 0^+} \frac{\hat{C}(q, q)}{q} \quad (2.4.11)$$

on $\hat{C}$ és la còpula de supervivència de C .

Per últim, es dona una taula que recull els diferents valors teòrics de la τ de Kendall, de la ρ de Spearman i de la dependència de cues de les diferents famílies de còpules que s'han vist prèviament.

Còpula	ρ_τ	ρ_S	λ_u	λ_l
C_θ^{Ga}	$(2/\pi) \arcsin \theta$	$(6/\pi) \arcsin \frac{1}{2}\theta$	0	0
$C_{\nu, \theta}^t$	$(2/\pi) \arcsin \theta$		$2t_{\nu+1}(-\sqrt{\frac{(\nu+1)(1-\theta)}{1+\theta}})$	$2t_{\nu+1}(-\sqrt{\frac{(\nu+1)(1-\theta)}{1+\theta}})$
C_θ^{Gu}	$1 - \frac{1}{\theta}$		$2 - 2^{1/\theta}$	0
C_θ^{Cl}	$\frac{\theta}{\theta+2}$	$\theta/(\theta+3)$	0	$\begin{cases} 2^{-1/\theta}, & \theta > 0, \\ 0, & \theta \leq 0, \end{cases}$
C_θ^{Fr}	$1 - 4\theta^{-1}(1 - D_1(\theta))$		0	0

Nota: La informació que no apareix en alguna de les columnes és degut a que algunes de les còpules no compten amb una fórmula simple tancada.

Nota: $D_1(\theta) = \int_1^\theta \frac{t/\theta}{e^t - 1} dt$.

Capítol 3

Dades i processament

3.1 Fonts de dades i criteris de selecció d'empreses concursades

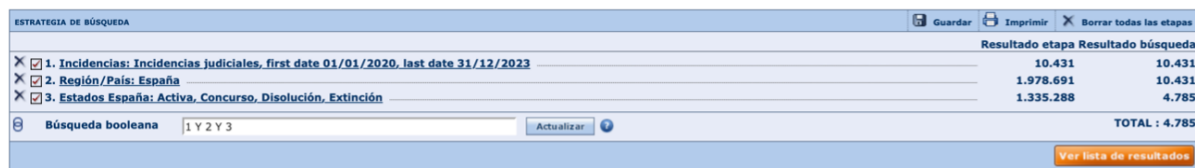
SABI (**Sistema de Análisis de Balances Ibéricos**) és una base de dades que proporciona la companyia *Bureau Van Dijk*, una companyia global dedicada a la recopilació i l'anàlisi de dades. Al ser membres de la *Universitat de Barcelona (UB)*, disposem d'una subscripció bàsica que ens permet accedir a l'esmentada base de dades.

SABI conté informació comptable i financera sobre gairebé 3 milions d'empreses Espanyoles i 900.000 empreses Portugueses.

L'objectiu d'aquest estudi és identificar empreses espanyoles que hagin iniciat un procediment concursal durant el període **2020 – 2023**. L'elecció de l'interval temporal respon a les moratòries concursals aprovades pel Govern espanyol entre 2020 – 2022, que suspènien temporalment l'obligació legal de sol·licitar concurs tot i trobar-se en situació d'insolvència. Aquestes mesures van provocar una acumulació d'empreses en risc d'insolvència que no van iniciar el procediment fins que es va finalitzar el període excepcional. Com a resultat, els anys 2020–2023 concentren l'impacte immediat de la crisi, fet que justifica l'elecció d'aquest període.

Tot i això, cal tenir present que les dades comptables d'empreses insolvents poden presentar limitacions de fiabilitat. Durant el procés concursal moltes empreses redueixen la seva activitat, presenten retards en el dipòsit dels comptes, registres incomplets o irregularitats en les partides de balanç i compte de resultats. A més, les empreses extingides deixen d'aportar informació a partir del moment de la dissolució. Per tant, les dades disponibles reflecteixen els comportaments financer anteriors al concurs, però no sempre reflecteixen de manera completa la trajectòria posterior.

SABI no ofereix l'opció d'aplicar un filtre “*d'any d'entrada en concurs*”. Així doncs, l'estratègia adoptada consisteix en filtrar per empreses Espanyoles en estat **Actiu, Concurs, Dissolució o Extinció** que hagin registrat **una incidència judicial** en aquest interval temporal, donat que l'obertura d'un procediment concursal implica necessàriament l'existència d'una incidència judicial (veure Figura 3.1).



The screenshot shows the 'ESTRATEGIA DE BÚSQUEDA' (Search Strategy) interface. It includes a table of filters and a results table. The filters are: 1. Incidencias judiciales, first date 01/01/2020, last date 31/12/2023; 2. Región/País: España; 3. Estados España: Activa, Concurso, Disolución, Extinción. The results table shows the number of results for each filter and a total of 4,785.

ESTRATEGIA DE BÚSQUEDA		Resultado etapa Resultado búsqueda	
☒ 1. Incidencias judiciales, first date 01/01/2020, last date 31/12/2023		10.431	10.431
☒ 2. Región/País: España		1.978.691	10.431
☒ 3. Estados España: Activa, Concurso, Disolución, Extinción		1.335.288	4.785
Búsqueda booleana 1 Y 2 Y 3		TOTAL : 4.785	

Figura 3.1: Filtre aplicat

De cada una de les empreses que complien amb les característiques del filtre aplicat, s'ha extret:

- **Informació no financera:** NIF, CNAE, Comunitat Autònoma, Estat (Activa, Concurs, Dissolta, Extingida), Detalls de l'Estat, Data de canvi d'estat, Últim any amb informació disponible i Nombre d'Empleats (veure Figura 3.2).
- **Informació financera:** Actiu Total, Actiu Circulant, Deutors, Tresoreria, Fons Propis, Passiu Total, Passiu líquid, Deutes financers, Creditors, Resultat d'Explotació, Import Net de la Xifra de Vendes, Resultat de l'Exercici, Fons de Maniobra, Període de cobrament, Període de crèdit... (veure Figura 3.3).

The screenshot shows a table with columns for company name, CNAE, autonomous community, and various financial metrics. The table is filtered to show companies in the 'Activa' status. The data includes company names like ORANGE ESPAGNE SA, FCC CONSTRUCCION SA, and others, along with their respective financial figures.

Figura 3.2: Informació no financera

The screenshot shows a table with columns for company name, CNAE, autonomous community, and various financial metrics. The table is filtered to show companies in the 'Activa' status. The data includes company names like ORANGE ESPAGNE SA, FCC CONSTRUCCION SA, and others, along with their respective financial figures.

Figura 3.3: Informació financera

A partir de l'exportació d'aquesta informació s'ha construït la base de **Raw Data**, que serveix com a punt de partida per a la fase de neteja i tractament.

3.2 Neteja, tractament de valors extrems i normalització

Un cop obtinguda la base de dades es van seguir els següents passos:

1. **Eliminació d'empreses amb canvi d'estat anterior a 2020**, ja que no poden haver iniciat un concurs durant el període d'interès.
2. **Eliminació d'empreses sense informació financera com a mínim fins el 2020**, requisit mínim per tal de calcular ràtios. S'ha escollit el 2020 com a data mínima perquè les empreses que han entrat en concurs durant el 2020 poden no tenir informació els anys següents.

3. **Filtrat d'empreses en estat Actiu:** Es van eliminar aquelles que en l'Estat Detallat no s'esmentava que havia finalitzat el concurs (Veure Figura 3.4).
4. **Substitució de valors extrems per NaN:** Alguns camps presentaven valors no realistes (p.e. 5000 dies de període cobrament).
5. **Validacions comptables:** Es va comprovar que hi hagués una relació coherent entre les partides de Actiu, Patrimoni Net i Passiu (p.e. $\text{Actiu} = \text{Passiu} + \text{Patrimoni Net}$).
6. **Eliminació d'empreses amb partides impossibles:** No s'han tingut en compte empreses que presentessin valors negatius en partides que, per criteri comptable, no poden ser negatives (Tresoreria, Deutors, Existències...). Agafem com a exemple la partida de Tresoreria, aquesta no pot ser negativa donat que si ho fos, hauria d'estar registrada com a deute financer al Passiu. Un raonament similar es pot aplicar a la resta de partides del balanç les quals tampoc poden ser negatives.

	Nombre	Código NIF	Código primario CNAE 2009	Comunidad autónoma	Estado	Estado detallado
101. ☒	SUNSET BEACH CLUB SA	A29182458	5510	Andalucía	Activa	Activa
102. ☒	ONE FACILITY SERVICES SL	B20572871	8121	País Vasco	Activa	Activa
103. ☒	CONSCYTEC SOCIEDAD LIMITADA	B02335347	4121	Castilla-La Mancha	Activa	Activa
104. ☒	ARTIFEX INFRAESTRUCTURAS SL	B65675365	4121	Cataluña	Activa	Activa
105. ☒	SOCIEDAD MUNICIPAL DE APARCAMIENTOS Y SERVICIOS SA	A29178902	5221	Andalucía	Activa	Activa
106. ☒	TRAZO GESTION INTEGRAL DE PROYECTOS SL	B87014973	4121	Madrid	Activa	Reapertura de hoja registral...
107. ☒	OBRES I SERVEIS ROIG SA	A08743098	6831	Cataluña	Activa	Activa
108. ☒	NOPAR SPAIN SL	B65737264	5320	Cataluña	Activa	Reapertura de hoja registral...
109. ☒	COTS I CLARET S.L.	B08136905	4121	Cataluña	Activa	Activa
110. ☒	FOODBOX SA	A87240420	5610	Madrid	Activa	Activa
111. ☒	TEGESTACIN SL (EN LIQUIDACION)	B70245386	1054	Madrid	Disuelta	Disolución: La empresa se e...
112. ☒	MOTOR GOMEZ VILLALBA SA (EXTINGUIDA)	A81695702	4511	Madrid	Extinta	Extinción: La empresa se e...
113. ☒	FOMENTO VALENCIA MANTENIMIENTO Y LIMPIEZA, SOCIEDA...	A96062948	8129	Comunidad Valenciana	Activa	Activa
114. ☒	ESDITRA SA	A38037438	8531	Canarias	Activa	Activa
115. ☒	GEOTECNIA Y CIMENTOS SA (EXTINGUIDA)	A28208874	4299	Madrid	Extinta	Extinción: La empresa se e...
116. ☒	BILFINGER INDUSTRIAL SERVICES SPAIN SA (EXTINGUIDA)	A28115319	4329	Madrid	Extinta	Extinción: La empresa se e...
117. ☒	MAJORICA SAU (EN LIQUIDACION)	A07333560	3213	Baleares	Disuelta	Disolución: La empresa se e...
118. ☒	GRUPO ADL S.COOP. ANDALUZA	F14961262	8811	Andalucía	Activa	Finalización del concurso: L...
119. ☒	CONSTRUCCIONES Y PROMOCIONES GRANULARES SOCIEDAD...	B76198670	4121	Canarias	Disuelta	Disolución: La empresa se e...
120. ☒	DIASA INDUSTRIAL SAU (EN LIQUIDACION)	A31064033	2059	Murcia	Disuelta	Liquidación concursal: La e...
121. ☒	SYS OUTSOURCING SA (EXTINGUIDA)	A61668349	8299	Madrid	Extinta	Extinción: La empresa se e...
122. ☒	ELEROC SERVICIOS SOCIEDAD LIMITADA	B26533257	8812	La Rioja	Activa	Activa
123. ☒	EXTERNA SERVICIOS GENERALES DE EMPRESA SL	B70061577	4619	Madrid	Activa	Activa
124. ☒	SORTIGUE-ACSA CONSERVACION DE INFRAESTRUCTURAS SA	A85110294	4211	Madrid	Activa	Activa
125. ☒	SALDODIPJE SL	B92941558	6499	Cataluña	Activa	Activa

Figura 3.4: Exemple d'estat detallat vàlid

Aquest procés garanteix que l'anàlisi es realitzi a partir d'empreses amb informació consistent i suficient. Així doncs, a l'inici la base de *Raw Data* comptava amb 4785 empreses, al finalitzar-lo la base de dades final compta amb **1308 empreses**.

3.3 Construcció de “buckets” CNAE X Comunitat Autònoma

L'aplicació de còpules requereix sèries temporals amb un **nombre mínim d'observacions** per obtenir estimacions robustes de dependència. No obstant, a mesura que avança el període temporal, el nombre d'empreses amb informació disponible disminueix notablement (Figura 3.5), fet que impedeix analitzar sèries individualment per empresa.

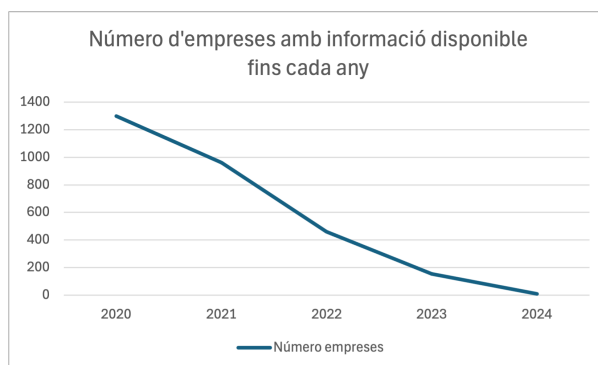


Figura 3.5: Gràfica que mostra la disminució d'informació disponible per any

Per aquest motiu, les empreses s'han agrupat segons la seva ubicació mitjançant **“bucket” per Comunitat Autònoma**. D'aquesta manera obtenim 18 “bucket” (veure Figura 3.6 on queda representat el nombre d'empreses que hi ha en cada “bucket”) el que ens permet obtenir sèries temporals amb més observacions i reduir la volatilitat.

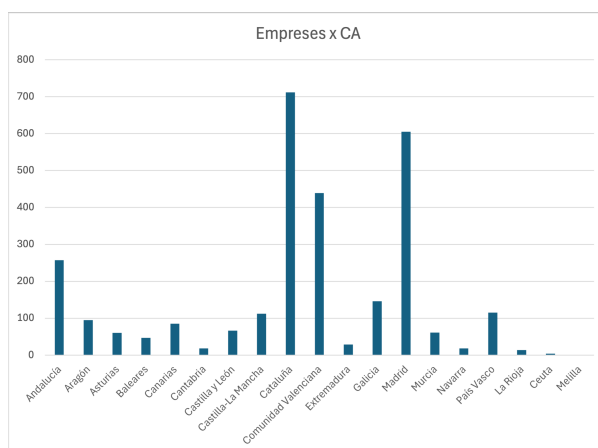


Figura 3.6: Nombre d'empreses que figuren en cada “bucket”

3.4 Càlcul de les ràtios financeres

A partir de la mostra, s'han calculat les següents ràtios prèviament introduïdes en el marc teòric:

$$R1 = \frac{\text{Actiu Corrent} - \text{Existències}}{\text{Actiu Total}}$$

$$R2 = \frac{\text{Actiu Corrent}}{\text{Passiu Corrent}}$$

$$R3 = \frac{\text{Passiu}}{\text{Patrimoni Net}}$$

$$R4 = \frac{\text{Compte de Proveïdors}}{\text{Comptes}} - \frac{\text{Compte de Clients}}{\text{Vendes}} = PMP - PMC$$

$$R5 = ROA = \frac{\text{Resultat d'Explotació}_{31/12/X+1}}{\text{Actiu Total}_{31/12/X}}$$

$$R6 = \text{Marge Brut} = \frac{\text{Resultat d'Explotació}_{31/12/X+1}}{\text{Vendes}_{31/12/X+1}}$$

$$R7 = ROE = \frac{\text{Benefici Net}_{31/12/X+1}}{\text{Patrimoni Net}_{31/12/X}}$$

$$R8 = \text{Marge Net} = \frac{\text{Benefici Net}_{31/12/X+1}}{\text{Vendes}_{31/12/X}}$$

$$R9 = \text{Palanquejament} = \frac{\text{Actiu Total}_{31/12/X}}{\text{Patrimoni Net}_{31/12/X}}$$

[illegible]

Figura 3.7: Base de dades amb el càlcul de les ràtios

Cal remarcar que, un cop obtinguda la base de dades amb les ràtios, i amb l'objectiu d'augmentar el rigor dels resultats, s'han identificat aquelles ràtios que presentaven valors poc realistes (p. ex. un valor de palanquejament de -103063) i s'han tractat com a *NaN*.

A la pràctica, tal com s'ha exposat prèviament, no totes les empreses disposen de la informació financera necessària per calcular les nou ràtios en tots els períodes analitzats. Per aquest motiu, en comparar dues ràtios concretes, només es consideraran aquelles empreses que disposin simultàniament de les dades requerides per a ambdues ràtios (és a dir, que les dues ràtios tinguin valors diferents a NaN). Aquesta restricció garanteix que les estimacions de dependència no es vegin afectades per problemes de dades mancants ni per diferències en la composició de la mostra.

A més, s’ha decidit que es donaran conclusions basades en les còpules calculades amb un mínim de 30 observacions per “bucket”. Aquest llinar s’ha escollit perquè, amb menys dades, els resultats serien massa inestables i podrien variar molt només per l’efecte de l’atzar. Tot i que 30 observacions no garanteixen una precisió perfecta, és un valor que s’utilitza sovint en estudis empírics i ens permet confiar que les conclusions obtingudes són prou consistents i útils.

Aquests criteris de selecció reforcen la robustesa del resultat final i assegurin que les conclusions extretes siguin rigoroses, comparables i basades en informació suficientment fiable.

3.5 Segmentació temporal (pre, COVID, post)

Per tal d'estudiar l'evolució de la dependència entre les ràtios, s'han definit tres períodes:

1. **Pre – COVID:** Període previ a la crisi durant els anys 2017 – 2019. S’espera observar ràtios dins dels llinars de normalitat així com una dependència moderada.
2. **COVID:** Període d’estrès financer durant els anys 2020 - 2021. La part teòrica suggereix valors de ràtios més extrems indicant risc d’insolvència i major dependència de cues.
3. **Post – COVID:** Període de recuperació durant els anys 2022 - 2023. S’espera que les empreses que han sobreviscut experimentin una millora en ràtios i una reducció de la dependència extrema, tot i que la informació disponible disminueix per l’extinció d’empreses insolvents.

Aquest esquema temporal permet capturar patrons d'avaluació i comparar l'impacte de la crisi sobre la dependència entre ràtios.

Capítol 4

Metodologia Empírica

4.1 Modelització de la distribució marginal de cada ràtio

En situacions pràctiques, en nombroses ocasions necessitem estimacions de les marginals per formar un model multivariant complet, és a dir, ens veiem obligats a estimar les marginals per tal d'estimar la còpula.

En aquesta secció, seguint [16] descriurem breument alguns enfocaments generals per a estimar les marginals i construir una mostra de pseudo-observacions a partir de la còpula.

Definim $\hat{F}_1, \dots, \hat{F}_d$ com l'estimació de les funcions de distribució de les marginals. La mostra de pseudo-observacions de la còpula consisteix en els vectors $\hat{\mathbf{U}}_1, \dots, \hat{\mathbf{U}}_n$ on

$$\hat{\mathbf{U}}_t = (\hat{U}_{t,1}, \dots, \hat{U}_{t,d})' = (\hat{F}_1(X_{t,1}), \dots, \hat{F}_d(X_{t,d}))'$$

Cal remarcar que tot i que els vectors originals $X_1, \dots, X_n$ siguin independents, la mostra de pseudo-observacions és generalment dependent, donat que les estimacions marginals $\hat{F}_i$ estan construïdes en la majoria de casos a partir de tots els vectors de dades originals a través de les mostres univariants $X_{1,i}, \dots, X_{n,i}$.

Passem a descriure els possibles mètodes per tal d'obtenir l'estimació de les funcions definides.

1. **Estimació paramètrica.** Aquest mètode consisteix en escollir un model paramètric (t-Student, Gaussià,...) adequat per a les dades observades i ajustar-lo mitjançant *Màxima Versemblança* (en la següent secció es tractarà aquest mètode).
2. **Estimació no paramètrica amb variant de la funció de distribució empírica.** Podríem estimar F_j utilitzant

$$F_{i,n}^*(x) = \frac{1}{n+1} \sum_{t=1}^n I_{\{X_{t,i} \leq x\}},$$

aquest mètode es discutirà amb més profunditat més endavant.

3. **Teoria de valors extrems per a les cues.** Les funcions de distribució empíriques poden ser en alguns casos estimadors poc fiables en quant a la distribució subjacent a les cues. Una alternativa en aquests cas és utilitzar la **Teoria de Valors Extrems**, on les cues es modelen semiparamètricament utilitzant una distribució de Pareto Generalitzada (GPD).

4.1.1 Transformació a escala uniforme (pseudo-observacions)

Tot i que es menciona la Teoria de Valors Extrems, no s'aprofundirà en la seva aplicació, ja que es tracta d'un tema molt ampli que queda fora de l'abast d'aquest treball.

El lector interessat pot consultar [17] per obtenir més informació sobre la Teoria dels Valors Extremes.

El primer mètode consisteix a ajustar una família paramètrica de distribució a cada ràtio i seleccionar la funció que millor s'adapti a les dades. No obstant, aquesta estratègia presenta diverses dificultats. Les ràtios financeres mostren distribucions fortament asimètriques, amb valors extrems i comportaments diferents segons el “bucket” i el període, fet que dificulta identificar una família paramètrica que sigui adequada de manera consistent. Aquest fet obligaria a realitzar un procés d'ajust, comparació i validació paramètrica molt ampli, cosa poc eficient i estadísticament fràgil donada la mida reduïda de moltes mostres. Per aquest motiu, el mètode paramètric s'ha descartat.

Per tant, en aquest estudi s'utilitzarà el segon mètode, que consisteix a estimar les funcions de distribució $\hat{F}_i$ mitjançant l'estimador empíric basat en rangs.

Passem a descriure com implementarem el **mètode 2**. Partim d'una mostra $X_1, \dots, X_n \in \mathbb{R}^d$, $X_i = (X_{i1}, \dots, X_{id})$. Per a cada component j (és a dir, cada variable marginal), tenim les dades $(X_{1j}, \dots, X_{nj})$.

Definim la funció de distribució empírica de la marginal j de la següent manera

$$F_{n,j}(x) = \frac{1}{n} \sum_{k=1}^n I_{\{X_{kj} \leq x\}},$$

on $I\{\cdot\}$ és la funció indicadora. En altres paraules, la part dreta de la igualtat ens dona informació sobre la fracció de les observacions X_{kj} que són menors o iguals que x .

Seguidament, prenem un dels valors de la mostra, X_{ij} , per exemple, i hi avaluem la funció de distribució empírica

$$F_{n,j}(X_{ij}) = \frac{1}{n} \sum_{k=1}^n I_{\{X_{kj} \leq X_{ij}\}},$$

en aquesta cas, la suma $\sum_{k=1}^n I_{\{X_{kj} \leq X_{ij}\}}$ compta quants X_{ij} són menors o iguals que X_{ij} . Per definició, aquesta suma és el rang de X_{ij} dins la columna j

$$R_{ij} = \#\{k : X_{kj} \leq X_{ij}\}.$$

Així doncs, $F_{n,j}(X_{ij}) = \frac{R_{ij}}{n}$. Ara, per la *Proposició 2.3.3*, tenim que $F_{n,j}(X_{ij}) \sim U(0,1)$. Si prenguèssim simplement $U_{ij} = F_{n,j}(X_{ij}) = \frac{R_{ij}}{n}$, un dels valors de U_{ij} podria ser exactament 1, cosa que comportaria problemes en l'avaluació de la densitat al implementar estimació de màxima versemblança per trobar els paràmetres de la còpula. Així doncs, fem un petit canvi i obtenim

$$U_{ij} = \frac{n}{n+1} F_{n,j}(X_{ij}) = \frac{1}{n+1} \sum_{k=1}^n I_{\{X_{kj} \leq X_{ij}\}} = \frac{R_{ij}}{n+1}$$

Les construccions $U_1, \dots, U_n$ són conegudes com pseudo-observacions i considerades com a observacions de la còpula C .

4.2 Estimació de paràmetres de còpules per MLE

Aquesta secció es dedicarà a exposar el mètode que s'utilitzarà a la part pràctica per tal d'estimar els paràmetres θ de les còpules paramètriques C_θ .

Passem a presentar el mètode d'estimació de **Màxima Versemblança**.

Segui C_θ una còpula paramètrica, on θ és un vector de paràmetres a estimar. L'estimador de Màxima Versemblança (*Maximum Likelihood Estimator (MLE)*) s'obté al maximitzar

$$\ln L(\theta; \hat{U}_1, \dots, \hat{U}_n) = \sum_{t=1}^n \ln c_\theta(\hat{U}_t)$$

respecte a θ , on c_θ denota la densitat de la còpula (en la següent nota hi figura la definició) i $\hat{U}_t$ denota una pseudo-observació de la còpula.

Nota: Definim la densitat d'una còpula c de la següent manera

$$c(u_1, \dots, u_d) = \frac{\partial C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d}$$

Òbviament, la qualitat estadística de les estimacions dels paràmetres de la còpula depèn en gran mesura de la qualitat de les estimacions de les distribucions marginals utilitzades en la creació de la pseudo-mostra de la còpula.

El lector pot referir-se a [18] per tal de trobar exemples sobre com s'aplica MLE a les còpules Gaussians i t-Student.

4.3 Selecció de models: AIC, BIC i tests de bondat d'ajust

Nota: En estadística, un model es considera imbricat dins d'un altre quan es pot obtenir com una versió més simple del model general imposant restriccions sobre alguns dels seus paràmetres, habitualment fixant-los a valors concrets (sovint zero).

Abans de continuar amb l'explicació de la part metodològica, cal aclarir uns quants conceptes estadístics que ens serviran més endavant per entendre com el programa RStudio produeix els resultats. D'acord amb [19], definirem els models de bondat d'ajust anomenats **criteris d'informació**. Aquests criteris intenten aconseguir un equilibri entre l'estimació de màxima versemblança i el nombre de paràmetres, penalitzant l'addició de paràmetres per augmentar l'estimació de màxima versemblança. Passem a explicar-los breument.

A la pràctica, sovint volem comparar models no imbricats amb un nombre de paràmetres possiblement força diferents. Suposem que tenim m models $M_1, \dots, M_m$ i que el model j té k_j paràmetres denotats per $\theta_j = (\theta_{j1}, \dots, \theta_{jk_j})'$ i una funció de probabilitat $L_j(\theta_j; \mathbf{X})$.

En l'enfocament del model denominat **Akaike Criteri d'Informació (AIC)**, escollim aquell model que minimitza

$$AIC(M_j) = -2 \ln L_j(\hat{\theta}_j; X) + 2k_j,$$

on $\hat{\theta}_j$ denota el **MLE** de θ_j . En el programa que es presentarà més endavant, aquest serà el criteri que s'utilitzarà per tal de determinar quina còpula s'ajusta millor a les dades.

Una alternativa al model anterior, és l'anomenat **Bayesian Criteri d'Informació (BIC)** definit de la següent manera

$$BIC(M_j) = k_j \ln(n) - 2 \ln L_j(\hat{\theta}_j; X)$$

on n fa referència a la mida de la mostra del model M_j . En aquest cas, també es seleccionen aquells models que presentin un BIC més baix.

El lector pot referir-se a [20] per obtenir més informació sobre aquests models.

4.4 Ajust paramètric (Gaussiana, t-Student i Arximedianes)

4.4.1 Algoritmes de simulació

En aquesta secció, deixem uns algoritmes en dimensió 2 resumint els passos a seguir partint de variables uniformes per tal de simular les diferent còpules.

Nota: A la pràctica, no utilitzarem aquests algoritmes concrets per tal de simular les còpules donat que el programari **RStudio** ofereix alternatives més eficients. Tot això, és rellevant mostrar com es realitzaria la simulació sense les facilitats del programa.

Algoritme simulació de la Còpula Gaussiana

1. Generem $\tilde{U}_1, \tilde{U}_2 \sim U(0, 1)$.

2. Definim

$$\begin{aligned} Z_1 &= \phi^{-1}(\tilde{U}_1) \\ Z_2 &= \phi^{-1}(\tilde{U}_2) \end{aligned}$$

on ϕ és la funció de distribució d'una normal estàndard, de manera que $Z_1, Z_2 \sim N(0, 1)$

3. Definim

$$\begin{aligned} Y_1 &= Z_1 \\ Y_2 &= \rho Z_1 + \sqrt{1 - \rho^2} Z_2 \end{aligned}$$

Cal remarcar que en aquest pas desconexim el valor de ρ . A partir d'aquí, obtenim

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \mathbf{A} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho^2} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \Rightarrow \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N(0, \rho)$$

on $\rho = \mathbf{A}\mathbf{A}^T = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. En aquest pas $\rho = \text{Corr}(Y_1, Y_2)$

4. Definim

$$\begin{aligned} U_1 &= \phi(Y_1) \\ U_2 &= \phi(Y_2) \end{aligned}$$

Així doncs, obtenim $U_1, U_2 \sim C_\rho^{Ga}$.

Algoritme simulació de la Còpula t-student

1. Generem $\tilde{U}_1, \tilde{U}_2, \tilde{U}_3 \sim U(0, 1)$.

2. Definim

$$\begin{aligned} Z_1 &= \phi^{-1}(\tilde{U}_1) \\ Z_2 &= \phi^{-1}(\tilde{U}_2) \end{aligned}$$

on ϕ és la funció de distribució d'una normal estàndard, de manera que $Z_1, Z_2 \sim N(0, 1)$

3. Introduïm correlació

$$Y_1 = Z_1$$

$$Y_2 = \rho Z_1 + \sqrt{1 - \rho^2} Z_2$$

de manera que $(Y_1, Y_2) \sim N(0, R)$ amb $R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Cal remarcar que $\rho = \text{Corr}(Y_1, Y_2)$.

4. Passem a generar una variable $W \sim \chi_\nu^2$ que segueixi una distribució χ^2 amb ν graus de llibertat, independent de (Y_1, Y_2) (d'aquesta manera, la còpula que generem tindrà cues pesades)

$$W = F_{\chi_\nu^2}^{-1}(\tilde{U}_3)$$

5. Definim la t-Student bivariant

$$T_1 = \frac{Y_1}{\sqrt{W/\nu}}$$

$$T_2 = \frac{Y_2}{\sqrt{W/\nu}}$$

Així doncs, (T_1, T_2) tenen una distribució t-Student bivariant amb ν graus de llibertat, marginal t_ν i correlació ρ .

6. Seguidament, utilitzem la funció de distribució d'una t-Student univariant F_{t_ν} per simular la còpula t-Student

$$U_1 = F_{t_\nu}(T_1)$$

$$U_2 = F_{t_\nu}(T_2)$$

Així doncs, obtenim $U_1, U_2 \sim C_{\rho, \nu}^t$.

Algoritme simulació Còpules Arximedians

Partim de la base que les còpules Arximedians tenen la següent forma $C(u_1, \dots, u_d) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d))$. A partir de l'aplicació de l'algoritme, es vol generar $\mathbf{U} = (U_1, \dots, U_d)$ tal que $U_i \sim \text{Uniforme}(0, 1)$ i la dependència conjunta tingui la mateixa estructura que C .

Fer una inversió directa de C és molt complicat, per tant, es vol buscar una representació estocàstica més simple que, en la teoria, generi la mateixa C .

Comencem donant una sèrie de resultats els quals necessitarem més endavant (veure [21] per més informació respecte els següents resultats).

Teorema 4.4.1. (Teorema de Bernstein en funcions monòtones). Una funció $f : (0, \infty) \rightarrow \mathbb{R}$ és completament monòtona si i només si pot ser expressada com una transformada de Laplace d'una mesura de Borel positiva μ a $[0, \infty)$, és a dir,

$$f(x) = \int_0^\infty \exp^{-xt} \mu(dt),$$

on μ satisfà la condició d'integrabilitat $\int_0^\infty (1 \wedge t) \mu(dt) < \infty$.

Propietat exponencial. Recordem que una exponencial amb paràmetre λ compleix la següent propietat,

$$Y \sim \text{Exp}(\lambda) \Rightarrow P(Y > y) = e^{-\lambda y}, y \geq 0.$$

Si apliquem el Teorema anterior al cas de les copules, sabent que ψ és completament monòtona tenim que

$$\psi(t) = E(e^{-tV}) = \int_0^\infty e^{-tv} dF_V(v),$$

per a alguna variable aleatòria positiva V . Així doncs, existeix una $V > 0$ tal que ψ és la **transformada de Laplace** de V .

Així doncs l'algoritme consisteix en

1. Generar una variable latent V tal que ψ és la transformada de Laplace.
2. Construir variables $Y_1, \dots, Y_d$ tals que la seva distribució tingui forma $\psi(y_1, \dots, y_d)$.
 - (a) Si prenem $Y_i|V = v \sim \text{Exp}(v)$, tal que les exponencials siguin independents (en implementació $E_i \sim \text{Exp}(1)$ i $Y_i = E_i/V$). Condicionant $V = v$ obtenim per la propietat de l'exponencial prèviament esmentada

$$P(Y_1 > y_1, \dots, Y_d > y_d | V = v) = \prod_{i=1}^d e^{-vy_i} = e^{-v(y_1 + \dots + y_d)}$$

Aquesta forma encaixa amb la transformada de Laplace.

- (b) Passem a calcular la distribució conjunta de les variable Y_i . Si traiem la condició de condicionalitat, fent la mitjana de V , obtenim

$$P(Y_1 > y_1, \dots, Y_d > y_d) = E[e^{-V(y_1 + \dots + y_d)}].$$

Prenent $t = y_1 + \dots + y_d$ i per la definició de la transformada de Laplace, tenim

$$E[e^{-V(y_1 + \dots + y_d)}] = E[e^{-tV}] = \psi(t).$$

- (c) Passem a transformar les variable $Y_i \in [0, \infty)$ a uniformes $U_i \in (0, 1)$. Definim la transformació $U_i = \psi(Y_i)$. Com que ψ és decreixent,

$$U_i \leq u_i \Leftrightarrow \psi(Y_i) \leq u_i \Leftrightarrow Y_i \geq \psi^{-1}(u_i).$$

Per tant,

$$P(U_1 \leq u_1, \dots, U_d \leq u_d) = P(Y_1 \geq \psi^{-1}(u_1), \dots, Y_d \geq \psi^{-1}(u_d)).$$

Aleshores,

$$P(U_1 \leq u_1, \dots, U_d \leq u_d) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)).$$

On hem aplicat el resultat anterior $P(Y_1 > y_1, \dots, Y_d > y_d) = P(Y_1 \geq y_1, \dots, Y_d \geq y_d) = \psi(y_1 + \dots + y_d)$ (la segona igualtat es dona perquè la distribució és contínua) i la transformació $y_i = \psi^{-1}(u_i)$.

Aquest últim resultat és justament la definició d'una Còpula Arximediana.

Nota: En la transformació $U_i = \psi(Y_i)$, tenim que ψ és decreixent, contínua i $\psi(0) = 1, \psi(\infty) = 0$, per definició del generador d'una Còpula Arximediana. Per tant, ψ envia $[0, \infty)$ a $(0, 1]$ i U_i són uniformes, tal i com buscàvem.

S'aplica el mateix algoritme per a la generació de qualsevol Còpula Arximediana estudiada en aquest treball (amb ψ els seus respectius generadors). Concretament

- **Clayton:** $V \sim \text{Gamma}(1/\theta, 1)$.
- **Gumbel:** $V = \frac{\sin(\alpha S)}{\sin S^{1/\alpha}} \left(\frac{\sin(1-\alpha)S}{W} \right)^{(1-\alpha)/\alpha}$, on $S \sim U(0, \pi)$ i $W \sim \text{Exp}(1)$.
- **Frank:** $P(V = k) = \frac{1}{k} \frac{(1-e^{-\theta})^k}{\theta}$, $k = 1, 2, 3, \dots$

4.4.2 Comparació de models de còpula

La següent taula resumeix les propietats més rellevants de les famílies de còpules utilitzades en aquest estudi, destacant la seva capacitat per capturar dependència global o extrema, i la seva utilitat en contextos de fallida empresarial.

Família de còpules	Propietats principals	Tipus de dependència que capten	Dependència de cua
Gaussiana	Simètrica, modela dependència global però no capta asimetries. Estimació senzilla a partir de la matriu de correlació.	Dependència moderada en situacions normals.	No presenta dependència ni superior ni inferior.
t-Student	Simètrica, amb cues pesades. El paràmetre de graus de llibertat controla la intensitat dels extrems.	Adequada per capturar moviments conjunts en crisi.	Sí, dependència de cua superior i inferior.
Clayton	Asimètrica, molt sensible a moviments conjunts en valors baixos.	Captura dependència forta en situacions de mínims simultanis.	Sí, dependència de cua inferior.
Gumbel	Asimètrica, orientada a valors elevats.	Captura dependència forta en situacions de màxims simultanis.	Sí, dependència de cua superior.
Frank	Simètrica, flexible per dependència moderada però sense sensibilitat extrema.	Dependència global sense biaix cap a extrems.	No presenta dependència de cua.

4.5 Estimació de mesures de dependència

L'estudi de la dependència entre les ràtios financeres és una de les parts fonamentals d'aquest treball, ja que permet analitzar com aquestes es mouen conjuntament, en particular durant un període d'estrès com la crisi del COVID-19. En aquesta secció es descriuen les mesures de dependència utilitzades i la seva rellevància en el context d'insolvència empresarial.

4.5.1 Kendall's τ i Spearman's ρ

KENDALL'S τ

La τ de Kendall mesura el grau de concordança entre parelles d'observacions. Concretament

- $\tau > 0$ indica que la probabilitat de concordança és major, i per tant, que R_i tendeix a variar en el mateix sentit que R_j , per a $1 \leq i, j \leq 9$ amb $i \neq j$.
- $\tau = 0$ absència de dependència monòtona.
- $\tau < 0$ indica que la probabilitat de disconcordança és major, i per tant, que quan R_i augmenta, R_j tendeix a disminuir, per a $1 \leq i, j \leq 9$ amb $i \neq j$.

En contextos de fallida, una τ elevada i positiva pot indicar que dues ràtios tendeixen a empitjorar simultàniament, fet rellevant per anticipar un deteriorament financer conjunt.

SPEARMAN'S ρ

Valors positius del signe de la ρ de Spearman indiquen concordança en els rangs, mentre que valors negatius reflecteixen disconcordança. Tot i això, la ρ de Spearman no admet una interpretació probabilística directa com la τ de Kendall.

4.5.2 Dependència de cua

A diferència de la τ de Kendall's i la ρ de Spearman, la dependència de cua descriu la probabilitat que dues variables prenguin valors extrems simultàniament.

Per una banda, la cua superior, permet identificar si dues variables augmenten conjuntament. D'altra banda, la cua inferior, permet identificar si dues variables cauen simultàniament. En el cas d'aquest treball, on l'objectiu és l'estudi de situacions de disminucions extremes, s'espera que la **cua inferior** sigui la més rellevant.

4.6 Estimació de mesures de risc: VaR i ES**4.6.1 Transformació prèvia de ràtios per ús en VaR/ES**

Per a les ràtios R1 - R9, el valor teòric òptim es defineix en funció de si un increment de la ràtio presenta una millora o un empitjorament de la situació financera. Concretament, degut a que estem realitzant un estudi sobre el risc d'insolvència, ens interessa estudiar aquells valors que reflecteixen un empitjorament de la situació financera. Tal i com es descriu a [14]:

- $R1 < 0,1 - 0,15$. Senyal que el volum d'actius realment líquids és molt baix respecte al total d'actius. Això indica tensió de tresoreria i incapacitat potencial de convertir actius en efectiu quan es necessita. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.
- $R2 < 0,8$. Indica que els actius corrents no cobreixen el passiu exigible a curt termini que pot provocar insolvència imminent de flux de caixa. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.
- $R3 < 0,25$. Mostra una forta dependència del deute. En el cas en que la ràtio sigui negativa (< 0), seria un indicatiu de que el patrimoni net és negatiu i per tant l'empresa es troba en situació d'insolvència patrimonial. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.
- $R4 < 0$. Si l'empresa paga als proveïdors abans de cobrar dels clients, pot quedar-se sense tresoreria operativa, una de les causes més habituals de concurs. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.
- $R5 < 0\%$. ROA negatiu indica que el negoci no és capaç de crear valor i probablement com a resultat ha de consumir caixa. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.
- $R6 < 0\% - 10\%$. Marges baixos dificulten cobrir costos fixos, condueixen a pèrdues i deterioren el patrimoni net. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.
- $R7 < 0\%$. Un ROE negatiu sostingut condueix a un patrimoni net negatiu, indicador d'insolvència en molts casos. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.
- $R8 < 0\%$. Marge net negatiu indica que l'empresa no genera prou benefici per sostenir-se. Per tant, una disminució de la ràtio és un indicatiu d'empitjorament.

- $R9 > 5$. Massa deute fa que petites caigudes en beneficis amplifiquin pèrdues i augmentin molt la probabilitat d'insolvència. En el cas en que el patrimoni net sigui proper a 0, el Palanquejament tendeix a infinit, indicador d'insolvència patrimonial. Per tant, un increment de la ràtio és un índex d'empitjorament.

Al treballar amb còpules, l'objectiu és estudiar com dues variables es mouen conjuntament. Així doncs, per tal d'analitzar la dependència entre ràtios financeres mitjançant còpules, és necessari unificar la direcció econòmica de totes les variables. Algunes ràtios empitjoren quan disminueixen (R1, R2, R3, R4, R5, R6, R7, R8), mentre que d'altres empitjoren quan augmenten (R9). Per tal de fer les variables comparables, la ràtio R9 (Palanquejament) on un valor més alt representa una situació més desfavorable s'ha transformat multiplicant els valors per -1 . D'aquesta manera, totes les ràtios empitjoren quan disminueixen, i les mesures de dependència (Kendall's τ , coeficients de cua,...) tenen una interpretació econòmicament coherent. Si no s'hagués efectuat aquesta transformació aleshores

- Kendall's τ pot sortir negatiu només perquè les ràtios tenen orientacions oposades.
- Les còpules que capturen cues inferiors/superiors no tindran interpretació econòmica clara i per tant, la selecció òptima de còpula podria ser incorrecta.
- Comparar dependències entre períodes quedaria distorsionat.

Les ràtios més interessants per estudiar amb còpules són aquells que representen dimensions diferents del risc financer però que, en situacions de crisi, s'espera que tendeixen a deteriorar-se simultàniament. En concret, ara s'exposaran les parelles de ràtios seleccionades que es creu que aportaran més valor afegit al treball.

- **ROA (R5) - Marge Brut (R6)**. El fet que el ROA es pugui expressar com a producte del marge i la rotació no implica correlació lineal entre aquestes ràtios. Es tracta d'una relació multiplicativa, no lineal i influenciada per múltiples factors operatius, de manera que marge i rotació poden evolucionar de manera independent o mostrar dependència només en situacions extremes. L'ús de còpules permet analitzar si, durant períodes de crisi, ambdues ràtios convergeixen simultàniament cap a la cua inferior.
- **ROE (R7) - Palanquejament (R9)**. Així com en el cas anterior, aquesta parella de ràtios també presenten una relació multiplicativa. El palanquejament actua com a mecanisme amplificador dels resultats, tant positius com negatius. Així doncs, en contextos adversos, increments del palanquejament poden intensificar les caigudes del ROE. L'aplicació de còpules permet examinar si les caigudes extremes del ROE s'intensifiquen quan hi ha nivells elevats d'endeutament.
- **PMP - PMC (R4) - (R1)**. El que es pretén estudiar a partir d'aquesta parella és la transició d'una tensió de caixa puntual a un risc de liquiditat estructural. La ràtio PMP-PMC capta desequilibris de caixa, i la seva comparació amb la ràtio de liquiditat (R1) permet analitzar si una tensió operativa puntual pot evolucionar cap a un risc de solvència a curt termini.
- **ROA (R5) - ROE (R7)**. En situacions d'insolvència, una caiguda del ROA tendeix a afectar al ROE, especialment en empreses endeutades. Les còpules ens permeten identificar com de lligat està el deteriorament de la rendibilitat operativa amb el de la rendibilitat per als accionistes.
- **Marge Net (R8) - Palanquejament (R9)**. Quan el Marge Net es redueix de manera sostinguda, és un indicatiu de que l'empresa presenta poca capacitat de generar benefici.

Si aquesta situació es combina amb un nivell elevat de palanquejament, qualsevol reducció addicional en els beneficis pot accelerar la situació d'insolvència. Les còpules permeten analitzar si, en situacions extremes, el deteriorament del marge net coincideix amb nivells elevats d'endeutament.

- **ROA (R5) - (R2)**. El deteriorament de la rendibilitat sovint va acompanyat de tensions de liquiditat, ja sigui per una reducció dels marges, de l'actiu corrent... L'ús de còpules permet analitzar si la dependència entre baixa rendibilitat i problemes de liquiditat s'intensifica en períodes adversos.
- **ROA (R5) - (R3)**. Aquesta parella permet estudiar com el deteriorament operatiu es transmet al patrimoni net. Rendibilitats baixes o negatives deterioren el patrimoni i empitjoren la ràtio de solvència patrimonial. Donat que aquesta relació no és lineal, les còpules són una bona eina per capturar possibles dependències extremes.

Les darreres combinacions permeten capturar dependència no lineal i dependència de cua inferior. A més a més, també proporcionen informació rellevant sobre la propagació del risc entre tensió operativa, deteriorament de caixa i insolvència patrimonial.

4.6.2 Càlcul del VaR i ES empírics

Les mesures de risc **Value at Risk (VaR)** i **Expected Shortfall (ES)** ens permeten quantificar els comportaments més desfavorables de les ràtios estudiades.

Per aplicar correctament aquestes mesures, cal identificar quin costat de la distribució reflecteix deteriorament. Tal com s'ha explicat prèviament, les ràtios R1–R8 empitjoren com més baix és el seu valor. En el cas de R9, després del canvi de signe aplicat, segueix el mateix patró. Per tant, en totes les ràtios el deteriorament s'associa a valors més petits, i per tant és necessari analitzar la **cua inferior** de la distribució.

El VaR ens permet determinar el llindar a partir del qual es troba l' $\alpha\%$ de les empreses amb els pitjors valors de la ràtio. Així, el VaR_α indica que aproximadament un $\alpha\%$ de les empreses presenten una ràtio inferior a aquest llindar. En aquest treball s'analitzarà el VaR per a $\alpha = 0.05$ i $\alpha = 0.1$ (és a dir, el pitjor 5% i 10% de les empreses).

L'ES quantifica la severitat mitjana dels pitjors casos. En concret, ES_α representa el valor mitjà de la ràtio quan aquesta es troba dins de l' $\alpha\%$ d'empreses amb pitjors resultats. De nou, s'avaluarà per a $\alpha = 0.05$ i $\alpha = 0.1$.

Així, un VaR més extrem (més negatiu o més baix) indica que els pitjors valors observats de la ràtio són més desfavorables en aquell període. Per altra banda, un ES més sever indica que, un cop l'empresa entra en una situació "problemàtica", el grau de deteriorament és més elevat.

4.7 Implementació computacional

4.7.1 Estructura del codi

Per implementar els mètodes descrits en seccions anteriors, s'ha utilitzat el programari **RStudio**, un entorn gratuït orientat a l'anàlisi estadística. A l'apartat de resultats no es mostrarà el codi complet; aquest es troba adjunt a l'Apèndix.

Prèviament a la presentació de resultats, cal introduir les llibreries i funcions utilitzades que incorpora **RStudio**. Concretament, s'han utilitzat les llibreries següents:

- **tidyr**. Permet organitzar i netejar dades desestructurades. S’ha utilitzat, entre d’altres, per eliminar files amb valors nuls mitjançant la funció *drop_na*.
- **deplyr**. Proporciona una nova sintaxi molt útil per a la manipulació i operacions amb *data frames*. Alguns exemples inclosos en el paquet que s’han utilitzat són: *select* (retorna un conjunt de columnes), *filter* (retorna un conjunt de files segons uno o varies condicions lògiques), *mutate* (afegeix noves variables/columnes o transforma variables existents)...
- **VineCopula**. Inclou eines per a l’anàlisi exploratòria de dependències bivariades, la selecció i estimació de còpules, i la construcció de models *vine*. Tot i que en aquest treball només s’utilitza per a còpules bivariants, el paquet és molt potent per modelar dependències complexes en altes dimensions.
- **readxl i writexl**. Permeten importar dades des d’Excel i exportar-hi els resultats de manera directa i senzilla.

A més d’aquestes llibreries, s’han definit i utilitzat diverses funcions específiques

- **VaR**. S’ha definit una funció auxiliar que calcula el VaR mitjançant la funció interna *quantile*, que retorna el valor associat al quantil α de la distribució empírica.
Nota: S’utilitza el mètode d’interpolació *type = 7*, adequat per a mostres que segueixen una distribució contínua.
- **ES**. Funció auxiliar que calcula l’ES de manera que retorna la mitjana de tots els rendiments que estan per sota del valor del VaR (aplicant la definició de l’ES vista al Lema 2.2.5).
- **prob_tail_cond**. Funció auxiliar que calcula la probabilitat condicionada

$$P(U_1 < \alpha | U_2 < \alpha) = \frac{P(U_1 < \alpha \cap U_2 < \alpha)}{P(U_2 < \alpha)} = \frac{C(\alpha, \alpha)}{\alpha}.$$

L’última igualtat surt de la definició de còpula. Aquesta probabilitat mesura fins a quin punt, quan una ràtio es troba en els seus pitjors $\alpha\%$ valors, un altra ràtio també tendeix a situar-se en la mateixa zona. Serveix com a indicador de moviments conjunts en situacions adverses i ajuda a identificar possibles ràtios “transmissores de risc”.

- **rank**. Funció interna que retorna els rangs dels valors d’un vector.
- **BiCopSelect**. Selecciona de manera automàtica una família de còpules bivariants adequada per a les dades donades. Les estimacions dels paràmetres s’obtenen mitjançant *MLE* i el criteri d’ajust és per defecte l’*AIC*. Algunes de les famílies de còpules disponible són

- 0 = Còpula independent
- 1 = Copula Gaussiana
- 2 = Còpula t-Student
- 3 = Còpula de Clayton
- 4 = Còpula de Gumbel
- 5 = Còpula de Frank

- **BiCopPar2Tau**. Calcula el valor teòric de la τ de Kendall d’una còpula bivariant per a certs valors de paràmetres (aquells calculats mitjançant *MLE*).

- **BiCopPar2TailDep**. Calcula els coeficients teòrics de dependència de la cua d'una còpula bivariant per a certs valors de paràmetres (aquells calculats mitjançant *MLE*).

Més informació sobre les llibreries i funcions figura a [23].

A continuació s'exposarà de manera general el **flux** que segueixen els diferents codis. Tot i que l'estructura no és idèntica, sí que comparteixen una lògica comuna i per això s'ha optat per presentar un únic flux.

1. Importació de la base de dades en Excel.
2. Definició de paràmetres: ràtios utilitzades, períodes estudiats, valors d' α ...
3. Definició de funcions auxiliars descrites anteriorment.
4. Aplicació de filtres per tal de seleccionar les columnes desitjades i eliminar aquells valors classificats com a "NaN".
5. Transformació de les ràtios en variables uniformes mitjançant rangs.
6. Selecció i estimació de la còpula que millor s'ajusta a cada parella de ràtios.
7. Càlcul de les mesures de dependència com la τ de Kendall (empírica i teòrica, de manera que es pugui comprovar si l'ajust de la còpula és rigorós), ρ de Spearman, dependència de cua inferior i superior.
8. Càlcul de VaR, ES i probabilitat condicionada de la cua.
9. Exportació de la taula generada amb els resultats finals a un arxiu Excel.

Capítol 5

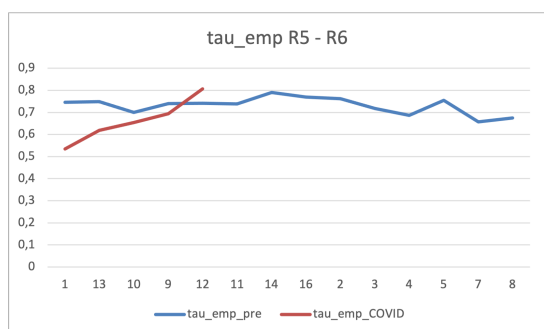
Resultats

Nota: A l'Apèndix figura una llegenda amb el número de “bucket” que correspon a cada Comunitat Autònoma.

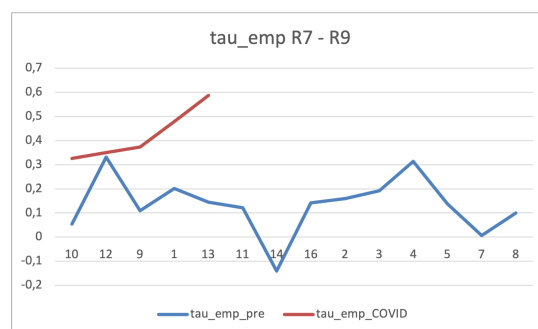
Abans d'analitzar els resultats, cal destacar que cap de les ràtios disposa de més de 30 observacions en el període post - COVID. Així doncs, els resultats corresponents a aquest període no es presenten ni s'inclouen en les comparacions amb els altres períodes, donada la falta de suficient robustesa estadística per extreure'n conclusions.

5.1 Dependència empírica entre ratís financers

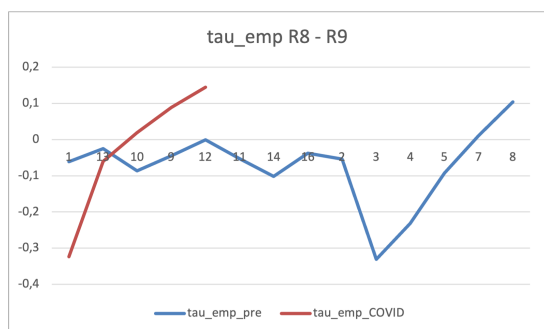
Tot i que s'ha calculat la τ Kendall empírica per a totes les parelles de ràtios, en aquesta secció només es presentaran aquelles relacions que aporten més cobertura en quant a rendibilitat operativa, rendibilitat financera, estructura financera i circulat i liquiditat. La resta de les mesures calculades figuren a l'Apèndix.



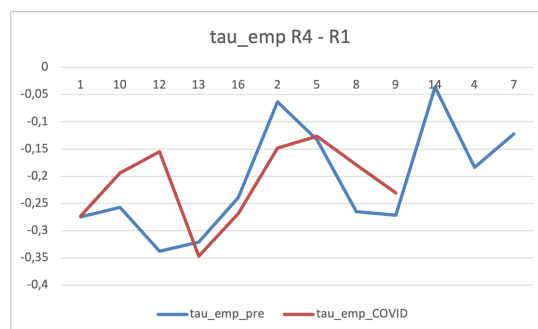
(a) τ de Kendall empírica parella R5 - R6



(b) τ de Kendall empírica parella R7 - R9



(c) τ de Kendall empírica parella R8 - R9



(d) τ de Kendall empírica parella R4 - R1

En una primera observació, s'aprecia que durant el període del COVID la τ de Kendall empírica

no segueix un patró uniforme d'increment o reducció de la dependència entre els diferents ratis analitzats. Aquest comportament es podria explicar principalment pel caràcter heterogeni de l'impacte de la crisi i per la menor disponibilitat d'observacions durant aquest període.

D'una banda, la pandèmia no va afectar de manera homogènia ni en el temps ni entre empreses, ni tampoc en totes les dimensions financeres. Així doncs, mentre algunes empreses van patir principalment tensions operatives, d'altres van experimentar dificultats financeres o de liquiditat, i amb intensitats i en temps diferents.

D'altra banda, la reducció del nombre de "buckets" amb informació durant el període COVID incrementa la volatilitat en l'estimació de la τ de Kendall, fent que petites variacions en les dades tinguin un impacte rellevant sobre els valors estimats.

Tot i això, s'observa un patró general. En el període pre-COVID les estimacions de dependència són més estables, mentre que durant la crisi esdevenen més disperses. Aquest fet suggereix que la crisi tendeix a trencar les relacions financeres donant lloc a comportaments heterogenis segons característiques com la mida, el territori o la situació prèvia de l'empresa.

Concretament, observem com la relació de dependència entre R4 i R1 és similar i negativa en els dos períodes, fet que implica que les tensions operatives no es tradueixen immediatament en problemes de liquiditat, probablement a causa de mesures extraordinàries com ajudes públiques o moratòries.

En quant a les relacions entre rendibilitat i estructura financera (R8 – R9 i R7 – R9), s'observa com aquestes no mostren una dependència lineal estable durant la crisi, sinó que mostren "buckets" que alternen dependències fortes, febles o negatives, reflectint el paper amplificador del palanquejament tant de pèrdues com de beneficis. En canvi, relacions de caràcter més estructural i operatiu, com la de R5 i R6, mantenen una dependència elevada i relativament estable en ambdós períodes.

Nota: També s'ha calculat la ρ de Spearman però degut a que els resultats obtinguts son redundants amb el càlcul de la τ de Kendall, les gràfiques només figuraran a l'Apèndix.

5.2 Ajust de còpules: selecció del model de dependència

En aquesta secció es proporcionarà una taula amb les diferents famílies de còpules que millor s'ajusten a les dades d'acord amb la selecció feta pel programari a partir del criteri AIC. A l'Apèndix queden recollides les gràfiques que mostren els ajusts de còpules diferenciats per famílies.

Parelles de ràtios	Família seleccionada pre - COVID	Família seleccionada COVID
R5- R6	Còpula de Frank	Còpula de Gumbell
R7 - R9	Còpula t-Student ($\nu \approx 2$)	Còpula de Gumbell
R5 - R7	Còpula t-Student ($\nu = 2$)	Còpula t-Student ($\nu = 2$)
R4 - R1	Còpula de Frank	Còpula t-Student ($\nu \approx 4$)
R5 - R3	Còpula de Clayton	Còpula de Clayton/t-Student ($\nu \approx 3$)
R8 - R9	Còpula Independent	Còpula Independent
R5 - R2	Còpula de Clayton	Còpula Independent

Nota: En el cas de la còpula t-Student, els graus de llibertat s'observen a través del segon paràmetre donat pel model. Valors baixos dels graus de llibertat ($\nu \approx 2-3$) indiquen la presència de cues pesades i una elevada probabilitat d'esdeveniments extrems conjunts.

Anàlisi R5 - R6

En el període pre-COVID, la dependència entre ROA i marge brut es descriu principalment amb una còpula de Frank, indicant una relació simètrica i homogènia, sense concentració a les cues, coherent amb un entorn econòmic estable. Durant el COVID, predomina la còpula de Gumbell, fet que mostra una dependència més forta a la cua superior i que indica que les empreses que mantenen marges bruts elevats durant la crisi també presenten un ROA elevat. Aquesta dependència no implica que la major part de les empreses presentin valors elevats de ROA i marge brut. El que realment implica és que, condicionat a valors extrems positius d'un indicador, la probabilitat que l'altre també ho sigui augmenta.

Anàlisi R7 - R9

Abans del COVID, la relació entre R7 i R9 mostra una dependència simètrica amb presència d'extrems conjunts, capturada per la còpula t-Student. Això reflecteix el paper amplificador del palanquejament, que intensifica tant guanys com pèrdues. Durant el COVID, la relació és capturada per la còpula de Gumbell, indicant una dependència asimètrica concentrada a la cua superior, el que suggereix que el palanquejament passa a tenir un paper amplificador només en el cas que hi hagi guanys.

Anàlisi R5 - R7

En el període pre-COVID, la dependència és simètrica i relativament intensa, amb presència d'esdeveniments extrems conjunts, modelada majoritàriament per la còpula t-Student, coherent amb el fet que el ROE està relacionat amb la rendibilitat operativa. Durant el COVID, aquesta dependència esdevé més heterogènia, amb major predomini de la t-Student però amb aparicions de còpules com la Independent o Gumbell, reflectint més varietat en el comportament empresarial com a conseqüència de la diferent gestió davant de la crisi.

Anàlisi R4 - R1

Abans del COVID, la relació entre R4 (PMP-PMC) i R1 es descriu mitjançant una còpula de Frank, indicant una dependència simètrica i sense concentració a les cues. El signe negatiu de la τ de Kendall suggereix una relació inversa, coherent amb el fet que si no es disposa de liquiditat el pagament a proveïdors s'enderrareix i R4 esdevé més elevat. Durant la crisi, el predomini de la còpula t-Student indica la presència d'extrems conjunts, reflectint un context on la situació i gestió de cada empresa és única.

Anàlisi R5 - R3

En el període pre-COVID, la dependència és asimètrica i predomina l'ajust per còpula de Clayton, reforçant la idea que la pèrdua de capacitat de generar beneficis impacta directament en la capacitat de fer front a les obligacions a curt termini. Durant el COVID, la dependència es manté però esdevé més diversa, amb presència similar de Clayton i t-Student, suggerint comportaments més extrems i menys asimètrics, amb empreses que van experimentar deterioraments conjunts i d'altres que van mostrar una major capacitat de resistència.

Anàlisi R8 - R9

Ni abans ni durant el COVID s'observa una dependència estadísticament significativa entre R8 i R9, és a dir, els extrems d'un indicador no aporten informació sobre l'altre, fins i tot en un context de crisi. Econòmicament, això suggereix que el marge net, que sol estar influït per

factors extraordinaris, no manté una relació sistemàtica amb l'estructura de finançament.

Anàlisi R5 - R2

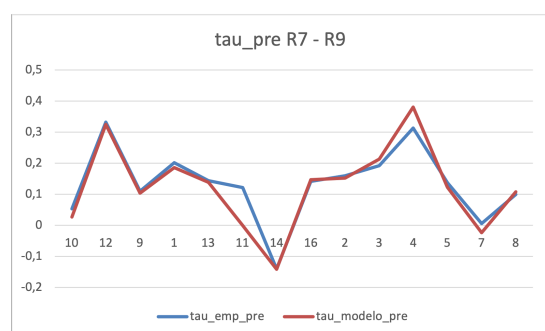
Abans del COVID, s'identifica una dependència feble però persistent i asimètrica, principalment a la cua inferior, on una caiguda del ROA s'associa a una pèrdua de liquiditat. Durant el COVID, aquesta relació pràcticament desapareix, amb predomini de la còpula independent i valors de la tau propers a zero, coherent amb la distorsió causada per les ajudes públiques sobre la relació habitual entre rendibilitat i liquiditat.

Finalment, cal destacar que els resultats obtinguts mitjançant l'anàlisi de còpules són coherents a nivell qualitatiu amb els patrons observats a partir de la τ de Kendall empírica. En particular, les parelles de ratis amb dependència empírica més estable tendeixen a mostrar estructures de dependència més definides, mentre que aquelles amb τ feble o altament dispersa presenten una selecció de còpules més heterogènia o propera a la independència.

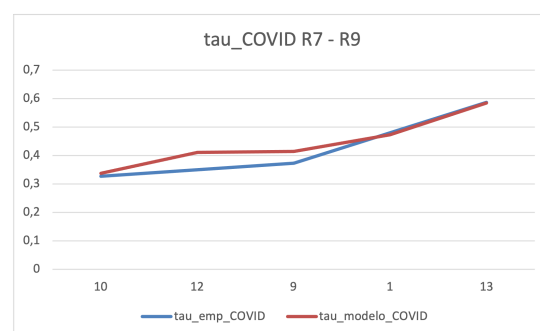
5.3 Validació del models: dependència observada vs modelada

En aquesta secció es compara la dependència observada en les dades, mesurada mitjançant la τ de Kendall empírica, amb la dependència que capturen els models de còpula estimats. L'objectiu principal no és reproduir exactament la dependència observada en cada "bucket", sinó comprovar si el model és capaç de captar adequadament el comportament general de la dependència entre les ràtios analitzades.

De manera general, s'observa que, per a la majoria de parelles de ràtios, la τ de Kendall modelada presenta un comportament molt similar al de la τ empírica, tant en el període previ al COVID com durant la crisi. Les diferències observades són, en general petites i no segueixen cap patró, fet que suggereix que els models de còpula seleccionats reproduïen l'estructura de dependència que presenten les dades.

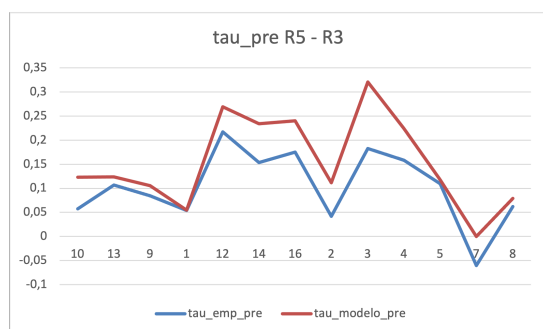


(a) Comparació pre - COVID de la τ del model amb l'empírica de la parella R7 - R9

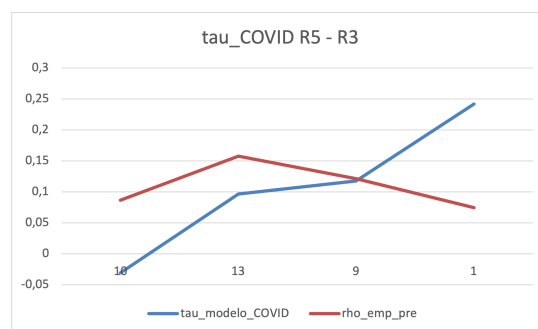


(b) Comparació COVID de la τ del model amb l'empírica de la parella R7 - R9

En les darreres dos figures es mostra un exemple representatiu del cas genèric que s'ha comentat prèviament.



(a) Comparació pre - COVID de la τ del model amb l'empírica de la parella R5 - R3



(b) Comparació COVID de la τ del model amb l'empírica de la parella R5 - R3

Per contra, en aquestes figures es mostra el cas de la parella R5 - R3, on s'observa un comportament diferent (sobretot en el període de COVID). Aquest comportament es pot donar perquè en un context de crisi, trobem empreses que presenten deterioraments conjunts i d'altres que mostren una major capacitat de resistència, fet que pot no quedar completament reflectit per la τ de Kendall.

Nota: Les demès gràfiques queden recollides a l'Apèndix donat que mostren un comportament similar a l'exemple de la parella R7 - R9.

5.4 Mesures de risc: VaR i Expected Shortfall

Si ens fixem en les gràfiques del VaR al 5%, s'observa que durant el període COVID el risc estimat tendeix a ser més elevat i més homogeni entre "buckets", mentre que en el període pre-COVID el VaR presenta una major dispersió i comportaments més irregulars, especialment en ràtios com R1, R2 o R3. Aquest patró és coherent amb el context econòmic on abans de la crisi les diferències entre empreses (mida, sector, estructura...) es reflecteixen clarament en el risc extrem, mentre que durant el COVID el xoc és més generalitzat i redueix la variabilitat entre empreses.

En les ràtios de liquiditat i estructura financera (R1, R2 i R4), el VaR durant el COVID mostra una tendència creixent però relativament estable, fet que pot interpretar-se com l'efecte combinat de la caiguda de l'activitat i de les mesures extraordinàries adoptades, que moderen els valors més extrems sense eliminar el risc. En canvi, en els altres ràtios (R3, R5, R6, R7 i R8) el VaR durant el COVID es situa en valors més negatius, indicant un augment del risc de pèrdues severes.

Pel que fa al palanquejament (R9), l'evolució del VaR entre els dos períodes podria indicar que el risc associat a l'endeutament esdevé més sistemàtic i menys dependent de casos puntuals.

Al estudiar l'ES ens centrem en la severitat mitjana de les pèrdues un cop s'entra a la cua. En termes generals, durant el COVID l'ES al 5% tendeix a ser més negatiu en la majoria de ràtios, indicant que els escenaris adversos són més severs que en el període pre-COVID.

Una excepció rellevant és el palanquejament (R9). Aquest resultat insinua que els escenaris extrems associats al palanquejament eren de mitjana, més severs abans de la crisi. Durant el COVID, aquestes situacions extremes semblen més contingudes, possiblement com a conseqüència de polítiques de suport i refinançament.

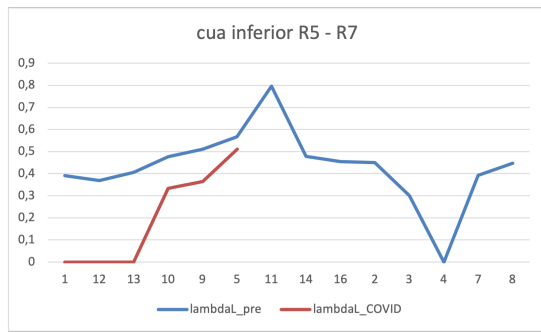
En contrast, les ràtios de rendibilitat (R5, R7 i R8) sí que mostren una intensificació de les pèrdues mitjanes dins la cua en la majoria dels "buckets", indicant que el xoc s'ha transmès principalment a través dels resultats operatius i nets més que no pas mitjançant un deteriorament

generalitzat de l'endeutament extrem.

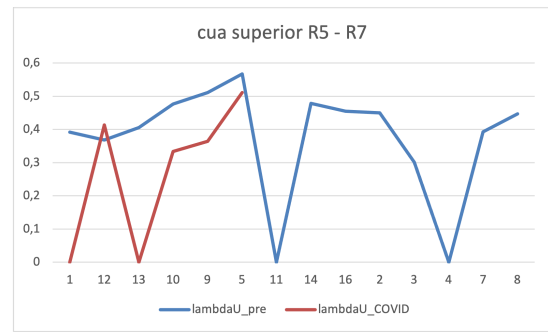
Nota: Degut a l'extensió de la mostra de gràfics, aquests queden recollits en l'Apèndix tant els valors del VaR per $\alpha \in \{5\%, 10\%\}$ com els de l'ES per $\alpha \in \{5\%, 10\%\}$. Al llarg de la secció només es discuteixen els resultats per $\alpha = 5\%$, ja que els obtinguts per $\alpha = 10\%$ són qualitativament similars.

5.5 Dependència de cua i risc extrem conjunt

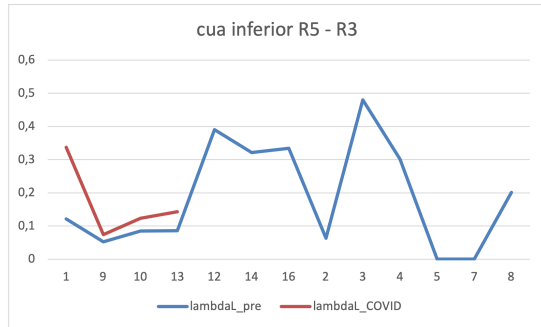
Aquest apartat analitza si els episodis extrems d'un rati tendeixen a produir-se simultàniament amb episodis extrems d'un altre, posant atenció en la dependència a les cues. Així doncs, només es rellevant estudiar el cas de les parelles R5 - R7 i R5 - R3, donat que en l'ajust de còpules són els únics que presenten dependència de cua tant al període pre - COVID com en el període COVID.



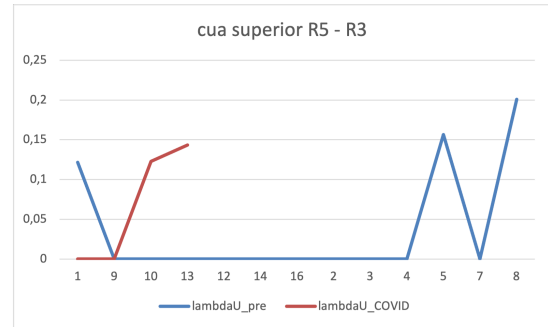
(a) λ_L R5 - R7



(b) λ_U R5 - R7



(c) λ_L R5 - R3



(d) λ_U R5 - R3

En la parella R5 - R7, la dependència de cua inferior i superior presenta valors molt similars, fet coherent amb l'ajust d'una còpula t-Student, que permet dependència simètrica en ambdues cues. No obstant això, durant el període COVID aquests valors són lleugerament inferiors als del període pre-COVID, indicant una reducció de la dependència extrema conjunta. Aquest comportament és consistent amb els resultats de l'ES, ja que, tot i que aproximadament la meitat dels "buckets" presenten valors d'ES més extrems durant el COVID, aquestes pèrdues severes no es produeixen de manera sincronitzada entre ambdues ràtios.

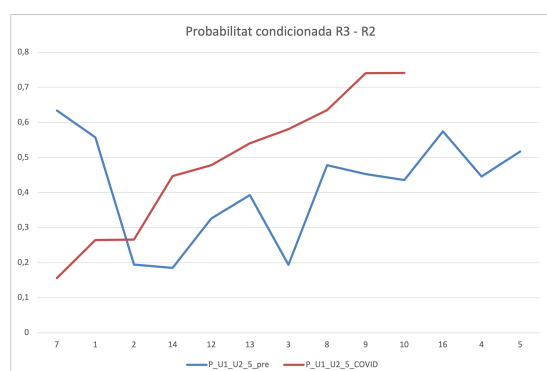
Pel que fa a la parella R5 - R3, s'observa un lleuger augment de la dependència tant a la cua inferior com a la superior durant el COVID, tot i que en ambdós casos els valors es mantenen molt baixos, indicant una dependència extrema feble. Aquest resultat suggereix que durant la crisi, els episodis extrems de rendibilitat i solvència tendeixen a coincidir més sovint que abans, però sense arribar a establir una relació forta. Aquesta observació és coherent amb el comportament de l'ES de la ràtio R3, que mostra valors clarament més extrems durant el COVID, reflectint un

empitjorament de les pèrdues severes en solvència, tot i que aquestes no es traslladin de manera sistemàtica a R5.

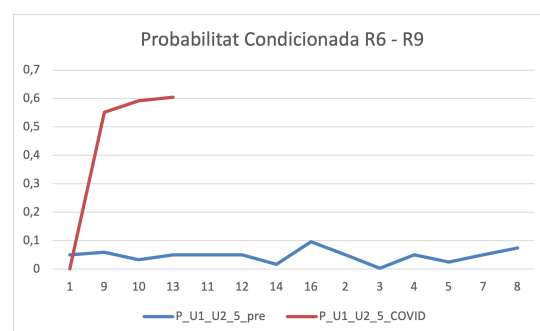
5.6 Identificació de ràtios tractores

Aquest apartat, a diferència dels anteriors, no es centra en l'estudi de l'ajust ni en l'estructura d'un model de dependència, sinó que té com a objectiu identificar si existeixen ràtios que, quan es troben en valors extrems, tendeixen a arrossega-ne d'altres.

Per analitzar aquestes relacions, s'ha estudiat la probabilitat condicionada $P(R_i < 5\% | R_j < 5\%)$ entre totes les ràtios.



(a) Probabilitat condicionada R3 - R2



(b) Probabilitat condicionada R6 - R9

En el cas de R3 – R2, durant el període COVID s'observa que quan la solvència (R3) entra en el pitjor 5% dels valors, la probabilitat de que la liquiditat (R2) també ho faci és clarament més elevada que en el període pre-COVID. Això indica que en un context de crisi, els problemes de solvència tendeixen a traslladar-se amb més facilitat a tensions de liquiditat.

Un altre cas a destacar es el de la parella R6 - R9. Mentre que abans del COVID la probabilitat condicionada és baixa i força estable, durant la crisi augmenta de manera molt significativa. Això indica que una caiguda extrema del marge brut (R6) acostuma a anar acompanyada d'un empitjorament del palanquejament (R9), indicant un trasllat del risc en l'àmbit financer cap a l'estructura financera.

Finalment, tot i que hi ha altres parelles que també mostren un paper rellevant, com R5 – R7, degut a que ja s'ha discutit en seccions prèvies la seva relació, en aquesta no es profunditzarà.

Aquests resultats indiquen que durant períodes de crisi, és especialment rellevant vigilar les ràtios R3 i R6, ja que actuen com a ràtios tractores, transmetent situacions de risc extrem cap a la liquiditat (R2) i el palanquejament (R9).

Capítol 6

Conclusions i Línies futures

L'objectiu principal d'aquest treball era analitzar el comportament del risc d'insolvència empresarial i la relació entre diferents ràtios financeres abans i durant la crisi provocada pel COVID-19, posant especial èmfasi en l'estructura de dependència, el risc extrem i els possibles mecanismes de transmissió del risc entre variables.

Una de les expectatives inicials era que, en un context de crisi sistèmica i davant d'un conjunt d'empreses en situació de risc d'insolvència, la dependència entre les ràtios financeres associades a aquesta situació augmentés de manera generalitzada. No obstant això, els resultats obtinguts mostren que aquesta hipòtesi no es compleix de forma sistemàtica. En molts casos, la dependència mesurada mitjançant la τ de Kendall no augmenta durant el COVID i, en alguns "buckets", fins i tot es redueix. Aquest resultat mostra que la crisi no enforteix necessàriament les relacions entre ràtios, sinó que tendeix a fragmentar-les i a incrementar l'heterogeneïtat en el comportament empresarial.

Aquest trencament de relacions mitjanes és coherent amb el fet que la crisi del COVID no va afectar a totes les empreses de la mateixa manera ni amb la mateixa intensitat. A més, la presència de factors externs com les ajudes públiques, les moratòries financeres o els ERTOS probablement va alterar la relació habitual entre alguns indicadors financers. Tot i que aquests factors no s'han modelitzat explícitament, els resultats obtinguts són compatibles amb un escenari en què aquests mecanismes actuen com a amortidors de valors extrems de les ràtios i per tant modifiquen les relacions esperades entre ells.

L'anàlisi de l'ajust de còpules reforça aquesta idea. En el període pre-COVID, predominen estructures de dependència més simples i simètriques, coherents amb un entorn econòmic relativament estable. Durant el COVID, en canvi, emergeixen estructures més heterogènies, amb presència de còpules amb dependència de cua en algunes parelles concretes. Aquest resultat no indica un augment generalitzat de la dependència, sinó un canvi en la seva naturalesa, amb una major rellevància dels comportaments extrems en determinades relacions estructurals.

Pel que fa a les mesures de risc extrem, els resultats del VaR i de l'Expected Shortfall mostren que la crisi del COVID no només incrementa el nivell de risc, sinó que també en modifica l'estructura. El VaR al 5% evidencia que, durant la crisi, el risc extrem tendeix a ser més homogeni entre empreses, reflectint l'impacte d'un xoc que afecta de manera semblant a les diferents empreses, mentre que en el període previ predominen comportaments més volàtils. L'ES permet reforçar aquest resultat, ja que mostra que la severitat mitjana de les pèrdues augmenta clarament en la majoria de les ràtios exceptuant el palanquejament, que presenta un comportament més heterogeni i, en alguns casos, menys sever que abans de la crisi.

Aquesta diferència entre risc extrem i dependència és un dels resultats centrals del treball. Tot i que el risc extrem augmenta durant el COVID, això no implica necessàriament un increment generalitzat de la dependència entre ràtios. En alguns casos, l'empitjorament de l'ES respon principalment a un deteriorament marginal més sever i no a una intensificació de la dependència

extrema conjunta.

L'anàlisi de probabilitats condicionades permet identificar possibles mecanismes de transmissió del risc extrem que no es capturen mitjançant mesures de dependència global. En particular, els resultats suggereixen que, durant el COVID, les ràtios associades a la solvència (R3) i al marge brut (R6) actuen com a ràtios tractores, ja que quan entren en situacions extremes tendeixen a arrossegar la liquiditat (R2) i el palanquejament (R9). Aquest resultat mostra que la transmissió del risc extrem no és simètrica ni generalitzada, sinó que es produeix a través de ràtios específiques.

Finalment, cal esmentar les limitacions del treball. Les principals han estat la reduïda disponibilitat de dades en alguns períodes, especialment durant el COVID i post-COVID, i la falta d'informació sobre factors externs. Aquests fets han limitat la possibilitat d'establir relacions causals fortes, però no han impedit l'obtenció de resultats robustos a nivell descriptiu i comparatiu, el que ha permès oferir una visió coherent del comportament del risc i de la dependència en un context de crisi sistèmica.

Com a possibles línies futures de recerca, seria interessant aprofundir en l'estudi de les ràtios que s'han trobat com a transmissores del risc incorporant informació addicional sobre factors externs com el sector d'activitat, la mida empresarial, les ajudes externes... Aquest enfocament permetria comprendre millor fins a quin punt les relacions capturades es deuen a característiques individuals de les empreses o al risc sistemàtic provocat pel context de crisi.

Com a possible extensió del treball, es podria realitzar l'anàlisi sobre empreses insolvents en períodes econòmics no marcats per una crisi sistemàtica com va ser la del COVID-19. L'estudi en períodes més estables, permetria que els resultats estiguessin més centrats en les relacions entre les ràtios financeres, reduint la influència de factors extraordinaris que poden alterar la dinàmica habitual entre aquests indicadors. A més, aquest anàlisi podria contrastar-se amb els resultats obtinguts en el període de crisi amb l'objectiu d'identificar quines relacions de dependència són característiques del risc d'insolvència i quines responen a situacions excepcionals.

Finalment, una última extensió seria estimar les distribucions marginals de les ràtios mitjançant models basats en la Teoria dels Valors Extremes, com la distribució de Pareto Generalitzada. Donat que el treball es centra en el comportament extrem de les variables, utilitzar aquest mètode podria proporcionar una representació més adequada de les cues i per tant podríem obtenir una millor modelització de les dades.

Bibliografía

- [1] Instituto Nacional de Estadística. (n. d.). *Estadística del procedimiento concursal. Datos trimestrales*.
<https://www.ine.es/jaxiT3/Tabla.htm?t=3182>, 2025.
- [2] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 7.2). Princeton University Press.
- [3] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 7.4). Princeton University Press.
- [4] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., p. 225). Princeton University Press.
- [5] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 6, p. 184, Lema 6.5). Princeton University Press.
- [6] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 7.2, p. 236). Princeton University Press.
- [7] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 7.2, p. 238). Princeton University Press.
- [8] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., apartat 7.2.3). Princeton University Press.
- [9] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 2.3). Princeton University Press.
- [10] Pessoa de Oliveira, A. K. (2015). *Análisis de solvencia de las empresas: Modelo estático versus modelo dinámico* (Tesi doctoral, Universidad de Zaragoza). Repositorio Zaguán.
https://zaguan.unizar.es/record/48319/files/TUZ_0863_pessoa_analisis.pdf, 2016.
- [11] Instituto Nacional de Estadística. (n. d.). *Insolvencia*.
<https://www.ine.es/DEFIne/concepto.htm?c=5172>, 2025.
- [12] Banco Santander. (n. d.). *Insolvencia: qué es y tipos*.
<https://www.santander.com/es/stories/insolvencia>, 2023.
- [13] Altman, E. I., Avery, R. B., Eisenbeis, R. A., & Stinke y, J. (1981). *Application of classification techniques in business, banking and finance* (Vol. 3). JAI Press.
- [14] Bernstein, L. A. (1989). *Financial statement analysis: Theory, application, and interpretation* (5a ed.). Richard D. Irwin.
- [15] Rojo Ramírez, A. A. (2011). *Análisis económico-financiero de la empresa: Un análisis desde los datos contables*. Garceta.
- [16] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 7.5). Princeton University Press.
- [17] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 5). Princeton University Press.
- [18] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. 7.5, p. 272). Princeton University Press.

- [19] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques, and tools* (2a ed., cap. A.3, p. 650). Princeton University Press.
- [20] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer pringer Science+Business Media.
- [21] Widder, D. V. (1946). *The Laplace transform*. Princeton University Press.
- [22] Di Clemente, A., & Romano, C. (2021). *Calibrating and simulating copula functions in financial applications*. *Frontiers in Applied Mathematics and Statistics*, 7, Article 642210. <https://www.frontiersin.org/journals/applied-mathematics-and-statistics/articles/10.3389/2021>.
- [23] RDocumentation. (n. d.). *R: Documentation for the R language*. <https://www.rdocumentation.org>
- [24] Xaus Pariente, A. (2016). *Insolvency risk: Characterisation and prediction* (Treball Final de Grau). Universitat de Barcelona, Facultat de Matemàtiques i Informàtica / Facultat d'Economia i Empresa.

Apèndix

Llegenda Comunitats Autònomes

Comunitat Autònoma	Número de “bucket”
Andalusia	1
Aragó	2
Astúries	3
Illes Balears	4
Canàries	5
Cantàbria	6
Castella i Lleó	7
Castella-la Manxa	8
Catalunya	9
Comunitat Valenciana	10
Extremadura	11
Galícia	12
Comunitat de Madrid	13
Múrcia	14
Navarra	15
País Basc	16
La Rioja	17
Ceuta	18
Melilla	19

Codi R utilitzat per obtenir els ajusts de les còpules

```
install.packages("readxl")
install.packages("writexl")
install.packages("dplyr")
install.packages("purrr")
install.packages("VineCopula")
library(readxl)
library(writexl)
library(dplyr)
library(purrr)
library(VineCopula)

#Llegir dades
dades <- read_xlsx("/Users/mariaaiguade/Desktop/DADES/BD_FINAL.xlsx", sheet = "RATIS")
```

```

#Definició Paràmetres
perioodes <- c("pre", "COVID", "post")
ràtio1 <- "Ri"
ràtio2 <- "Rj"
alphas <- c(0.05, 0.1)

resultats_total <- map_dfr(perioodes, function(periode_objectiu) {

#Filtrar i seleccionar columnes
d_filtre <- dades %>%
  filter(periode == periode_objectiu) %>%
  select(NIF, bucket, periode, all_of(ràtio1), all_of(ràtio2)) %>%
  mutate(bucket = as.character(bucket)) %>%
  tidyr::drop_na(bucket)

#Loop per bucket
resultats <- d_filtre %>%
  group_by(bucket) %>%
  group_modify(function(df, key) {
    df2 <- df %>%
      mutate(
        across(all_of(c(ràtio1,ràtio2)), ~ na_if(., "NaN")),
        across(all_of(c(ràtio1,ràtio2)), ~ as.numeric(.))
      ) %>%
      filter(complete.cases(across(all_of(c(ràtio1,ràtio2)))))

    n <- nrow(df2)

    if (n < 2) return(tibble())

    X1 <- df2[[ràtio1]]
    X2 <- df2[[ràtio2]]

    U1 <- rank(X1)/(n+1)
    U2 <- rank(X2)/(n+1)

    fit <- BiCopSelect(U1, U2, familyset = c(0, 1, 2, 3, 4, 5), rotations = FALSE)

    tau_model <- BiCopPar2Tau(fit$family, fit$par, fit$par2)
    tau_emp <- cor(U1, U2, method = "kendall")
    rho_emp <- cor(U1, U2, method = "spearman")
    tails <- BiCopPar2TailDep(fit$family, fit$par, fit$par2)

    tibble(
      periode = periode_objectiu,
      Ràtio_1 = ràtio1,
      Ràtio_2 = ràtio2,
      n = n,
      family = fit$family,
      family_name = BiCopName(fit$family),

```

```

    par = fit$par, par2 = fit$par2,
    AIC = fit$AIC, BIC = fit$BIC,
    tau_emp = tau_emp, tau_model = tau_model, rho_emp = rho_emp,
    lambdaL = as.numeric(tails[1]), lambdaU = as.numeric(tails[2]),
  )
}) %>%
ungroup()

return(resultats)

})

write_xlsx(resultats_total, "/Users/mariaaiguade/Desktop/resultats_copules.xlsx")

```

Codi R utilitzat per obtenir els càlculs del VaR, ES i probabilitats condicionades

```

install.packages("readxl")
install.packages("writexl")
install.packages("dplyr")
install.packages("purrr")
install.packages("VineCopula")
library(readxl)
library(dplyr)
library(tidyr)
library(purrr)
library(writexl)
library(VineCopula)

dades <- read_xlsx("/Users/mariaaiguade/Desktop/DADES/BD_FINAL.xlsx", sheet = "RATIS (2)")

perioodes <- c("pre", "COVID", "post")
ratio <- "Ri"
ratios_totals <- paste0("R", 1:9)
ratios_restants <- setdiff(ratios_totals, ratio)
alphas <- c(0.05, 0.1)

#Funcions auxiliars
VaR<- function(x, a) quantile(x, probs = a, type = 7, na.rm = TRUE)
ES<- function(x, a) {
  mean(x[x <= VaR(x, a)], na.rm = TRUE)
}
prob_tail_cond <- function(fit, a) {
  C <- BiCopCDF(a, a, family = fit$family, par = fit$par, par2 = fit$par2)
  C/a
}

#Filtrar i seleccionar columnes sense NaN
d_filtre <- dades %>%
  select(NIF, bucket, periode, all_of(ratios_totals)) %>%

```

```

mutate(bucket = as.character(bucket)) %>%
drop_na(bucket) %>%
mutate(across(all_of(ratios_totals), ~ na_if(., "NaN"))) %>%
mutate(across(all_of(ratios_totals), as.numeric))

#Càlcul VaR i ES
resultats_var_es <- map_dfr(periodes, function(p) {

  d_filtre %>%
    filter(periode == p) %>%
    group_by(bucket) %>%
    group_modify(function(df, key) {

      df2 <- df %>% filter(!is.na(.data[[ratio]]))
      n <- nrow(df2)
      if (n < 2) return(tibble())

      X1 <- df2[[ratio]]

      tibble(
        periode = p,
        Ratio = ratio,
        n = n,

        VaR_5 = VaR(X1, 0.05),
        ES_5 = ES(X1, 0.05),
        VaR_10 = VaR(X1, 0.10),
        ES_10 = ES(X1, 0.10)
      )
    }) %>%
    ungroup()
})

#càlcul probabilitats
resultats_prob <- map_dfr(periodes, function(p) {

  d_filtre %>%
    filter(periode == p) %>%
    group_by(bucket) %>%
    group_modify(function(df, key) {

      # si el rati objectiu no té dades suficients, no té sentit calcular probabilitats
      df2 <- df %>% filter(!is.na(.data[[ratio]]))
      if (nrow(df2) < 2) return(tibble())

      map_dfr(ratios_restants, function(ratio_alt) {

        df3 <- df %>%
          filter(!is.na(.data[[ratio]]),
                 !is.na(.data[[ratio_alt]]))
      })
    })
})

```

```

n <- nrow(df3)
if (n < 2) return(tibble())

X1 <- df3[[ratio]]
X2 <- df3[[ratio_alt]]

if (length(unique(X1)) < 2 || length(unique(X2)) < 2) return(tibble())

U1 <- rank(X1) / (n + 1)
U2 <- rank(X2) / (n + 1)

fit <- BiCopSelect(U1, U2,
                   familyset = c(0, 1, 2, 3, 4, 5),
                   rotations = FALSE)

prob_5 <- prob_tail_cond(fit, 0.05)
prob_10 <- prob_tail_cond(fit, 0.10)

tibble(
  periode = p,
  Ratio = ratio,
  Ratio_alt = ratio_alt,
  n = n,

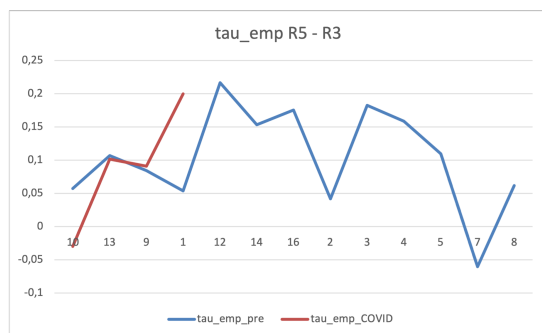
  family_name = BiCopName(fit$family),
  AIC = fit$AIC,

  P_U1_U2_5 = prob_5,
  P_U1_U2_10 = prob_10
)
}) %>%
ungroup()
})

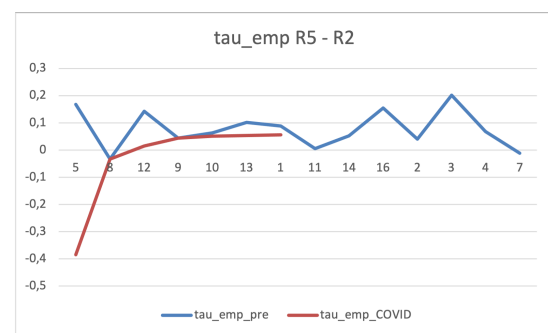
#Exportatr excel
write_xlsx(
  list(
    "VaR_ES" = resultats_var_es,
    "Probabilitats" = resultats_prob
  ),
  "/Users/mariaaiguade/Desktop/resultats_per_rati.xlsx"
)

```

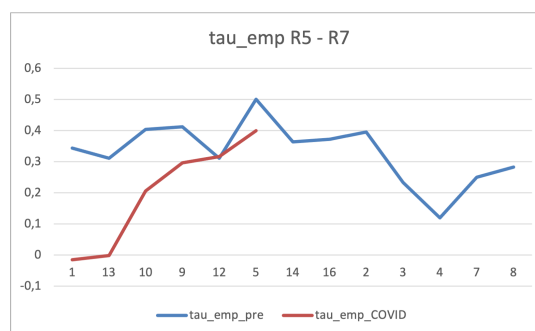

Gràfiques de τ de Kendall empírica



τ de Kendall empírica parella R5 - R3

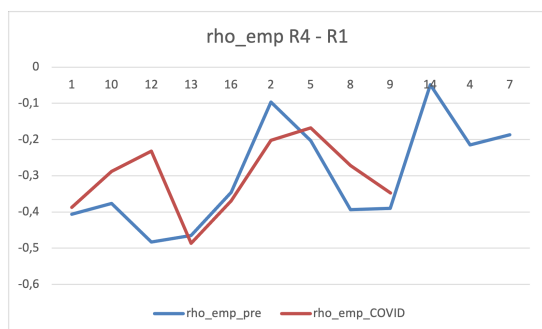


τ de Kendall empírica parella R5 - R2

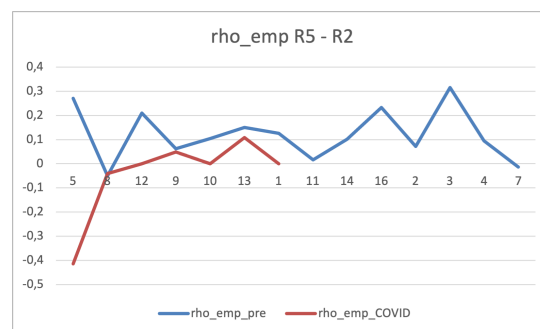


τ de Kendall empírica parella R5 - R7

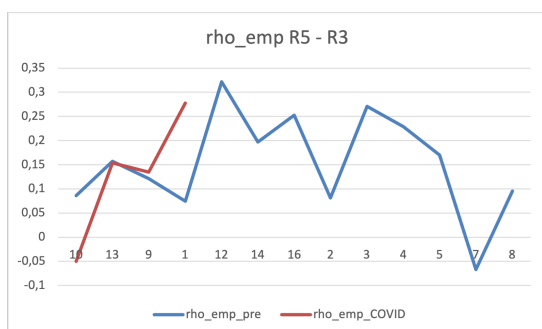
Gràfiques de ρ de Spearman empírica



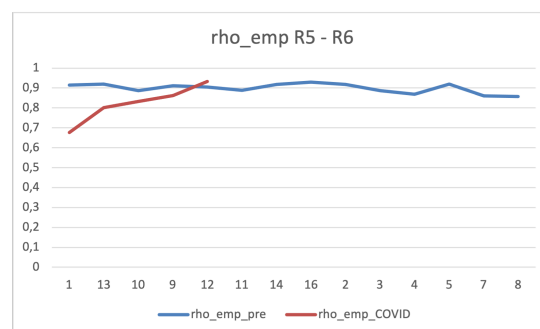
ρ de Spearman empírica parella R4 - R1



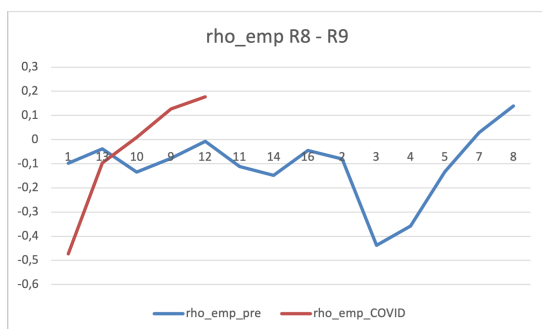
ρ de Spearman empírica parella R5 - R2



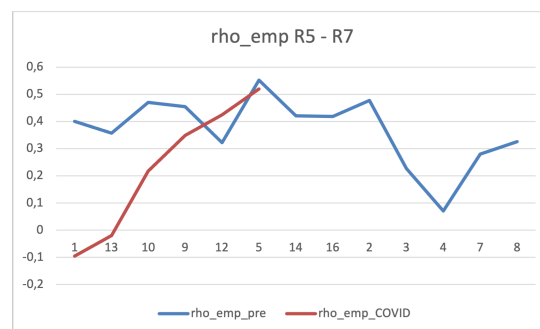
ρ de Spearman empírica parella R5 - R3



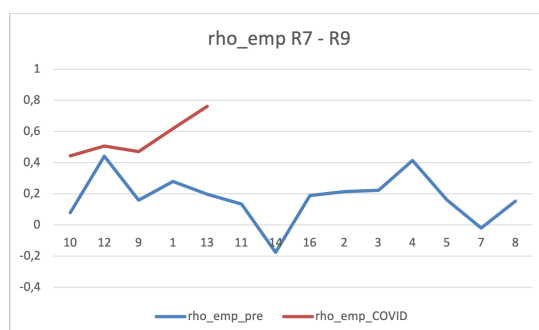
ρ de Spearman empírica parella R5 - R6



ρ de Spearman empírica parella R8 - R9

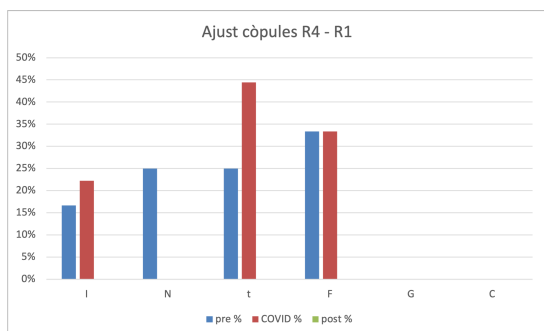


ρ de Spearman empírica parella R5 - R7

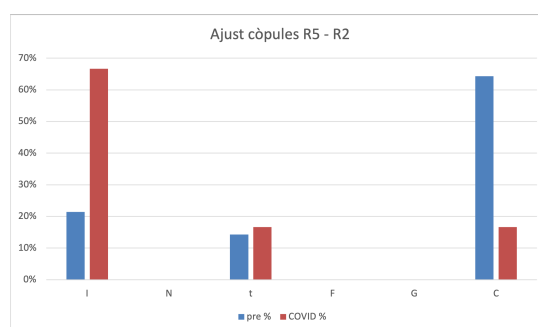


ρ de Spearman empírica parella R7 - R9

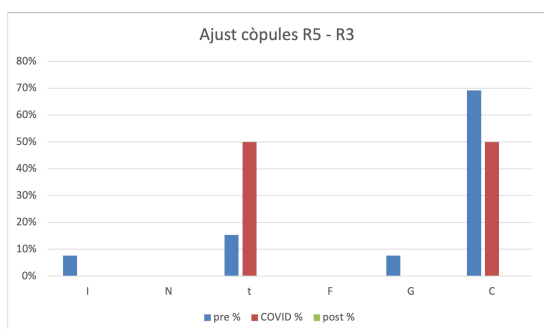
Ajust de Còpules



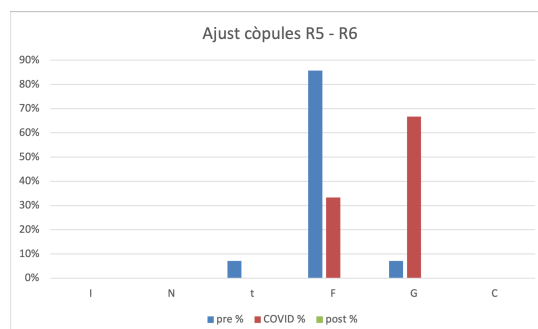
Ajust còpules parella R4 - R1



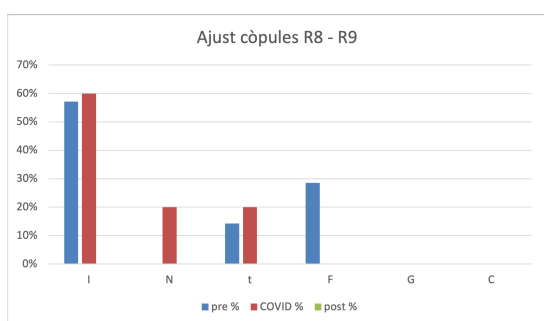
Ajust còpules parella R5 - R2



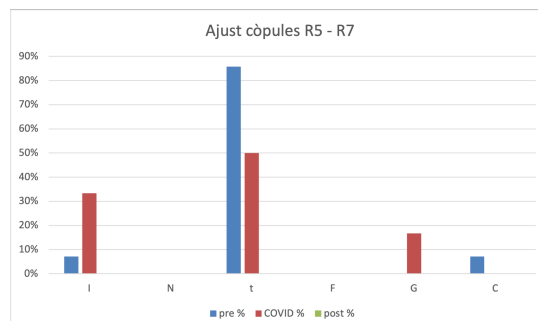
Ajust còpules parella R5 - R3



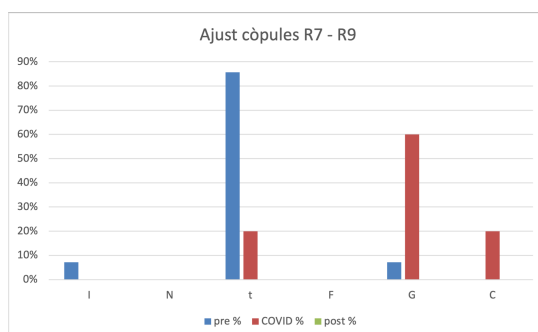
Ajust còpules parella R5 - R6



Ajust còpules parella R8 - R9

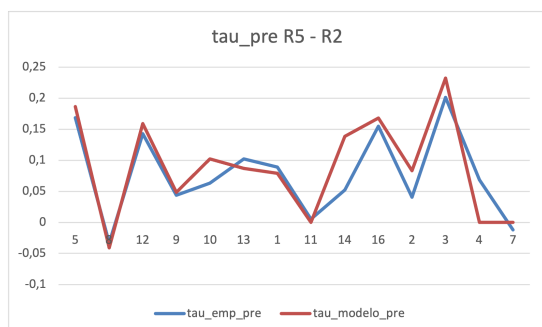


Ajust còpules parella R5 - R7

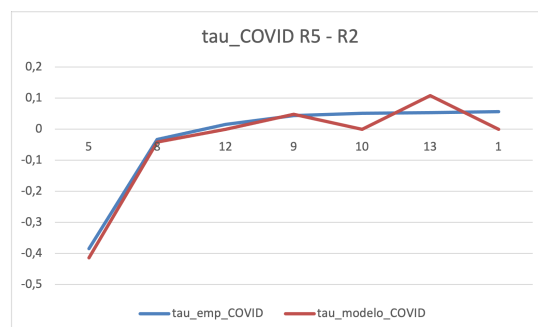


Ajust còpules parella R7 - R9

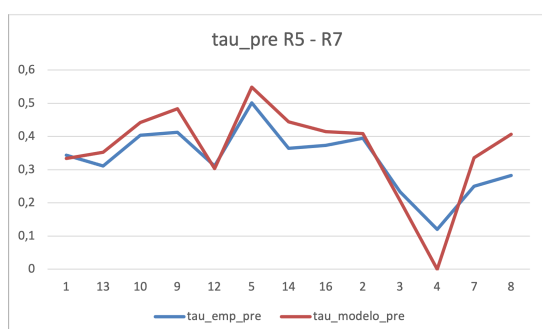
Gràfiques validació de models: dependència observada vs modelada



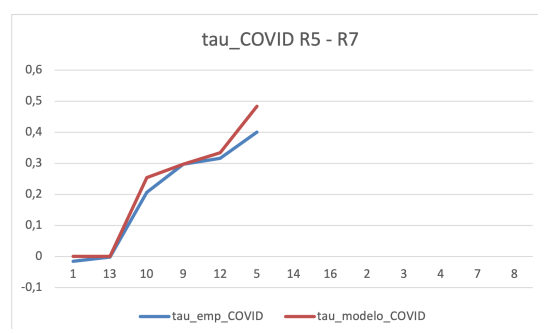
Comparació pre - COVID de la τ del model amb l'empírica de la parella R5 - R2



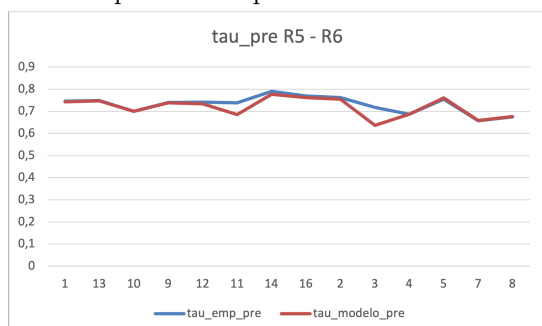
Comparació COVID de la τ del model amb l'empírica de la parella R5 - R2



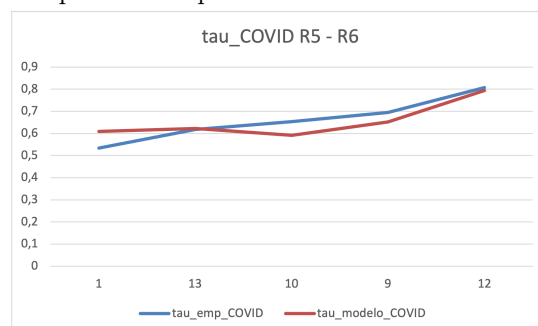
Comparació pre - COVID de la τ del model amb l'empírica de la parella R5 - R7



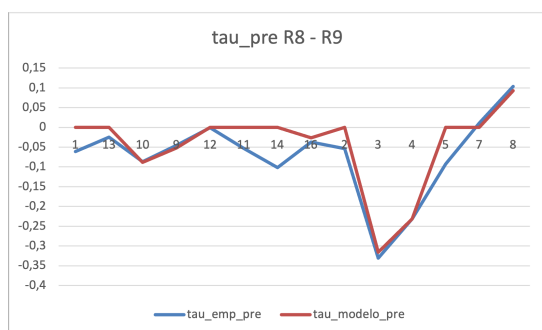
Comparació COVID de la τ del model amb l'empírica de la parella R5 - R7



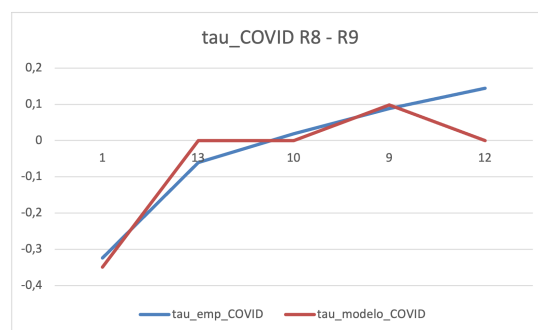
Comparació pre - COVID de la τ del model amb l'empírica de la parella R5 - R6



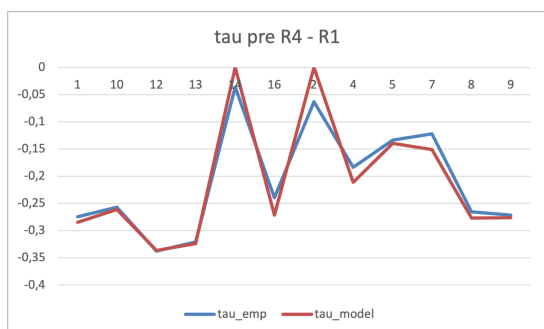
Comparació COVID de la τ del model amb l'empírica de la parella R5 - R6



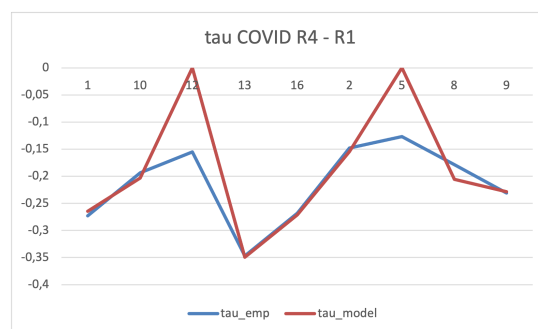
Comparació pre - COVID de la τ del model amb l'empírica de la parella R9 - R8



Comparació COVID de la τ del model amb l'empírica de la parella R9 - R8

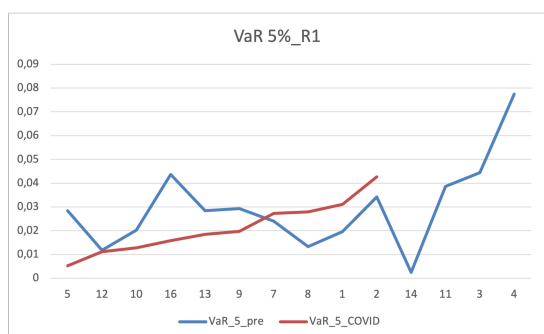


Comparació pre - COVID de la τ del model amb l'empírica de la parella R4 - R1

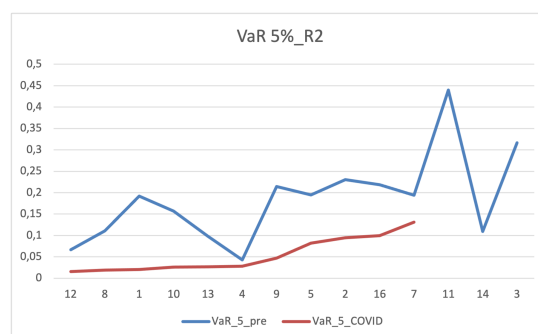


Comparació pre - COVID de la τ del model amb l'empírica de la parella R4 - R1

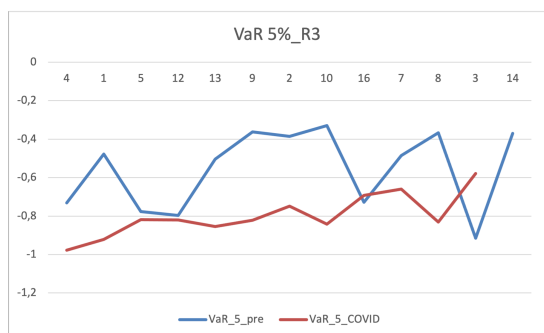
Gràfiques càlculs VaR i ES



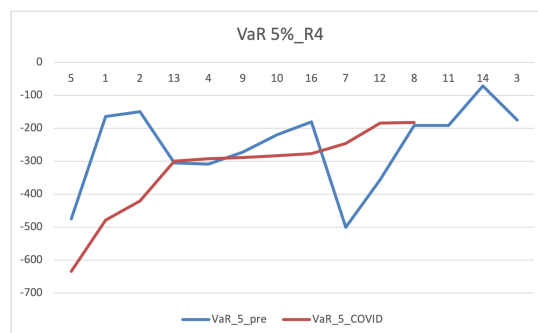
VaR 5% Ràtio 1



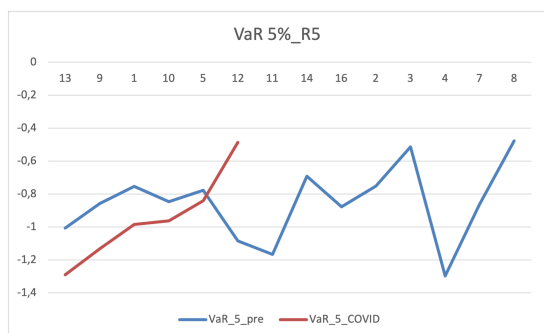
VaR 5% Ràtio 2



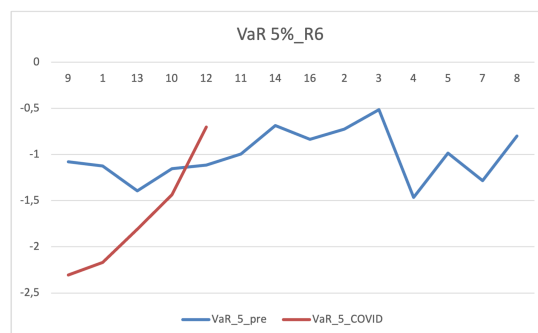
VaR 5% Ràtio 3



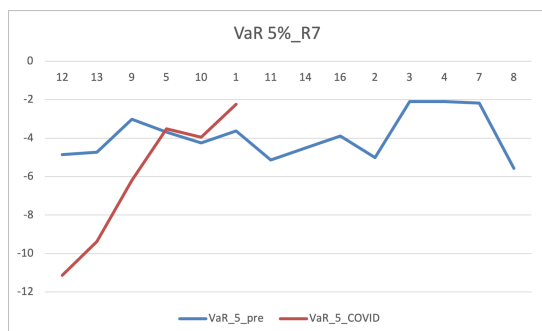
VaR 5% Ràtio 4



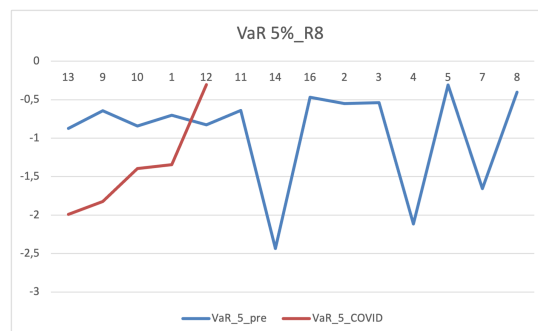
VaR 5% Ràtio 5



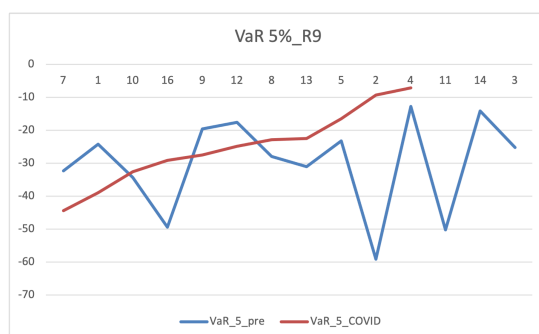
VaR 5% Ràtio 6



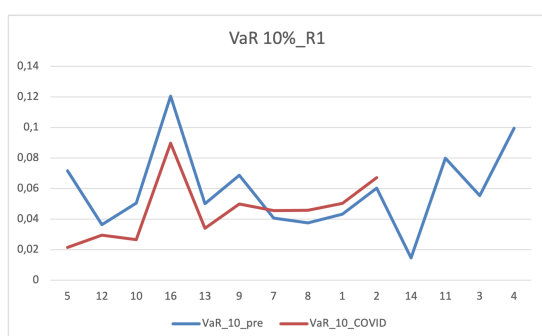
VaR 5% Ràtio 7



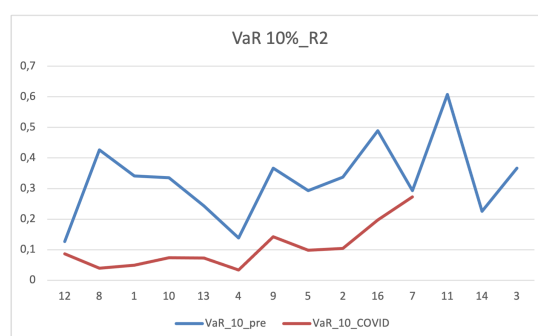
VaR 5% Ràtio 8



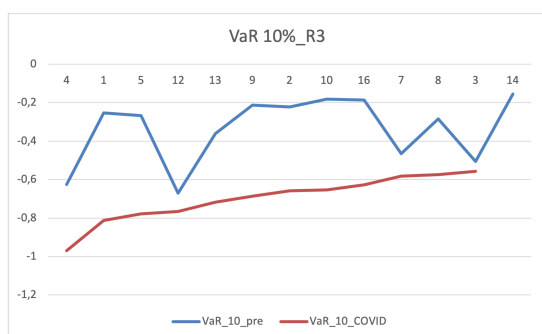
VaR 5% Ràtio 9



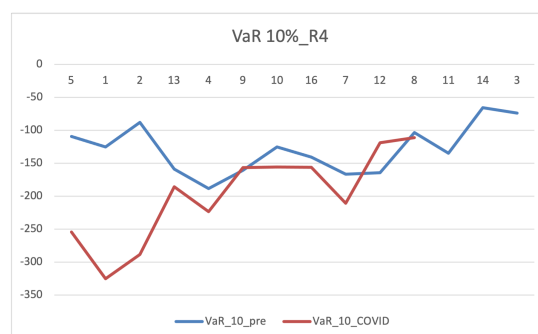
VaR 10% Ràtio 1



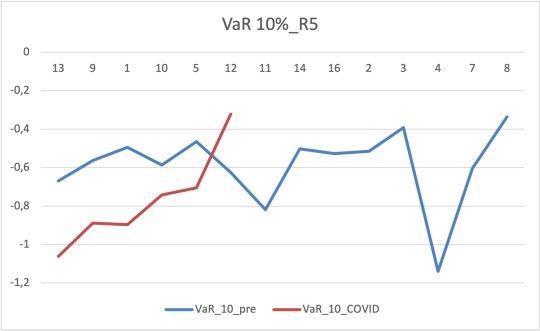
VaR 10% Ràtio 2



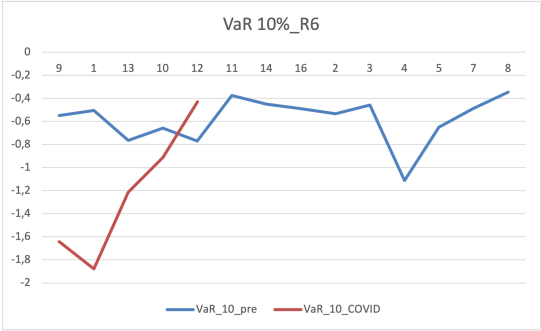
VaR 10% Ràtio 3



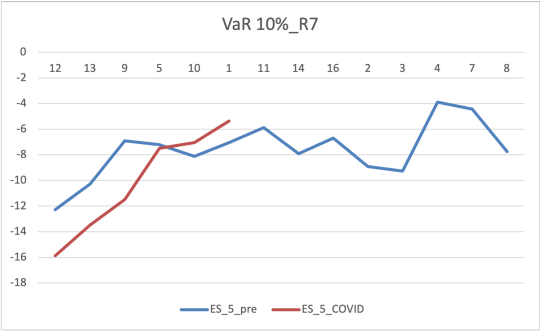
VaR 10% Ràtio 4



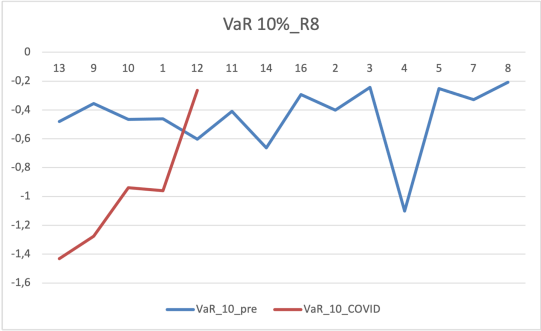
VaR 10% Ràtio 5



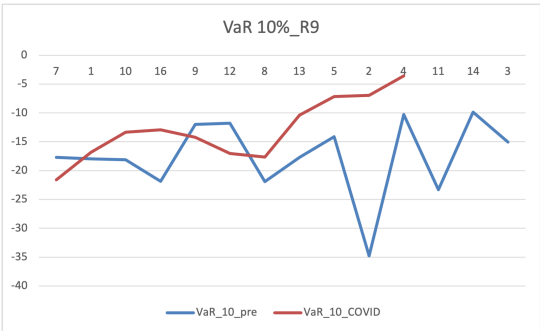
VaR 10% Ràtio 6



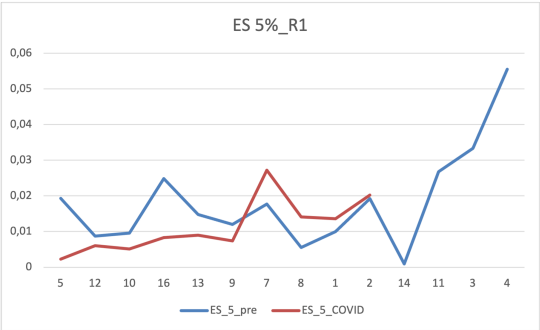
VaR 10% Ràtio 7



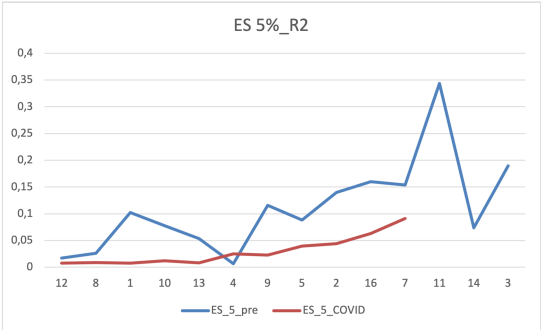
VaR 10% Ràtio 8



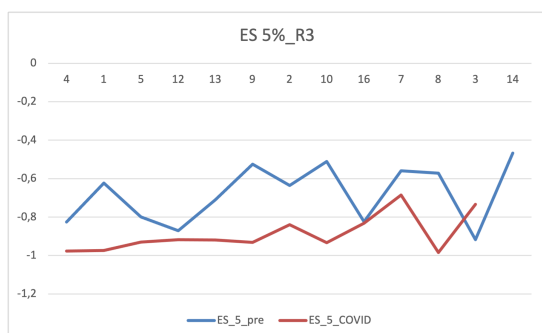
VaR 10% Ràtio 9



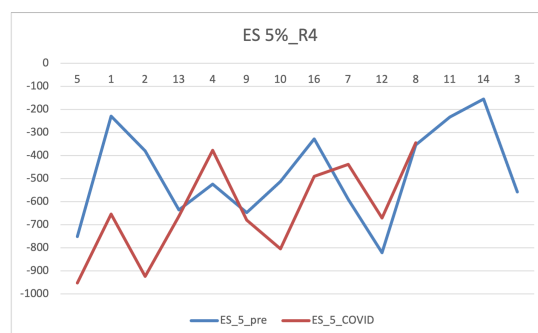
ES 5% Ràtio 1



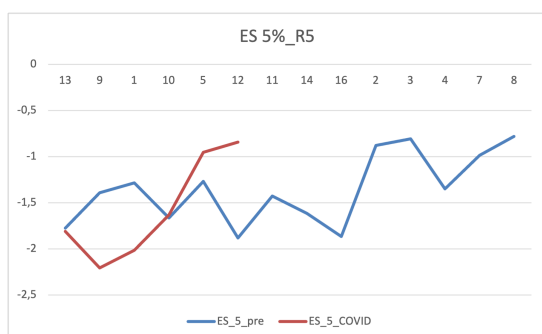
ES 5% Ràtio 2



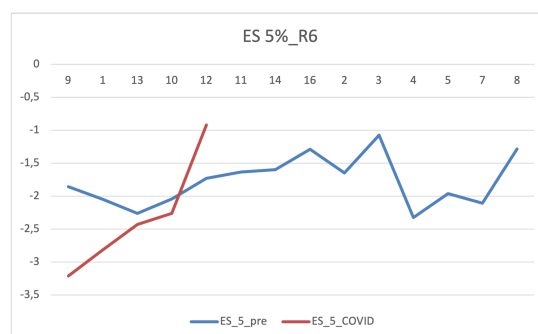
ES 5% Ràtio 3



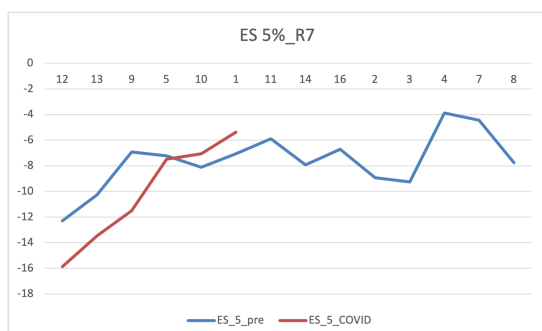
ES 5% Ràtio 4



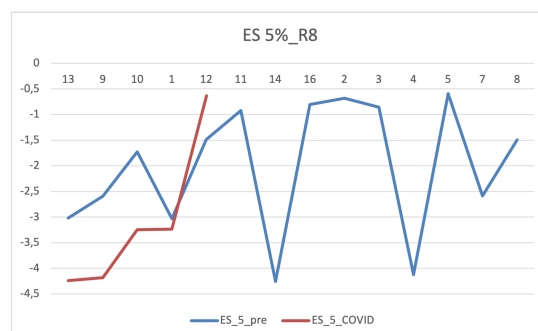
ES 5% Ràtio 5



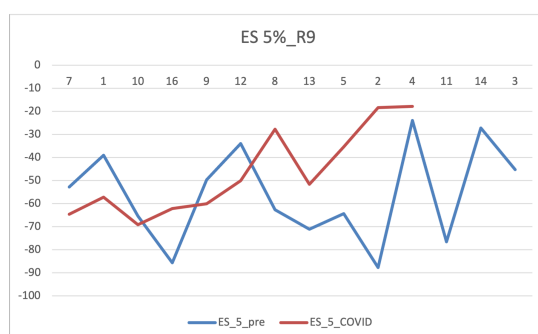
ES 5% Ràtio 6



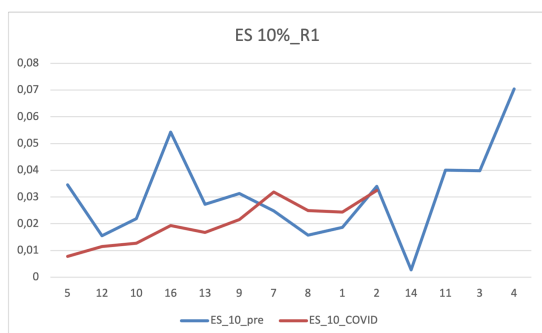
ES 5% Ràtio 7



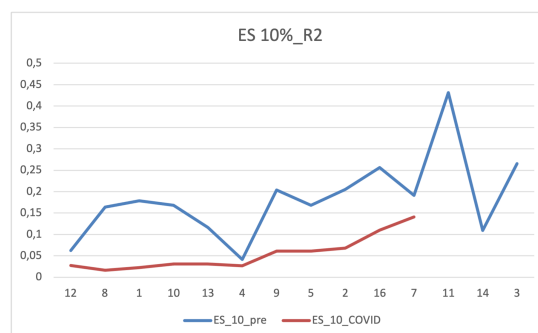
ES 5% Ràtio 8



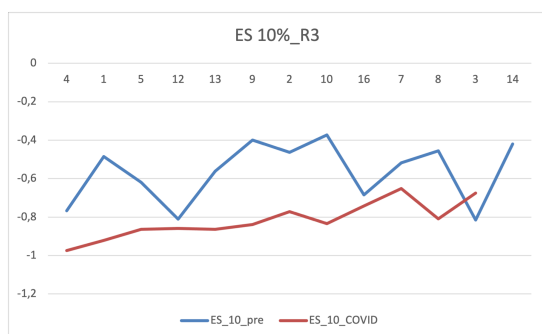
ES 5% Ràtio 9



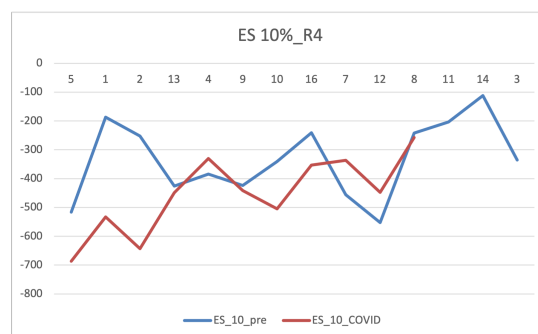
ES 10% Ràtio 1



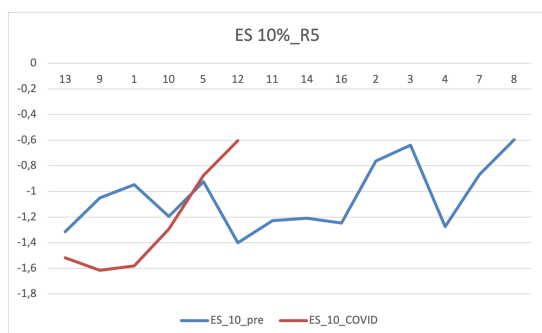
ES 10% Ràtio 2



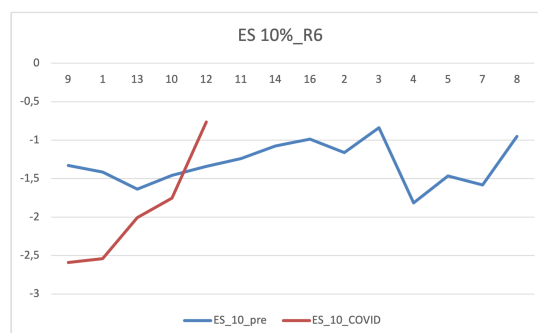
ES 10% Ràtio 3



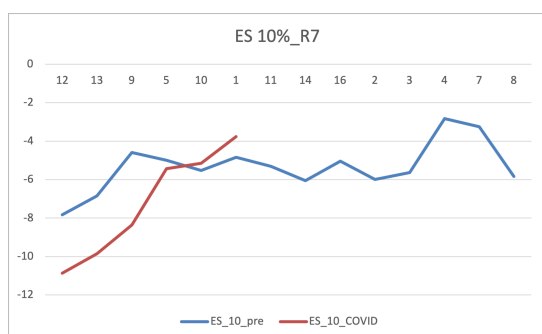
ES 10% Ràtio 4



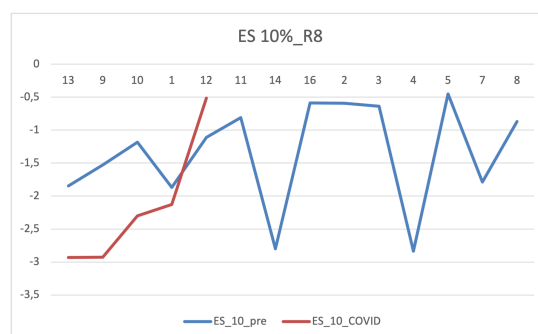
ES 10% Ràtio 5



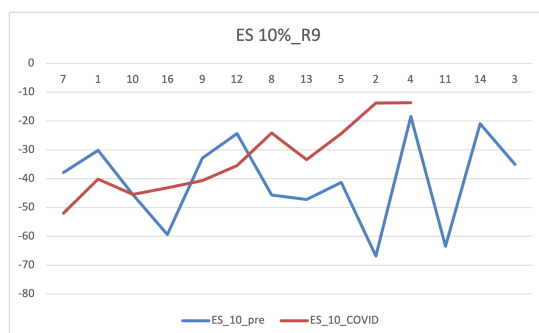
ES 10% Ràtio 6



ES 10% Ràtio 7

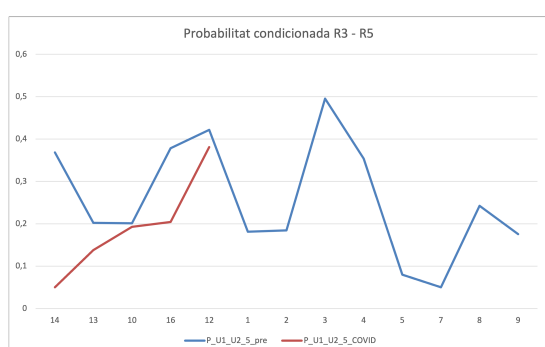


ES 10% Ràtio 8

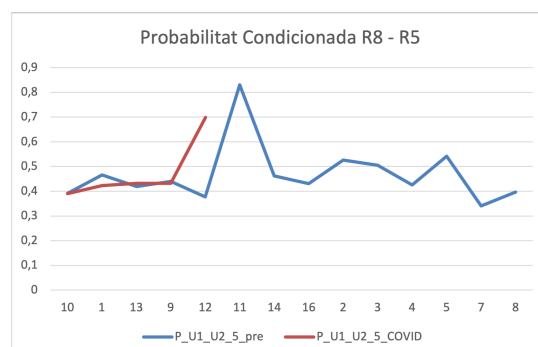


ES 10% Ràtio 9

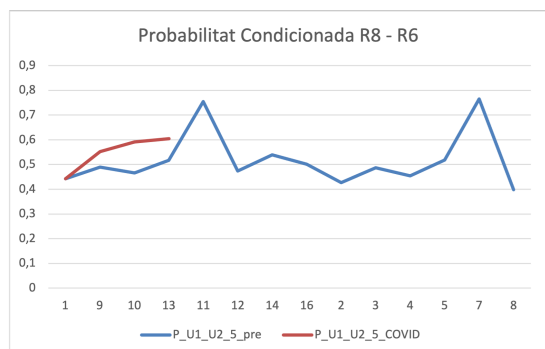
Gràfiques probabilitats condicionades



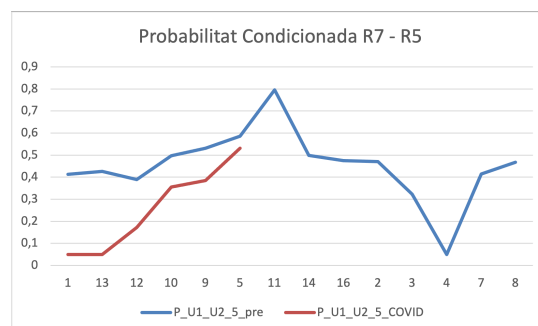
Probabilitat condicionada R3 - R5



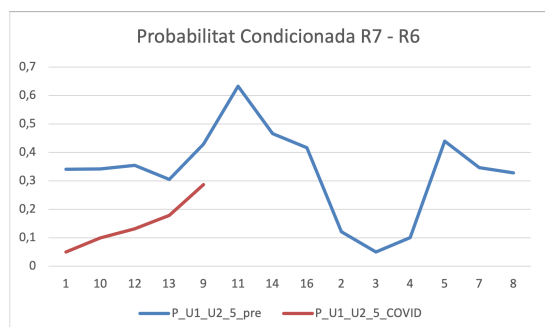
Probabilitat condicionada R8 - R5



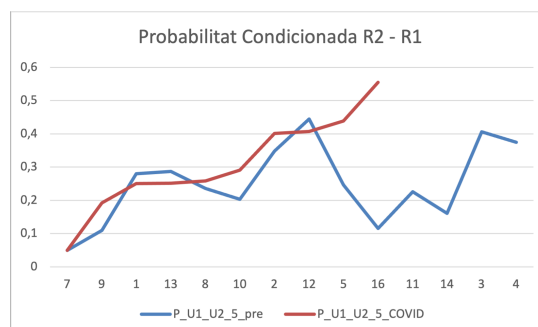
Probabilitat condicionada R8 - R6



Probabilitat condicionada R7 - R5



Probabilitat condicionada R7 - R6



Probabilitat condicionada R2 - R1