

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : **Statistique**

Par

BAHRI Yousra

Titre :

Simulation des données dépendantes avec logiciel R

Membres du Comité d'Examen :

Dr. BRAHIMI Brahim	UMKB	Encadreur
Dr. SOLTANE Louiza	UMKB	Présidente
Dr. BENBRAIKA Ghazlene	UMKB	Examineur

Juin 2020

DÉDICACE

Je dédie cette mémoire

À Ma chère Mère.

À Mon chère père.

À mes soeurs.

À mes frères.

À mes chères amies.

À tous ceux qui étaient à côtés de moi et qui m'ont soutenu dans ma carrière universitaire.

REMERCIEMENTS

Mes premiers remerciements à Dieu tout-puissant pour la volonté et la patience qu'il m'a données pour achever mon humble travail.

Mes sincères remerciements, mon respect à mes parents.

Je tiens à exprimer toute ma reconnaissance à mon Directeur de mémoire **Dr-Brahimi Brahim.**, professeur de mathématique à l'université de Biskra.

Mes vifs remerciements vont également aux membres du Jury : **Dr-Soltani Louiza**,
,Dr-Benbraika Ghazlene

Je remercie à tous les enseignants du département de Mathématiques.

Liste des abréviations

Ω :	L'espace fondamental.
C :	Copule.
F :	Fonction de distribution.
F_X :	La fonction de répartition de la variable aléatoire X .
$i : i : d$:	Indépendantes et identiquement distribuées.
ssi :	Si et seulement si.
$v.a$:	Variable aléatoire définie sur l'espace (Ω, F, P) .
Var :	Variance mathématique.

Table des matières

Dédicace	i
Remerciements	ii
Table des matières	iv
Introduction	1
1 Notions et définitions	2
1.1 Introduction :	2
1.1.1 Variables aléatoires discrètes :	2
1.1.2 Variable aléatoire continue :	2
1.2 Couples et vecteurs aléatoires :	4
1.2.1 Indépendance de variables aléatoires :	5
1.2.2 Définition de la fonction de répartition bivariable :	5
1.3 Loi jointe :	6
2 Copule et la mesures d'association :	7
2.1 Les mesures de dépendance :	7
2.1.1 Coefficient de corrélation linéaire :	7
2.1.2 Les coefficients de corrélation de Kendall et de Spearman :	8
2.2 Copules : définition et propriétés :	12
2.2.1 Théorème de Sklar :	13
2.2.2 La densité d'une copule :	13

2.2.3 Les familles de copules :	13
Conclusion	22
Bibliographie	23

Introduction

Les relations de la dépendance entre les variables aléatoires sont l'un des sujets les plus étudiés dans la probabilité et la statistique, Mesurer la dépendance entre deux ou plusieurs variables aléatoires est une pratique largement utilisées par les statisticiens. Beaucoup des ces mesures a été proposées tel que ; le coefficient de corrélation de Pearson, le tau de Kendall et le rho de Spearman. Bien que ces mesures sont simples à calculer et peuvent être facilement interprétées, elles ne sont pas en mesure de détecter toutes les formes de dépendances, donc il était indéniable de trouver un autre moyen pour résoudre ce problème. En effet, la fonction copule a l'avantage de modéliser complètement la dépendance entre les variables. et chaque fois que je donne une application dans \mathbb{R} .

Ce mémoire est constitué de deux chapitres :

Le premier chapitre se veut comme une introduction à certaines notions de base de probabilité et du concept de dépendance.

Le deuxième chapitre est constitué de deux parties. Dans la première partie nous présentons l'aspect des mesures d'association : la mesure de dépendance en forme de copule, telle que le rho de Spearman et le tau de Kendall. L'importance de l'utilisation des copules dans la modélisation de la dépendance, et nous rappeler la mesure de concordance.

Dans la deuxième partie nous introduisons les définitions de base des copules, leurs propriétés, ainsi que les principaux théorèmes liés à cette théorie, en particulier le théorème de Sklar. De même quelque exemple de familles

Chapitre 1

Notions et définitions

1.1 Introduction :

Dans ce chapitre on va introduit et étudié des notions de base du concept de dépendance et probabilité On y présentera quelques caractéristiques et propriétés fondamentale $v - a$.

Soient (Ω, A, P) un espace probabilisé et X une variable aléatoire définie sur cette espace.

1.1.1 Variables aléatoires discrètes :

Définition 1.1 On dit que X un $v - a$ si l'ensemble $X(\Omega)$ discret (\mathbb{N} ou \mathbb{Z})

Loi de probabilité :

La loi de probabilité d'une $v - a$ discrète est entièrement déterminée par les probabilités P_i d'évènements $\{X = x_i\}$, x_i parcourant l'univers image $X(\Omega)$. La loi de probabilité est donnée par les (x_i, P_i) .

Notation 1 $P(\{X = x_i\})$ équivalent $P(X = x_i)$ ou P_i .

1.1.2 Variable aléatoire continue :

Définition 1.2 On dit que X est une variable aléatoire absolument continue si sa fonction de répartition F est continue et dérivable à gauche et à droite de tout point de la fonction

dérivée f de F est dite fonction densité de probabilité de X et vérifie les relations :

$$\forall x \in \mathbb{R}, f(x) = F'(x).$$

Et

$$F(x) = \int_{-\infty}^x f(x)dx.$$

Le résultat suivant résume les principales propriétés de la fonction densité de probabilité f d'une variable aléatoire continue X .

1. $\forall x \in \mathbb{R}, f(x) \geq 0$.
2. L'intégrale $\int_{-\infty}^{+\infty} f(x)dx$ est convergente et on a $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Loi de probabilité d'une variable aléatoire continue :

Dans toute la suite X désigne une variable aléatoire continue et f représente sa fonction densité de probabilité.

Pour tout nombre réel on a :

$$P(X \leq \alpha) = F(\alpha) = \int_{-\infty}^{\alpha} f(x)dx$$

et

$$P(X > \alpha) = \int_{\alpha}^{+\infty} f(x)dx = 1 - F(\alpha)$$

Pour tous réels α et β tels que $\alpha \leq \beta$ on a :

$$P(\alpha \leq X \leq \beta) = F(\beta) - F(\alpha) = \int_{\alpha}^{\beta} f(x)dx$$

Simulation avec logiciel R :

Exemple 1.1 `> x <- round(runif(20, 0, 20), digits = 1)`

[1]6.9 0.4 13.6 2.2 13.1 12.0 16.1 12.7 19.9 11.9 16.6 7.5 12.4 3.8 14.4

[16]2.8 13.4 10.1 13.0 11.9

```
> #20 numbers, between 0 and 20
```

```
> #roundedat1e - 1
```

Exemple 1.2 `> n < -10`

```
> x < -rnorm(n)
```

```
> print(x)
```

```
[1] 0.28903195 -1.11601180 0.02802541 -0.18483669 -0.39935578 -1.64387063
```

```
[7] -1.92499026 -2.59820485 -1.24957524 0.03160353
```

1.2 Couples et vecteurs aléatoires :

Définition 1.3 *Les applications $X_1, X_2, X_3, \dots, X_n$ sont des variables aléatoires réelles appelées composantes du vecteur aléatoire X . On note alors $X = (X_1, X_2, X_3, \dots, X_n)$.*

Simulation avec logiciel R

Il existe un grand nombre de fonctions pour créer des vecteurs avec R.

Les principales sont :

- La fonction `c()` prend un nombre arbitraire d’arguments et retourne un vecteur en concaténant ces arguments.

```
> c(2,10,5,37,4,8)
```

```
[1] 2 10 5 37 4 8
```

```
> c(1,2,c(5,10))
```

```
[1] 1 2 5 10
```

- La fonction `rep()` prend un ou plusieurs éléments et les crée ‘ n ’ fois, elle retourne donc un vecteur qui sera la répétition de ces éléments.

```
> rep(1,5)
```

```
[1] 1 1 1 1 1
```

```
> rep(c(1,2),3)
```

```
[1] 1 2 1 2 1 2
```

- L'opérateur `' :'` génère un vecteur qui représente une suite de nombres ordonnée par ordre croissant ou décroissant.

```
> 2 :11
```

```
[1] 2 3 4 5 6 7 8 9 10 11
```

```
> 5 :-5
```

```
[1] 5 4 3 2 1 0 -1 -2 -3 -4 -5
```

- La fonction `seq()` est assez similaire à l'opérateur `' :'` avec comme particularité supplémentaire la possibilité de donner un pas à la séquence de nombres générée.

```
> seq(2,11)
```

```
[1] 2 3 4 5 6 7 8 9 10 11
```

```
> seq(1,25,by=2)
```

```
[1] 1 3 5 7 9 11 13 15 17 19 21 23 25
```

- D'autres méthodes existent pour créer des vecteurs notamment grâce aux fonctions `vector()`, `sequence()`, `scan()`...

1.2.1 Indépendance de variables aléatoires :

Soit deux variables aléatoires X_1 et X_2 continues, on dit que X_1 est indépendante de X_2 ssi :

- $F(x_1, x_2) = F_1(x_1)F_2(x_2)$.
- $F(x_1 | x_2) = F_1(x_1)$.

1.2.2 Définition de la fonction de répartition bivariée :

Définition 1.4 Une fonction $F : \mathbb{R}^2 \rightarrow [0, 1]$ continue à droite est dite une fonction de répartition bivariée ssi :

1.

$$\lim_{x_i \rightarrow -\infty} F(x_1, x_2) = 0 \text{ pour } i \in \{1, 2\}.$$

$$\lim_{x_i \rightarrow +\infty} F(x_1, x_2) = 1.$$

2. F est 2-croissante (inégalité de rectangle) : pour tout (α_1, α_2) et $(\beta_1, \beta_2) \in \mathbb{R}^2$ tels que :
 $\alpha_1 < \beta_1$ et $\alpha_2 < \beta_2$, on a :

$$F(\alpha_1, \alpha_2) + F(\beta_1, \beta_2) - F(\alpha_1, \beta_2) - F(\beta_1, \alpha_2) \geq 0.$$

1.3 Loi jointe :

Définition 1.5 Soient X et Y deux variables aléatoires. La loi jointe de (X, Y) est définie par sa fonction de répartition :

$$\begin{aligned} F_{(X,Y)} : \mathbb{R} \times \mathbb{R} &\rightarrow [0, 1] . \\ (x, y) &\rightarrow \mathbb{P}(X \leq x, Y \leq y) . \end{aligned}$$

Les **loi marginales** du couple (X, Y) sont la loi $F_1(X)$ de X et $F_2(Y)$ de Y .

On désignera par $C(F_1, F_2)$ la classe des fonctions de répartition bivariées de marginales F_1 et F_2 .

Les marginales s'obtiennent de la fonction de répartition conjointe telle que :

$$F_1(x) = \lim_{x_2 \rightarrow +\infty} F(x, y).$$

$$F_2(y) = \lim_{x_1 \rightarrow +\infty} F(x, y).$$

Chapitre 2

Copule et la mesures d'association :

1

2.1 Les mesures de dépendance :

2.1.1 Coefficient de corrélation linéaire :

Définition 2.1 Soient X et Y deux variables aléatoires ayant des variances finies.

Le coefficient de corrélation linéaire des variables X et Y est donné par :

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(x)Var(Y)}}.$$

Où

$$Cov(X, Y) = E(X.Y) - E(X)E(Y) .$$

$$Var(X) = E(X^2) - E(X)^2.$$

Néanmoins, ce coefficient ne de meure pas constant sous l'hypothèse d'une transformation croissante non-linéaire.

Et on a si X et Y deux $v - a$ indépendantes alors $\rho(X, Y) = 0$ mais la réciproque est fausse contre-exemple soient $X \rightsquigarrow N(0, 1)$ et $Y = X^2$ alors :

$$\rho(X, Y) = E(X^3) = 0.$$

Simulation avec logiciel R :

```
x<-c(90 ,120,150,180,210,240, 270,300 )
z<-c(5.77, 5.55, 5.31, 5.10 ,4.89 ,4.65, 4.42, 4.23)
r<-cor(x,z)
```

2.1.2 Les coefficients de corrélation de Kendall et de Spearman :

Notion de la concordance :

Définition 2.2 Notons (x_1, y_1) et (x_2, y_2) deux observations d'un vecteur aléatoire continue (X, Y) , alors (x_1, y_1) et (x_2, y_2) sont dites :

1. **Concordantes** si :

$$(x_1 - x_2)(y_1 - y_2) > 0 \iff (x_1 < x_2 \text{ et } y_1 < y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 > y_2).$$

2. **Discordantes** si :

$$(x_1 - x_2)(y_1 - y_2) < 0 \iff (x_1 < x_2 \text{ et } y_1 > y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 < y_2).$$

Plus généralement soit $\{(x_1, y_1), (x_2, y_2)\}$ un échantillon de n observations d'un couple (X, Y) Il existe :

$$C_n^2 = n!/2!(n-1)!$$

Pairs de distributions de couples $(x_i, y_i), (x_j, y_j)$ qui sont soit concordantes, soit discordantes.

Définition 2.3 La fonction de concordance entre les deux couples (X_1, Y_1) et (X_2, Y_2) est

définie par :

$$Q = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 2].$$

Simulation avec logiciel R

```
# Calculer des paires CONcordantes dans une table
# parcourir x [r, c] et multiplier par
# sum (x éléments en dessous et à droite de x [r, c])
# x = table
concordant <- fonction ( x )
{
  x <- matrice (as.numeric ( x ), dim ( x ))

  # get sum (valeurs de la matrice > r AND > c)
  # pour chaque matrice [r, c]
  mat.lr <- fonction ( r , c )
  {
    lr <- x [( rx > r ) & ( cx > c )]
    somme ( lr )
  }

  # obtenir l'index de ligne et de colonne pour chaque
  # élément de matrice
  rx <- ligne ( x )
  cx <- col ( x )

  # renvoie la somme de chaque matrice [r, c] * sommes
  # utilisation de mapply pour séquencer chaque matrice [r, c]
  somme ( x * mapply ( mat.lr , r = rx , c = cx ))
}

# Calculer les paires DIScordantes dans une table
```

```

# parcourir x [r, c] et multiplier par
# sum (x éléments en dessous et à gauche de x [r, c])
# x = table
discordant <- fonction ( x )
{
  x <- matrice (as.numeric ( x ), dim ( x ))

  # get sum (valeurs de la matrice > r AND <c)
  # pour chaque matrice [r, c]
  mat.ll <- fonction ( r , c )
  {
    ll <- x [( rx > r ) & ( cx < c )]
    somme ( ll )
  }

  # obtenir l'index de ligne et de colonne pour chaque
  # élément de matrice
  rx <- ligne ( x )
  cx <- col ( x )

  # renvoie la somme de chaque matrice [r, c] * sommes
  # utilisation de mapply pour séquencer chaque matrice [r, c]
  somme ( x * mapply ( mat.ll , r = rx , c = cx ))
}

```

Tau de Kendall :

Définition 2.4 (version échantillon) Soit une série de n observations $\{(x_i, y_i)\}_{1 \leq i \leq n}$ d'un couple (X, Y) . Le tau de Kendall est défini par :

$$\tau = \frac{2}{n(n-1)} (n_c - n_d); \tau \in [-1, 1].$$

Où

$$\left\{ \begin{array}{l} n_c : \text{nombre de paires concordantes.} \\ n_d : \text{nombre de paires discordantes.} \\ n : \text{nombre de totale de paires.} \end{array} \right.$$

Définition 2.5 (Version population) Soient (X_1, Y_1) et (X_2, Y_2) deux vecteurs aléatoires continues i.i.d de fonction de répartition conjointe H et des fonctions marginales F (pour X_1 et X_2) et G (pour Y_1 et Y_2). Le tau de Kendall noté est défini par :

$$\tau = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

Le tau de Kendall possède les propriétés suivantes :

- Le taux de Kendall est symétrique, c'est-à-dire : $\tau(X_1, X_2) = \tau(X_2, X_1)$.
- $-1 < \tau < 1$
- Si X et Y sont comonotones alors $\tau = 1$ (la concordance la plus solide).
- Si X et Y sont antimonotones alors $\tau = -1$ (la discordance la plus solide).
- Si X et Y sont indépendantes alors $\tau = 0$
- Si a et b sont des fonctions strictement croissantes, alors $\tau(a(X), b(Y)) = \tau(X, Y)$.

Rho de Spearman :

La valeur de ce coefficient dénoté par ρ_s , est équivalente au coefficient de corrélation de Pearson. Il a été développé par Charles Spearman (1904). La définition de cette mesure est la suivante :

Définition 2.6 Soient (X_1, Y_1) , (X_2, Y_2) et (X_3, Y_3) trois vecteurs aléatoires indépendants de même loi H . Le coefficient de corrélation de Spearman est défini par :

$$\rho_s = 3(\mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) < 0]).$$

Les propriétés du ρ_s sont les suivantes :

1. $-1 \leq \rho_s \leq 1$
2. Si X et Y sont comonotones alors $\rho_s = 1$.
3. Si X et Y sont antimonotones alors $\rho_s = -1$.
4. Si X et Y sont indépendantes alors $\rho_s = 0$.
5. Il est invariant sous des transformations non linéaires strictement croissantes. C'est-à-dire si a et b sont deux fonctions strictement croissantes, alors :

$$\rho_s(a(X), b(Y)) = \rho_s(X, Y).$$

2.2 Copules : définition et propriétés :

Du fait que la théorie multivariée est une extension du cas bivariée, nous étudions la théorie des couples bivariées.

Définition 2.7 (Copules bivariées) Une copule 2-dimensionnelle est une fonction C de $\mathbb{I}^2 = [0, 1]$ dans \mathbb{I} ayant les propriétés suivantes :

i)

$$C(u, 0) = C(0, v), \forall u, v \in \mathbb{I}.$$

ii)

$$C(u, 1) = u \text{ et } C(1, v) = v, \forall u, v \in \mathbb{I}.$$

iii) La copule C est 2-croissantes : $\forall u_1, v_1, u_2, v_2 \in \mathbb{I}$; telle que $u_1 \leq u_2$ et $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

Théorème 2.1 (continuité) Soit C une copule bivariée; pour tout $u_1, u_2, v_1, v_2 \in \mathbb{I}$ avec $u_1 \leq u_2$, et $v_1 \leq v_2$, on a :

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|.$$

2.2.1 Théorème de Sklar :

Si H est la fonction de répartition de (X, Y) alors il existe une copule C de dimension 2 telle que :

$$\forall (x, y) \in \mathbb{R}^2 : H(x, y) = C(F(x), G(y)).$$

Où F, G sont les fonctions de répartitions des X, Y respectivement.

2.2.2 La densité d'une copule :

Les copules admettent des densités de probabilités. Si la densité c associée à la copule C existe, alors elle est définie par :

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}.$$

Simulation avec logiciel R :

`dCopula()` donne la densité,

Exemple 2.1 `norm.cop <- normalCopula (0.5)`

`norm.cop`

one d-vector = matrice à 1 ligne, fonctionne aussi :

`dCopula (c (0.5 , 0.5), norm.cop)`

2.2.3 Les familles de copules :

De nombreuses familles de copules ont été proposées dans la littérature. Les plus utilisées, sont introduites dans cette section.

Exemples des familles de copules :

On commence par les copules usuelles :

Copules usuelles :

Copule d'indépende : La copule associée à deux $v - a$ indépendantes X et Y est donnée par le produit de ces marginales comme suit :

$$C(u, v) = \pi(u, v) = uv, \forall u, v \in \mathbb{I}.$$

Simulation avec logiciel R

```
### Independence copula
```

```
library(copula)
```

```
d <- 2
```

```
ic <- indepCopula(dim = d)
```

```
set.seed(2008)
```

```
u <- runif(d) # a random point in the unit hypercube
```

```
(Pi <- pCopula(u, copula = ic)) # the value of the independence copula at u
```

```
stopifnot(all.equal(Pi, prod(u))) # check numerical equality of the samples
```

Copule min : Définie par :

$$M(u, v) = \min(u, v).$$

Copule max : Définie par :

$$W(u, v) = \max(u + v - 1, 0), \forall u, v \in \mathbb{I}.$$

Copule Archimédienne :

Définition 2.8

On appelle copule Archimédienne de générateur la copule définie par :

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

Avec $\phi : \mathbb{I} \rightarrow [0, 1[$ une fonction continue, strictement décroissante vérifiant $\phi(1) = 0$.

On définit l'inverse de ϕ par $\phi^{[-1]}$ tel que :

$$\phi^{[-1]} = \begin{cases} \phi^{-1}(t) & \text{si } 0 \leq t \leq \phi(0). \\ 0 & \text{si } \phi(0) \leq t \leq +\infty. \end{cases}$$

La famille de Frank : L'une des expressions équivalentes possibles pour les membres de cette famille est :

$$C_{\theta}(u, v) = -\frac{1}{\theta} \ln \left[1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1} \right].$$

Où $u, v \in \mathbb{I}$ et $\theta \in \mathbb{R}$ et

$$\phi_{\theta}(t) = -\ln\left(\frac{\theta^t - 1}{\theta - 1}\right).$$

Simulation avec logiciel R

```
### Frank copula
d <- 2 # dimension
theta <- -9 # copula parameter
fc <- frankCopula(theta, dim = d) # define a Frank copula
set.seed(2010)
n <- 5 # number of evaluation points
u <- matrix(runif(n * d), nrow = n) # n random points in [0,1]^d
pCopula(u, copula = fc) # copula values at u
dCopula(u, copula = fc) # density values at u
```

La famille de Gumbel : Le standard expression de cette famille est :

$$C_{\theta}(u, v) = \exp\{-[(-\ln u)^{\theta} + (-\ln v)^{\theta}]^{\frac{1}{\theta}}\}.$$

Où $\theta \geq 0$ et

$$\phi_{\theta}(t) = (-\ln t)^{\theta}.$$

La famille de Clayton : Le standard expression de cette famille est :

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

Où $u, v \in \mathbb{I}$ et $\theta \geq -1$ et

$$\phi_{\theta}(t) = \frac{1}{\theta} (t^{-1} - 1).$$

Simulation avec logicial R :

```
#####CopuledeClayton:#####
log_vrais_clayton1 <- function(alpha, u, dom){
  l <- -0
  freq <- -table(dom)
  for(iin1:length(freq)){
    ni <- -freq[[i]]
    s <- -u[1:ni]
    u <- -u[-(1:ni)]
    if(freq[[i]] > 1){
      t <- -sum(copula::copClayton@iPsi(s, alpha))
      k <- -0:(ni - 1)
      lnc <- -sum(log(alpha * k + 1)) - (1 + alpha) * sum(log(s)) -
      (ni + 1/alpha) * log(1 + t)
      l <- -l + lnc
    }
  }
  return(l)
}
```

La famille de Joe Le standard expression de cette famille est :

$$C_{\theta}(u, v) = 1 - \left[(1 - u)^{\theta} + (1 - v)^{\theta} - (1 - u)^{\theta} (1 - v)^{\theta} \right]^{\frac{1}{\theta}}.$$

Où $u, v \in \mathbb{I}$ et $\theta \geq 1$ et

$$\phi_{\theta}(t) = -\ln \left[1 - (1 - t)^{\theta} \right].$$

La famille de ALI-MIKHAIL-HAQ (AMH) : Le standard expression de cette famille est :

$$C_{\theta}(u, v) = uv / (1 - \theta(1 - u)(1 - v)).$$

Où $u, v \in \mathbb{I}$ et $-1 \leq \theta \leq 1$. Copule normale

$$\phi_{\theta}(t) = \ln\left(\frac{1 - \theta(1 - t)}{t}\right).$$

Simulation avec logiciel R

```
R> copAMH@nestConstr
function (theta0, theta1)
{
  copAMH@paraConstr(theta0) && copAMH@paraConstr(theta1) &&
  theta1 >= theta0
}
<environment : namespace :nacopula>
R> copAMH@V01
function (V0, theta0, theta1)
{
  rnbinom(length(V0), V0, (1 - theta1)/(1 - theta0)) + V0
}
```

Copules Archimax : Une fonction bivariée est une copule Archimax ssi elle est de la forme :

$$C_{\phi, A}(u, v) = \phi^{-1}[(\phi(u) + \phi(v))] A\left(\frac{\phi(u)}{\phi(u) + \phi(v)}\right).$$

pour tout $u, v \in [0, 1]$ où

1. $A : [0, 1] \rightarrow [\frac{1}{2}, 1]$ telque $\max(t, t - 1) \leq A(t) \leq 1$ pour tout $0 \leq t \leq 1$

2. $\phi :]0, 1] \rightarrow [0, \infty[$ est une fonction convexe, décroissante qui vérifie $\phi(1) = 0$.

Avec la convention suivante : $\phi(0) = \lim_{x \rightarrow 0^+} \phi(x)$ et $\phi^{-1}(y)$ pour tout $x \geq \phi(0)$.

Copules elliptiques : Un vecteur aléatoire $X \in \mathbb{R}^d$ est dit avoir une distribution elliptique s'il admet la représentation stochastique

$$X = \mu + RAU.$$

où $\mu \in \mathbb{R}^d$ et R est positive indépendant de U , U est un vecteur aléatoire uniforme sur la sphère unitaire dans \mathbb{R}^d et A est une matrice $d \times d$ fixe. La famille des copules elliptiques pour laquelle on présente la copule gaussienne (normale) et celle de Student.

Copule normale : Le premier type de copules paramétriques que l'on présente ici est la copule gaussienne (ou normale) la copule de la distribution normale bivarée est donnée par :

$$C_\theta(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \exp\left(-\frac{x^2 - 2\theta xy + y^2}{2(1-\theta^2)}\right) dx dy.$$

où $\theta \in [-1, 1]$ et Φ^{-1} représente l'inverse de la distribution normale univariée.

Copule de Student : La copule de distribution de student donnée par :

$$C_\theta(u, v) = \int_{-\infty}^{T^{-1}(u)} \int_{-\infty}^{T^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \left(\frac{1 + x^2 - 2\theta xy + y^2}{v(1-\theta^2)} \right)^{-\frac{(v+2)}{2}} dx dy.$$

$\theta \in [-1, 1]$ et T^{-1} est la fonction inverse de la distribution standard de Student.

Copule de Gumbel-Morgenstern (FGM) : Le standard expression de cette famille est :

$$C_\theta(u, v) = uv + \theta uv(1-u)(1-v)$$

Où $u, v \in \mathbb{I}$ et $\theta \in [-1, 1]$.

Copule de Plackett : Le standard expression de cette famille est :

$$C_{\theta}(u, v) = \frac{1 + (\theta - 1)(u + v)}{2(\theta - 1)} - \frac{\sqrt{[1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)}.$$

Avec $u, v \in \mathbb{I}$ et $\theta \geq 0$.

Tau de kendall et rho de spearman :

Il existe deux relations utiles entre θ et tau de kendall τ et rho de spearman ρ_s .

La famille de Frank :

Le tau de kendall de la copule de Frank est donnée par la formule suivante :

$$\tau(\theta) = \frac{1 - 4D_1(-\ln \theta) - 1}{\ln \theta}.$$

Et le rho de spearman,

$$\rho_s(\theta) = 1 - 12 \frac{D_2(-\ln \theta) - D_1(-\ln \theta)}{\ln \theta}.$$

Où $D_k(x)$ est la fonction de Debye donnée par :

$$D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{\exp t - 1} dt, \forall k \geq 0.$$

La famille de Gumbel :

Le tau de kendall pour la famille de Gambel est définie par :

$$\tau(\theta) = \frac{\theta - 1}{\theta}.$$

La famille de Clayton :

Le tau de kendall pour la copule de Clayton est donnée par :

$$\tau(\theta) = \frac{\theta}{\theta + 2}.$$

La famille de ALI-MIKHAIL-HAQ (AMH) :

Le tau de kendall á cette famille donnée par la formule suivant :

$$\tau(\theta) = 1 - \frac{2(\theta + (1 - \theta)^2 \ln(1 - \theta))}{3\theta^2}.$$

La famille Archimax :

Pour une copule Archimax le tau de kendall est donnée par :

$$\tau(\phi, A) = \tau(A) + (1 - \tau(A))\tau(\phi).$$

Avec

$$\tau(A) = \int_0^1 t(1 - t) dA'(t).$$

Et

$$\tau(\phi) = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt.$$

La famille de copule elliptique :

Le tau de kendall de la copule elliptique est définie par :

$$\tau(\theta) = \frac{2}{\pi} \arcsin \theta.$$

Et le rho de spearman par :

$$\rho_s(\theta) = \frac{6}{\pi} \arcsin \frac{\theta}{2}.$$

Simulation avec logiciel R

```
### Spearman's rho and Kendall's tau for normal copulas
rho <- seq(-1, 1, by = 0.01) # correlation parameters of normal copulas
rho.s <- (6/pi) * asin(rho/2) # corresponding Spearman's rho
tau <- (2/pi) * asin(rho) # corresponding Kendall's tau
```

La famille de copule de Gumbel-Morgenstern (FGM) :

Le tau de kendall à cette famille donnée par la formule suivant :

$$\tau(\theta) = \frac{2\theta}{9}.$$

Et le rho de spearman par :

$$\rho_s(\theta) = \frac{\theta}{3}.$$

La famille de copule de Plackett :

Le rho de spearman de copule de Plackett s'écrit :

$$\rho_s(\theta) = \frac{\theta + 1}{\theta - 1} - \frac{2\theta \ln \theta}{(\theta - 1)^2}.$$

Conclusion

La copule est un concept de modélisation de la structure de dépendance. Contrairement aux différentes mesures de dépendance, la relation entre les variables est bien expliquée par la copule. Autrement dit, au lieu de résumer la structure de la dépendance par un seul nombre, par exemple le coefficient de corrélation, on peut utiliser la fonction copule qui décrit au mieux cette dépendance.

Dans ce mémoire on a abordé la notion des copules bivariée ainsi que les différentes propriétés qui lui sont associées. Au début, on a présenté quelques familles de copules paramétriques les plus utilisées, et nous avons établis de nouvelles mesures de dépendances : le tau de Kendall et le Rho de Spearman.

Bibliographie

- [1] Fadhila, L. (2011). Modélisation de la dépendance par les copules et applications. Mémoire de maîtrise, Université de Mouloud Mammeri, Tizi-Ouzou.
- [2] BENBRAIKA, G. DEPENDANCE DES RISQUES ET APPLICATION,(Doctoral dissertation, UNIVERSITÉ MOHAMED KHEIDER-BISKRA)
- [3] CANTONI, Eva, HUBER, Philippe, RONCHETTI, Elvezio. Maîtriser l'aléatoire Exercices résolus de probabilités et statistique. 2 éd. Paris : Springer, 2009. (Statistique et probabilités appliqués) <https://archive-ouverte.unige.ch/unige:22635>.
- [4] Toupin, M. H. (2008). Nouveau test d'adéquation pour les copules basé sur le processus de Spearman (Doctoral dissertation, Université du Québec à Trois-Rivières).
- [5] Fromont, M. Tests Statistiques, Année universitaire 2015-2016
- [6] Kadi, N. (2014). Estimation non-paramétrique de la distribution et densité de copules. Mémoire de maîtrise, Université de Sherbrooke
- [7] BENELMIR, I. (2018). Modélisation de la Dépendance par les Copules (Doctoral dissertation, UNIVERSITÉ MOHAMED KHEIDER-BISKRA).
- [8] Salvadori, G., & De Michele, C. (2007). On the use of copulas in hydrology : theory and practice. *Journal of Hydrologic Engineering*, 12(4), 369-380.

Résumé

Le but de ce travail est d'étudier les méthodes les plus importantes de simulation des données dépendent avec logiciel R.

Ce travail est composé de deux chapitres dans le premier chapitre on donne des notations générales que sera utilisé dans la suite, dans deuxième chapitre nous avons étudié la copule et les mesures d'association et chaque fois que je donne une application dans R.

Mot clés : Copule, Simulation, dépendent, concordantes, Tau de Kendall, Rho de Spearman.

تلخيص

الهدف من هذا العمل هو دراسة الطرق المهمة لمحاكاة المعطيات المرتبطة باستخدام البرنامج آر.
يتألف هذا العمل من فصلين في الفصل الأول نعطي مفاهيم عامة ستستخدم في ما يلي، في الفصل الثاني درسنا كوبيل و مقياس الارتباط وفي كل مرة أقدم فيها طلبا في آر.
الكلمات المفتاحية: كوبيل، المحاكاة، مرتبطة، متوافق، طو دو كاندل، رو دو سبيرمان.

Abstract

The purpose of this work is to study the most important methods of data simulation depend with R software.

this work is composed of two chapters in the first chapter we give general notation that will be used in the following chapter, in the second chapter we have studied the copula and the association measures and each time I give an application in R.

Keywords: Copula, Simulation depend, concordantes, Tau de Kendall, Rho de Spearman.

