

République Algérienne Démocratique et populaire **Ministère**
de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mouloud Mammeri de Tizi-Ouzou



Faculté de Génie Electrique et Informatique
Département Automatique

MEMOIRE DE MASTER

En Automatique

Option : Commande des systèmes

Présenté par :

ALLICHE Fadhila

BRAHIMI Kenza

Segmentation d'images couleurs basée sur les copules

Mémoire soutenu le 24/09/2017 devant le jury d'examen composé de :

M^{me} CHILALI Ouardia	M.A.A. à l'U.M.M.T.O.	Présidente
M^r HAMMOUCHE Kamal	M.A.A. à l'U.M.M.T.O.	Rapporteur
M^{me} DOORBANE Farida	M.A.A. à l'U.M.M.T.O.	Examineur
M^{me} BEDOUHANE Saida	M.C.A. à l'U.M.M.T.O.	Examineur

Remerciements

Nous tenons à remercier :

✚ *Le bon dieu qui nous a donné santé et la volonté durant notre travail.*

✚ *Nos parent, nos amis et tous ceux qui nous ont aidé.*

✚ *Notre promoteur HAMMOUCHE Kamal d'avoir accepté la charge de nous encadrer*

Les membres de jury qui fera l'honneur d'examiner notre travail.

✚ *Et tous les lecteurs.*

Dédicaces

Avec les sentiments d'amour et gratitude les plus sincères.

Je dédie ce modeste ce travail.

A mon père.

A ma très chère mère pour son amour, ses prière et ses sacrifices, que dieu les protège.

A mes frères et mes sœurs

A mes enseignants qui ont contribues à ma formation depuis mon jeune âge.

A mes très chères amies

A mon très cher binôme Fadhila

Kenza

LISTE DES FIGURES

1.1 Distribution Uniforme	8
1.2 Distribution Normale	9
1.3 Allure de quelques fonctions de densité continues	9
2.1 Exemple de copule Gaussienne	25
2.2 Exemple de copule Student.....	26
2.3 Fonction de densité d'une copule de Gumbel	28
2.4 Allure de la fonction de densité de copule de Clayton.....	30
2.3 Fonction de densité de copule de Frank.....	31
3.1 Images couleur tests.....	38
3.2 Image test1	38
3.3 Image test2	39
3.4 Image test3.....	39
3.5 Image test4	40
3.6 Images tests Segmenté	41

LISTE DES TABLEAUX

1.1 Lois de probabilité continues usuelles et leurs caractéristiques statistiques	10
3.2 Valeurs estimées des paramètres des copules	43
3.3 Taux d'erreur de classification.....	43

SOMMAIRE

Introduction générale

Chapitre 1 : Les Probabilités.

1.1 Introduction	1
1.2 Notion de probabilité	1
1.2.1 Ensemble fondamental	1
1.2.2 Probabilité.....	2
1.2.3 Probabilité conditionnelle.....	2
1.2.4 Indépendance en probabilité	3
1.2.5 Règle de Bayes	3
1.3 Variable aléatoire.....	4
1.3.1 Variable aléatoire discrète.....	4
1.3.2 Variable aléatoire continue	5
1.4 Moment d'une variable aléatoire.....	6
1.5 Lois usuelles	7
1.5.1 Lois usuelles discret.....	7
1.5.2 Lois usuelles continue	7
1.6 Estimation des paramètres d'une distribution	7
1.6.1 Méthode des moments	12
1.6.2 Méthode de maximum de vraisemblance.....	12
1.7 Distribution multivariées	13
1.7.1 Fonction de densité marginale	13
1.7.2 Fonction de densité conditionnelle.....	14
1.7.3 Exemple de densité de probabilité multivariée	15
1.8 Estimation d'une fonction de densité de probabilité	15
1.8.1 Estimation non paramétrique d'une fonction de densité de probabilité	16
1.8.2 Estimation non paramétrique d'une fonction de répartition	17
1.8.2 Estimation paramétrique d'une fonction de densité de probabilité	18
1.9 Conclusion	19

Chapitre 2 : Les Copules

2.1 Introduction	22
2.2 Définition et théorème d'existence.....	22
2.2.1 Définition	22
2.2.2 Théorème de Sklar	22
2.2.3 Borne de Fréchet	24
2.3 Propriétés d'une copule	25
2.4 Quelques copules	26
2.4.1 La copule Gaussienne	26
2.4.2 La copule Student	27
2.4.3 Copule Archimédienne	28
2.5 Conclusion.....	33

Chapitre 3 : Segmentation d'image couleur basé sur la classification Bayesienne

3.1 Introduction.....	34
3.2 Segmentation d'image couleur.....	34
3.2.1 Classification Bayesienne.....	34
3.2.2 Mélange Gaussiens.....	36
3.3 Méthode de segmentation adoptée pour la segmentation d'images couleur	37
3.4 Tests et résultats	38
3.5 Conclusion	39

Conclusion générale

Introduction générale

De nos jours l'analyse d'images est très utilisée dans plusieurs domaines très divers comme celui de la sécurité (reconnaissance des visages, d'empreintes, d'iris), en météorologie, en astronomie, en biologie, en médecine grâce à l'imagerie médicale (IRM de cellules), dans le tri postal (reconnaissance de caractères). En contrôle industriel, l'analyse d'images permet des automatisations (tri, détection des défauts, mesures, ...) sans manipulation directe sur les objets. D'une manière générale l'analyse d'image a pour but de reconnaître les éléments contenus dans l'image de façon objective, précise et souvent automatique. L'analyse d'images est effectuée en plusieurs étapes: acquisitions, prétraitement, segmentation, post-traitement. La segmentation est une étape critique car elle conditionne la qualité des mesures ou de l'interprétation du contenu de l'image. Elle permet d'isoler dans l'image les objets sur lesquels doit porter l'analyse ou séparer les régions d'intérêt du fond de l'image. Elle est également définie comme une opération qui a pour but de regrouper les pixels d'une image entre eux suivant des critères (couleur, texture) prédéfinis. Comme résultat de la segmentation, les pixels sont regroupés en régions homogènes, qui constituent un pavage ou une partition de l'image.

Contrairement aux images en niveaux de gris, une image couleur est de type multi spectrale, elle est composée de pixels ou chaque pixel est caractérisé par trois composantes au lieu d'une seule (niveau de gris). Les trois composantes les plus courantes, car issues des caméras couleur, sont le rouge (R), le vert (V) et le bleu (B). La segmentation d'images couleur tient en compte de toutes ces composantes couleur, elle est plus avantageuse par rapport à la segmentation d'images en niveaux de gris car des objets de couleurs différentes peuvent apparaître similaires dans une image en niveaux de gris.

Les méthodes de segmentation d'images couleurs peuvent être abordées selon plusieurs approches. L'approche basée sur la notion de classification considère une région comme un ensemble de pixels connexes appartenant à une même classe. Elles supposent donc que les pixels qui appartiennent à une même région possèdent des caractéristiques similaires et forment un nuage de points dans l'espace colorimétrique. La classification consiste alors à retrouver ces nuages de points qui correspondent aux classes des pixels présentes dans l'image. Cette opération peut être réalisée de deux manières: la première suppose l'existence de certains pixels dits prototypes dont l'appartenance aux classes est connue a priori. La seconde dite non supervisée (clustering), vise à regrouper automatiquement des pixels de l'image en classes sans aucune connaissance a priori sur l'appartenance des pixels aux classes.

Introduction générale

Dans ce mémoire nous nous sommes intéressées à la classification supervisée et plus exactement à la classification Bayésienne. Cette méthode fait appel à la règle de Bayes afin de classer les pixels les fonctions de densité de probabilité conditionnelle de chaque classe. Souvent ces fonctions sont décrites par une fonction multivariable Gaussienne. La méthode qui en découle est dite méthode de mélange Gaussienne, elle est considérée comme une méthode de référence dans le domaine de classification car elle minimise le risque d'erreur.

Or l'hypothèse selon laquelle chaque classe est décrite par une loi Gaussienne n'est pas toujours vérifiée. Pour remédier à ce problème, nous proposons dans ce mémoire de décrire une fonction de densité de probabilité conditionnelle multivariable de chaque classe ou d'une région de l'image par une composition d'une fonction copule et de fonctions de densité marginale relative à chaque composante couleur.

Les copules constituent un outil statistique assez récent permettant de modéliser la dépendance entre des variables aléatoires. Une copule est définie comme une fonction de répartition multivariable dont les lois marginales sont uniformes sur $[0,1]$. Elle offre un moyen efficace de relier une fonction de densité de probabilité conjointe de plusieurs variables à celles de ces marginales au lieu des fonctions de densité de probabilité conditionnelle.

Ce mémoire est principalement scindé en trois chapitres.

Dans le premier chapitre, nous présentons un rappel sur la théorie classique des lois de probabilités et des variables aléatoires continus et discrets. Quelques notions sur les lois de distributions mono variables, l'estimation de leurs paramètres, l'estimation de leur fonction de densité de probabilité et de répartition et leur généralisation au cas multidimensionnel sont également abordées.

Le deuxième chapitre est dédié à la notion de copules. Nous introduisons quelques définitions de base qui nous permettrons de décrire le concept de copule et ces principales propriétés, Quelques copules paramétriques les plus utilisées ainsi que leurs propriétés sont également décrites.

Dans le troisième chapitre, nous rappellerons très brièvement les différentes approches de la segmentation d'images, puis nous montrons comment utiliser le concept de copule pour segmenter une image couleur. Nous terminons le chapitre par la présentation des résultats de la segmentation de quelques images couleur par la méthode adoptée. Une conclusion générale clôturera le mémoire.

1.1 Introduction

Il est toujours possible d'associer à une variable aléatoire une probabilité et définir ainsi une loi de probabilité. Lorsque le nombre d'épreuves augmente indéfiniment, les fréquences observées pour le phénomène étudié tendent vers les probabilités et les distributions observées vers les distributions de probabilité ou loi de probabilité.

L'objectif de ce chapitre est de rappeler quelques notions probabilistes utilisées en statistique mathématique. Il présente certaines distributions des variables aléatoires continus et discrets.

1.2 Notion de probabilité

1.2.1 Ensemble fondamental

Avant toute formalisation, le résultat d'une expérience aléatoire s'appelle événement. La quantification des « chances » qu'un tel événement a de se réaliser correspond à la notion intuitive de probabilité. Pour effectuer cette quantification, il est nécessaire de décrire au préalable l'ensemble des résultats possibles, appelés événements élémentaires. Cet ensemble expérimental s'appelle ensemble fondamental (ou univers), il est noté Ω .

Exemple 1

Jet d'un dé à six faces numérotées : $\Omega = \{1,2,3,4,5,6\}$.

Chaque élément $\omega \in \Omega$ représente donc un événement élémentaire, et toute partie $A \subset \Omega$ (ou $A \in \mathcal{P}(\Omega)$) sera un événement. Parfois on dit que Ω est l'ensemble des éventualités possibles et les événements élémentaire sont alors les singletons, c'est-à-dire les ensembles réduits à un seul élément $\{\omega\}$, qui sont effectivement en toute rigueur des événements, puisque appartenant à $\mathcal{P}(\Omega)$, ce qui n'est pas le cas du point ω .

Le couple $(\Omega, \mathcal{P}(\Omega))$ s'appelle un espace probabilisable.

1.2.2 Probabilité

Une fois défini l'ensemble des événements auxquels on s'intéresse, on va essayer de traduire par un nombre leurs « possibilités » de réalisation. Cela revient à affecter une mesure de « croyance » à chaque événement, c'est-à-dire un degré de certitude que l'événement se produise ou non. Une probabilité est un nombre associé à un événement, compris entre 0 et 1,

pouvant se convertir en pourcentage de « chance ». L'événement certain se voit attribuer la probabilité 1 et l'événement impossible la probabilité 0. D'autre part, si deux événements sont incompatibles, c'est-à-dire ne pouvant pas se réaliser simultanément, la probabilité de réalisation de l'un des deux sera la somme de leurs probabilités respectives.

Définition

On appelle probabilité P sur (Ω, \mathcal{A}) une application $P: \mathcal{A} \rightarrow [0,1]$ telle que :

1. $P(\Omega) = 1$;
2. Pour toute suite A_n d'événements incompatibles, $A_n \cap A_m = \emptyset$ pour $m \neq n$:

$$P\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} P(A_n) \quad (1.1)$$

Propriété dite de σ -additivité.

Le triplet (Ω, \mathcal{A}, P) est appelé espace probabilisé.

Par exemple, dans un jet de dé, $P(1 \text{ ou } 2) = P(1) + P(2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

1.2.3 Probabilités conditionnelles

On considère l'espace probabilisé (Ω, \mathcal{A}, P) et un événement particulier B de \mathcal{A} tel que $P(B) > 0$. La connaissance de la réalisation de B modifie la probabilité de réalisation d'un événement élémentaire, puisque l'ensemble des résultats possibles est devenu B et non plus Ω .

Cette nouvelle probabilité, notée $P(\cdot|B)$, est définie sur la tribu conditionnelle $\mathcal{A}|B = \{A \cap B, A \cap B^c\}$ par :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.2)$$

Exemple

A l'issue d'un jet de dé on sait que le résultat est supérieur à trois et on s'intéresse à l'événement $A = \{\text{obtenir une face paire}\}$. Initialement on avait $P(A) = \frac{3}{6} = \frac{1}{2}$; maintenant $B = \{4,5,6\}$ est devenu Ω et $P(A|B) = \frac{2}{3} > \frac{1}{2}$.

1.2.4 Indépendance en probabilité

Deux événements A et B sont dits indépendants, relativement à la probabilité P , si :

$$P(A \cap B) = P(A)P(B) \quad (1.3)$$

La probabilité de réalisation simultanée de deux événements indépendants est égale au produit des probabilités que chacun de ces événements se produise séparément. En conséquence, si $P(B) > 0$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A) \quad (1.4)$$

La réalisation d'un événement ne modifie pas la probabilité de réalisation de l'autre.

1.2.5 Règle de Bayes

Soit (A_1, \dots, A_n) est un système complet d'événements dans \mathcal{B} tel que $P(A_k) > 0$ pour tout k compris entre 1 et n , on a alors pour tout événement $B \in \mathcal{B}$ de probabilité non nulle et tout entier j compris entre 1 et n :

$$P(A_j|B) = \frac{P(B|A_j) P(A_j)}{\sum_{k=1}^n P(B|A_k) P(A_k)} \quad (1.5)$$

Ainsi, pour tout j dans I on a

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j) P(A_j)}{\sum_{k=1}^n P(B|A_k) P(A_k)}$$

1.3 Variable aléatoire

Une des notions fondamentales des statistiques est celle de variable aléatoire. On considère un ensemble d'individus qui sera appelée Ω . Un individu de cet ensemble sera noté ω . On note $X(\omega)$ une caractéristique de l'individu ω . La quantité $X(\cdot)$ est appelée variable aléatoire. Les valeurs possibles que peut prendre $X(\omega)$ quand $\omega \in \Omega$ détermine la nature de la variable aléatoire.

1.3.1 Variable aléatoire discrète

Une variable aléatoire X ayant des valeurs dans un ensemble E fini ou dénombrable est appelée variable aléatoire discrète (v.a.d).

1.3.1.1 Loi de probabilité

On appelle distribution ou loi de probabilité de la v.a.d X l'ensemble des couples (x_i, p_i) avec $i \in \mathbb{N}$.

La loi de probabilité P_X de X est alors définie par les probabilités individuelles :

$$p_i = P_X(X = x_i) = P\{X^{-1}(x_i)\}, i \in \mathbb{N}$$

1.3.1.2 Fonction de répartition

On appelle fonction de répartition de la v.a.d X , la fonction F définie pour x réel par :

$$F(x) = P_X\{X < x\} = P\{\omega \in \Omega \mid X(\omega) < x\}$$

C'est une fonction en escalier, constante par morceaux continue à gauche, défini par :

$$F(x) = \sum_{\{x_i \mid x_i < x\}} p_i \quad (1.6)$$

C'est-à-dire que c'est la somme des poids de tous les points qui sont strictement à gauche de x .

1.3.2 Variable aléatoire continue

On dit que X est une variable aléatoire continue (v.a.c) si elle prend des valeurs dans \mathbb{R} .

Exemple

La durée de vie d'une lampe ou le salaire d'un individu tiré au sort dans une population sont représentés par des v.a. continues.

1.3.2.1 Fonction de répartition

La fonction de répartition $F(x)$ d'une v.a.c X est définie par:

$$F(x) = P_X\{X < x\}$$

- Elle est croissante au sens large

En effet, pour $x < y$ on a $\{X < x\} \subset \{X < y\}$ et par conséquent $F(x) = P_X(X < x) \leq P_X(X < y) = F(y)$.

- Elle prend ses valeurs entre 0 et 1 :

$$0 \leq F(x) \leq 1 \text{ avec } \lim_{x \rightarrow -\infty} F(x) = 0 \text{ et } \lim_{x \rightarrow +\infty} F(x) = 1$$

- La probabilité de l'intervalle $[a, b[$, pour $a < b$, se calcul par :

$$P_X(a \leq X < b) = F(b) - F(a)$$

- Elle est continue à gauche :

$$\lim_{h \rightarrow 0^+} F(x-h) = F(x)$$

1.3.2.2 Fonction de densité de probabilité

La fonction dérivée $f(x)$ de la fonction de répartition $F(x)$ est dite fonction de densité de probabilité de X et vérifie les relations :

$$x \in \mathbb{R}, f(x) = F'(x) \text{ et } F(x) = \int_{-\infty}^x f(t) dt$$

- Pour tout nombre réel on a :

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

- Pour tout réel α et β on a :

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

- $x \in \mathbb{R}, f(x) \geq 0$ (f est positive).
- f est continue sur \mathbb{R} sauf peut-être en un nombre fini de points ou elle admet une limite finie à gauche et une limite à droite.
- L'intégrale $\int_{-\infty}^{+\infty} f(x) dx$ est convergente et on a $\int_{-\infty}^{+\infty} f(x) dx = 1$.

1.4 Moments d'une variable aléatoire

1.4.1 Espérance mathématique

On appelle espérance mathématique de la v.a. X la quantité, si elle existe, définie par le moment d'ordre $k \in \mathbb{N}$

$$m_k = E(X)^k.$$

$$m_k = \sum x_i^k p_i \text{ pour une v.a. discrète.}$$

$$m_k = \int_{-\infty}^{+\infty} x^k f(x) dx \text{ pour une v.a. continue.}$$

1.4.2 La moyenne

Le moment d'ordre 1 ($k = 1$) définit la moyenne de la v.a. X , $m_1 = E(X)$.

Il s'agit d'une moyenne en probabilité, ou moyenne pondérée, des valeurs x_i que peut prendre la v.a. X . Notons que si X prend ses valeurs entre x_1 et x_n , on aura bien sur $x_1 \leq E(X) \leq x_n$.

Propriétés

- $E(aX + bY) = aE(X) + bE(Y)$.

- $E(X + Y) = E(X) + E(Y)$.
- Si $X \geq 0$, alors $E[X] \geq 0$
- Si $X \leq Y$, alors $E[X] \leq E[Y]$

1.4.3 Moments centrés

Le moment centré d'ordre $k \in \mathbb{N}$ d'une v.a X est définie par l'espérance mathématique de la variance centrée $(X - E[X])$ puissance k .

$$m_k^c = E[(X - E[X])^k] \quad (1.7)$$

1.4.3 La variance

Il s'agit d'un indicateur mesurant la dispersion des valeurs x_q que peut prendre la v.a. X , autour de la moyenne $E(X)$. Elle est définie par l'espérance mathématique du carré de la v.a. centrée $(X - E(X))$.

$$V(X) = E[(X - E(X))^2] \quad (1.8)$$

$$V(X) = \sum_{i \in \mathbb{N}} p_i (x_i - E[X])^2 \text{ pour une v.a discrète}$$

$$V(X) = \int (x - E[X])^2 f(x) dx \text{ pour une v.a continue}$$

Notons que l'écart type correspond à la racine carrée de la variance : $\sigma(X) = \sqrt{\text{Var}(X)}$

1.5 Covariance

La covariance permet d'estimer la dépendance entre deux variables aléatoires. Elle est définie par:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) \quad (1.9)$$

L'espérance $E(XY)$ est calculée à partir de la loi jointe du couple de variables (X, Y) .

1.6 Loïs usuelles

Il existe un grand nombre de lois statistiques. Les plus connues sont:

1.6.1 Loïs usuelles discrètes

a- Loi de Bernoulli

On appelle variable de Bernoulli ou variable indicatrice, la variable aléatoire X telle que :

$$X: \Omega \rightarrow \mathbb{R}$$

$$X: \Omega \rightarrow \{0,1\}$$

La loi de probabilité associée à la variable de Bernoulli X telle que,

$$\begin{aligned} P(X = 0) &= q \\ P(X = 1) &= p \quad \text{avec } p + q = 1 \end{aligned}$$

Est appelée loi de Bernoulli notée $\mathcal{B}(1, p)$

b- Loi de Binomiale

La variable Binomiale, représente le nombre de succès obtenus lors de la répétition de n épreuves identiques et indépendantes, chaque épreuve ne pouvant donner que deux résultats possibles. Est la loi binomiale de paramètres n et p . $\mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}$,

$$0 \leq k \leq n.$$

Ainsi la loi de probabilité suivie par la somme de n variables de Bernoulli ou la probabilité associée au succès est p , est la loi Binomiale de paramètres n et p .

$$S_n = \sum_{i=1}^n X_i \quad \mathbb{R}^n$$

$$S_n = \sum_{i=1}^n X_i \quad \mathcal{B}(1, p) \quad (1.10)$$

c- Loi de Poisson

Une variable aléatoire X a valeurs dans \mathbb{R} suit une loi de Poisson de paramètre λ ($\lambda > 0$) si les réels p_k sont donnés par

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (1.11)$$

On note : $X \sim \mathcal{P}(\lambda)$

6.2 Lois usuelles continues

Il existe un grand nombre de lois continues. Les plus importantes sont la loi uniforme et la loi normale.

a- Loi uniforme

La loi uniforme est la loi exacte de phénomènes continus uniformément répartis sur un intervalle. La variable aléatoire X suit une loi uniforme sur le segment $[a, b]$ avec $a < b$ si sa densité de probabilité est donnée par :

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon} \end{cases} \quad (1.12)$$

L'allure de cette fonction est donnée par la figure (1.1).

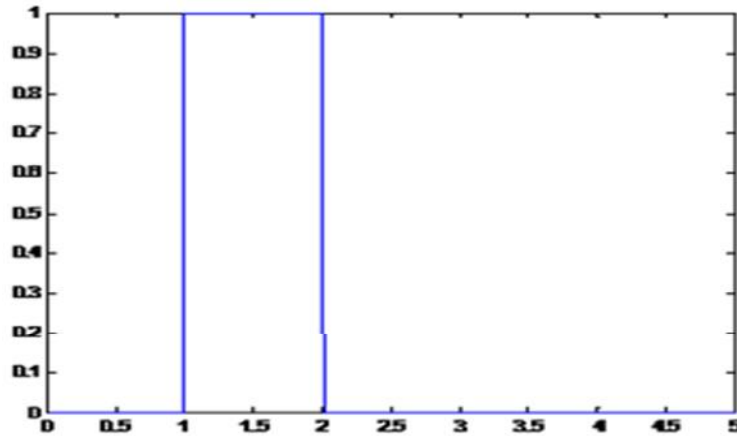


Figure (1.1) : Distribution Uniforme

Sa fonction de répartition est donnée par :

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}$$

Théorème

- 1- Soit U une v.a de loi uniforme sur $[0,1]$, alors $F^{-1}(U)$ a pour loi F .
- 2- Si une v.a X ayant pour loi F , alors $U = F(X)$ suit la loi uniforme sur $[0,1]$.

b- Loi normale ou de Gauss-Laplace

La loi normale joue un rôle particulièrement important dans la théorie des probabilités et dans les applications pratiques. La particularité fondamentale de la loi normale la distinguant des autres lois pour des conditions se rencontrant fréquemment en pratique.

La v.a.c. X suit une loi normale de paramètre $\mu \in \mathbb{R}$ et $\sigma > 0$, noté par $N(\mu, \sigma)$ si sa fonction densité est donnée par :

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], x \in \mathbb{R} \quad (1.13)$$

- Le paramètre μ représente la moyenne, il détermine la position de la courbe, l'axe $x = \mu$ étant un axe de symétrie
- Le paramètre σ désigne l'écart type, il détermine l'échelle (la position des valeurs
-
- autour de la moyenne).

La figure (1.2) montre l'allure de cette fonction pour différentes valeurs des paramètres μ et σ .

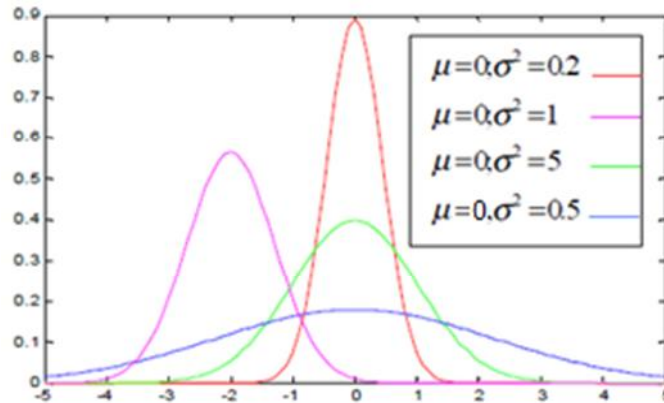


Figure (1.2) : Distribution Normale

Remarque : Il existe bien d'autres distributions continues comme la loi bêta, gamma, exponentielle, log-normale, Khi 2, Weibull., La figure 1.3 montre l'allure de ces distributions.

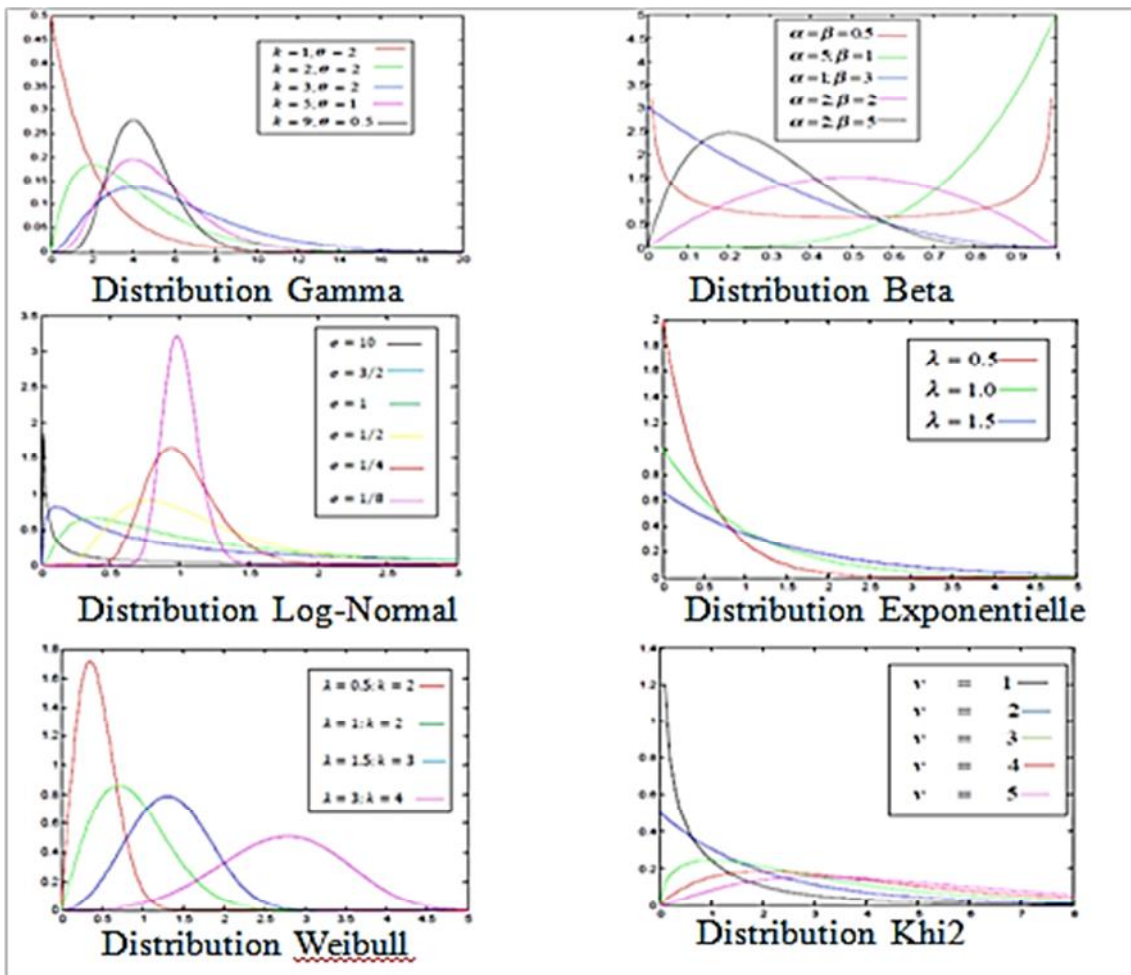


Figure (1.3) : Allure de quelques fonctions de densité continues

Les fonctions de densité de probabilité ainsi que les différentes caractéristiques statistiques principales (moyenne et variance) de quelque distribution citée précédemment sont résumés dans le tableau (2.1).

Distribution	Fonction de densité de probabilité	Fonction de répartition	Moyenne $E(X)$	Variance $V(X)$
Uniforme $U(a, b)$	$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{si } \text{non} \end{cases}$	$F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}$	$\frac{a+b}{2}$	$\frac{(a-b)^2}{12}$
Normale $N(u, \sigma^2)$	$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-u}{\sigma}\right)^2\right], x \in \mathbb{R}$	$F(x) = \frac{\theta^k}{\Gamma(k)} \int_0^x u^{k-1} e^{-\theta u} du$	u	σ^2
Gamma $\gamma(\theta, k)$	$f(x) = \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x}, x > 0$ Ou $\Gamma(k) = \int_0^{+\infty} e^{-x} x^{k-1} dx$	$F(x) = \begin{cases} 0, & \text{pour } x < 0 \\ \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) + \Gamma(\beta)} \int_0^x u^{\alpha-1} (1-u)^{\beta-1} du, & 0 \\ 1, & \text{pour } x > 1 \end{cases}$	$\frac{k}{\theta}$	$\frac{k}{\theta^2}$
Beta $\beta(\alpha, \beta)$	$f(x) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)}, x \in \mathbb{R}$ Ou $B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt, \alpha > 0, \beta > 0$	$F(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left[\frac{\ln(x) - \mu}{\sigma \sqrt{2}}\right]$	$\frac{\alpha}{\alpha + \beta}$	$\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$
Lognormal $L(u, \sigma)$	$f(x) = \frac{1}{x\pi} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x - u}{\sigma}\right)^2\right\}, x > 0$		$\left(u + \frac{\sigma^2}{2}\right)$	$[e^{\sigma^2} - 1]e^{2u + \sigma^2}$
Exponentielle $E(\lambda)$	$f(x) = \lambda e^{-\lambda x}$	$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Khi2 $\chi^2(v)$	$f(x; v) = \frac{1}{2^{v/2} \Gamma\left(\frac{v}{2}\right)} x^{\frac{v}{2}-1} e^{-\frac{x}{2}}, x > 0$		v	$2v$
Weibull $W(k, \lambda)$	$f(x; k, \lambda) = (k/\lambda)(x/\lambda)^{k-1} e^{-(x/\lambda)^k}$		$\lambda \left(1 + \frac{1}{k}\right)$	$\lambda^2 \left(1 + \frac{2}{k}\right) - E^2(X)$
Rayleigh $R(\sigma)$	$f(x; \sigma) = \frac{x}{\sigma^2} \exp\left[-\frac{x^2}{2\sigma^2}\right] \text{ avec } x > 0$	$F(x) = 1 - \left(\frac{x^2}{2\sigma^2}\right)$	$\sigma \sqrt{\frac{\pi}{2}}$	$\frac{4 - \pi}{2} \sigma^2$

Tableau 1.1 : Loïs de probabilité continues usuelles et leurs caractéristiques statistiques.

1.7 Estimation des paramètres d'une distribution

Soit $X = \{X_1, \dots, X_N\}$ un ensemble échantillons décrivant une distribution $f(x)$ de la v.a X , définie par un ensemble de paramètres $\theta = \{\theta_1, \theta_2, \dots, \theta_m\}$.

L'ensemble de ces paramètres peuvent être estimés à partir des échantillons de X suivant deux méthodes.

1.7.1 Méthode de maximum de vraisemblance

L'estimation du maximum de vraisemblance est une méthode de statistique couramment utilisée pour inférer les paramètres de la distribution de probabilité d'un échantillon donnée.

La fonction de vraisemblance, notée $L(X_1, \dots, X_n; \theta)$, est une fonction de probabilité conditionnelle qui décrit le paramètre θ d'une loi statistique en fonction des valeurs x_i supposées connues. Elle s'exprime à partir de la fonction de densité de probabilité conditionnelle $f(x|\theta)$ par :

$$L(X_1, \dots, X_n; \theta) = \prod_{i=1}^n f(x_i; \theta) \quad (1.14)$$

Cette formule n'est valable que si on suppose que les X_i sont indépendants entre eux.

L'estimation du vecteur paramètre θ revient à maximiser la fonction de vraisemblance pour que les probabilités des réalisations observées soient aussi maximales. Ceci constitue un problème d'optimisation dont la solution est celle du système suivant :

$$(I) \left\{ \begin{array}{l} \frac{\partial L}{\partial \theta} = 0 \\ \left(\frac{\partial^2 L}{\partial \theta^2} \right)_{\theta = \hat{\theta}} < 0 \end{array} \right.$$

A cause de la forme particulière des densités de probabilité des distributions usuelles de probabilité, il est plus aisé d'utiliser le logarithme de la vraisemblance, $\log L(X_1, \dots, X_n, \theta)$, si $f(x, \theta) > 0$, x , et θ .

Le système (I) est donc équivalent au système (II) suivant :

$$(II) \left\{ \begin{array}{l} \frac{\partial \log L}{\partial \theta} = 0 \\ \left(\frac{\partial^2 \log L}{\partial \theta^2} \right)_{\theta = \hat{\theta}} < 0 \end{array} \right.$$

La première équation du système (II), $\frac{\partial \log L}{\partial \theta} = 0$, s'appelle équation de vraisemblance.

Pour un échantillon indépendant X , l'équation de vraisemblance s'écrit :

$$\sum_{i=1}^N \frac{\partial \log f(x_i, \theta)}{\partial \theta} = 0$$

La résolution de cette équation par rapport à chaque paramètre $\theta_j (j = 1, \dots, M)$ permet d'aboutir à la solution $\hat{\theta}$.

1.7.2 Méthode des moments

Cette procédure d'estimation repose sur la propriété de convergence presque sûre des moments empiriques d'un échantillon $X_N = (X_1, \dots, X_n)$, extrait de X , vers les moments théoriques correspondants de X .

On appelle moment empirique, non centré d'ordre k la quantité suivante :

$$m_k = \frac{1}{n} \sum_{i=1}^n X_i^k$$

Le moment empirique, centré, d'ordre k est défini par :

$$m_k^c = \frac{1}{n} \sum_{i=1}^n (X_i - m_1)^k$$

Où m_1 est la moyenne empirique de l'échantillon.

Soit le vecteur paramètre $\theta = (\theta_1, \dots, \theta_m)$ à estimer : on note par $m_k(\theta)$ le moment théorique d'ordre k de X , qu'il soit centré ou non, et par m_k le moment empirique d'ordre k .

La méthode des moments permet d'estimer les paramètres $\theta_j (j = 1, \dots, m)$ en résolvant le système d'équation suivant :

$$\begin{cases} m_1(\theta) = m_1 \\ \vdots \\ m_m(\theta) = m_m \end{cases}$$

1.8 Distribution multivariées

Le concept de fonction de densité de probabilité et fonction de répartition peut être facilement généralisé à deux ou plusieurs variables.

Soient X_1, X_2, \dots, X_N , un ensemble de N variables réelles continues. Ces variables sont mutuellement indépendantes si pour tous intervalles A_1, \dots, A_N de \mathbb{R} on a :

$$P(X_1 \in A_1, \dots, X_N \in A_N) = \prod_{i=1}^N P(X_i \in A_i) \quad (1.15)$$

La loi de probabilité de X_1, X_2, \dots, X_N est déterminée par la fonction de répartition conjointe $F(x_1, x_2, \dots, x_N)$, définie par :

$$F(x_1, x_2, \dots, x_N) = \Pr\left[\bigcap_{j=1}^N (X_j \leq x_j)\right] \quad (1.16)$$

Si F est N fois dérivable par rapport aux variables, alors la loi de (X_1, X_2, \dots, X_N) absolument continue, de densité conjointe $f(x_1, x_2, \dots, x_N)$ définie par :

$$f(x_1, x_2, \dots, x_N) = \frac{\partial^N F(x_1, x_2, \dots, x_N)}{\partial x_1 \partial x_2 \dots \partial x_N} \quad (1.17)$$

$f(x_1, x_2, \dots, x_N)$ doit satisfaire les conditions suivantes :

- 1- $f(x_1, x_2, \dots, x_N) \geq 0$
- 2- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = 1$

La fonction de répartition se calcule alors par intégrations :

$$F(x_1, x_2, \dots, x_N) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_N} f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N \quad (1.18)$$

1.8.1 Fonction de densité marginale

Les fonctions de densité marginales de X_1, X_2, \dots, X_N sont des fonctions de densité monovariée, définies à partir de la fonction de densité conjointe par :

$$f_1(x_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_N) dx_2 dx_3 \dots dx_N$$

$$f_2(x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_N) dx_1 dx_3 \dots dx_N$$

$$f_N(x_N) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_N) dx_1 \dots dx_{N-1}$$

1.8.2 Fonction de densité conditionnelle

La fonction de densité de probabilité conjointe peut être également définie par les fonction de densité conditionnelles :

$$f(x_1, x_2, \dots, x_n, \dots, x_N) = f(x_1) f(x_2|x_1) f(x_3|x_1 x_2) \dots f(x_N|x_1, x_2, \dots, x_{n-1}, \dots, x_{N-1}) \quad (1.19)$$

$f(x_i|x_j)$ est la fonction de densité de probabilité conditionnelle de la variable X_i sachant la variable X_j .

Elle est définie pour un sous ensemble de r variables sachant les $(N - r)$ variables restantes par :

$$f(x_1, x_2, \dots, x_r | x_{r+1}, x_{r+2}, \dots, x_N) = \begin{cases} \frac{f(x_1, x_2, \dots, x_r)}{f(x_{r+1}, x_{r+2}, \dots, x_N)} & \text{if } f(x_{r+1}, x_{r+2}, \dots, x_N) > 0 \\ 0 & \text{if } f(x_{r+1}, x_{r+2}, \dots, x_N) = 0 \end{cases}$$

Si les variables X_1, X_2, \dots, X_N sont mutuellement indépendantes, alors :

$$f(x_1, x_2, \dots, x_N) = \prod_{j=1}^N f_j(x_j) \quad (1.20)$$

1.8.3 Exemples de fonctions de densité de probabilité multivariable

a) La loi normale multidimensionnelle

La fonction de densité de probabilité de la loi normale multidimensionnelle est :

$$f_X(X) = f_X(x_1, x_2, \dots, x_N) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T \Sigma^{-1} (X - \mu) \right\} \quad (1.21)$$

μ de dimension N est le vecteur moyen et Σ de dimension $(N \times N)$ est la matrice variance-covariance.

Pour $N = 2$, la densité de la loi normale de moyennes μ_1, μ_2 , de variance σ_1^2, σ_2^2 et de coefficient de corrélation ρ s'écrit :

$$f_{X_1, X_2}(x_1, x_2) = \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right)^{-1} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\}$$

b- La loi Student multidimensionnelle

Un vecteur aléatoire X de \mathbb{R}^N suit une loi de Student de paramètres (v, μ, Σ) s'il admet pour fonction de densité

$$t_d(X; v, \mu, \Sigma) = \frac{\Gamma\left(\frac{v+d}{2}\right)(\pi v)^{-d/2} |\Sigma|^{-1/2}}{\Gamma(v/2)[1+(X-\mu)^T \Sigma^{-1}(X-\mu)/v]^{v+d/2}} \quad (1.22)$$

Où v de \mathbb{R}^N ($v > 0$), μ et Σ désignent le degré de liberté, le vecteur moyen, une matrice inversible de dimension $N \times N$ et la fonction Gamma d'Euler, respectivement.

1.9 Estimation d'une fonction de densité de probabilité

L'estimation d'une fonction densité f ou d'une fonction de répartition F à partir d'un échantillon de variables aléatoires réelles X_1, X_2, \dots, X_n indépendantes est une procédure très importante dans la modélisation des processus. Elle occupe une place importante dans l'étude de nombreux phénomènes de nature aléatoire. Par exemple, en médecine, on cherche à étudier l'assimilation d'un traitement antibiotique administré par voie orale. Pour cela, on mesure, pour chaque patient $i = 1, \dots, n$, la concentration x_i de l'antibiotique qui est passée dans le sang du patient après 5 heures (temps moyen de digestion). On modélise alors le phénomène de la manière suivante : x_1, x_2, \dots, x_n sont les réalisations de n variables aléatoires indépendantes X_1, X_2, \dots, X_n ayant même densité f . Dans ce contexte médical, comprendre le processus d'assimilation de l'antibiotique dans le sang revient à connaître f .

Il existe deux approches très complémentaires pour réaliser ces estimations : l'approche paramétrique et l'approche non paramétrique. L'approche paramétrique suppose que les données sont issues d'une loi de probabilité de forme connue dont seuls les paramètres sont inconnus. Dans ce cas, l'estimation de densité se résume à l'estimation des paramètres de la distribution. Dans l'approche non paramétrique, la loi de probabilité est inconnue, il s'agit

alors de trouver la forme de cette loi sans à priori. L'estimation non paramétrique d'une fdp permet également de réaliser une estimation paramétrique de la fdp (tests d'adéquation).

1.9.1 Estimation non paramétrique d'une fonction d'une fonction de densité de probabilité

Soient x_1, \dots, x_n sont les réalisations de n variables aléatoires indépendantes X_1, X_2, \dots, X_n ayant la même fonction de densité f , inconnue. Les n variables X_i sont supposées indépendantes et identiquement distribuées. Il existe plusieurs méthodes d'estimation non-paramétrique d'une densité. Les plus connues sont la méthode des histogrammes et la méthode des noyaux.

a)- Méthode des histogrammes

C'est la méthode la plus simple et la plus ancienne. Elle consiste à partitionner l'intervalle de référence $I = [x_{min}, x_{max}]$ en $p \in \mathbb{N}$ cellules $A_k, k \in \{1, \dots, p\}$, et à compter le nombre d'observations appartenant à chaque cellule A_k . Si toutes les cellules de l'histogramme ont la même largeur, on dit que l'histogramme est régulier. On note $h \in \mathbb{R}^+$, la largeur des cellules qui est alors appelée le pas ou la granularité de l'histogramme. La valeur de h est donnée par :

$$h = \frac{x_{min} - x_{max}}{p}$$

Les cellules $A_k, k \in \{1, \dots, p\}$ forment une partition de I (c'est-à-dire $I = \bigcup_{k=1}^p A_k$ $\forall i, j \in \mathbb{N}: i \neq j, A_i \cap A_j = \emptyset$). Chaque cellule A_k de centre a_k , est définie par:

$$A_k = \left[a_k - \frac{h}{2}, a_k + \frac{h}{2} \right] \forall k \in \{1, \dots, p-1\} \text{ avec } a_{k+1} = a_k + h$$

Le nombre d'observations appartenant à chaque cellule A_k est, noté Acc_k . est appelé accumulateur de la cellule A_k . Il est défini par :

$$Acc_k = \sum_{i=1}^n \mathbb{1}_{A_k}(x_i), \tag{1.23}$$

Où $\mathbb{1}_{A_k}$ est la fonction caractéristique de l'ensemble A_k telle que :

$$\mathbb{1}_{A_k}(x) = \begin{cases} 1 & \text{si } x \in A_k, \\ 0 & \text{sinon.} \end{cases}$$

En émettant l'hypothèse que les observation se répartissent uniformément dans la cellule A_k , on peut alors construire un estimateur de la densité f . Pour tout $x \in I$, par :

$$\hat{f}_\Delta^n(x) = \frac{1}{h} \sum_{k=1}^p P(A_k) \mathbb{1}_{A_k}(x). \tag{1.24}$$

avec $P(A_k) = \frac{Acc_k}{n}$.

Cet estimateur peut aussi s'écrire

$$\hat{f}_\Delta^n(x) = \frac{1}{n} \sum_{k=1}^p A_{cc_k} A_k(x).$$

b)- Méthode des noyaux

C'est une méthode très populaire, elle a été introduite par Rosenblatt en 1956 puis par Parzen en 1962. Pour estimer une densité f sur \mathbb{R} , un estimateur à noyau (kernel en anglais) est donné, pour $h > 0$ et $x \in \mathbb{R}$, par :

$$\hat{f}_h(x) = \frac{1}{Nh} \sum_{j=1}^N k\left(\frac{x-X_j}{h}\right) \tag{1.25}$$

Où la fonction $k : \mathbb{R} \rightarrow \mathbb{R}$ est appelée noyau. Elle doit être intégrable telle que $\int_{\mathbb{R}} k(x) dx = 1$, et être positive pour que l'estimateur soit une densité. Les noyaux usuels sont :

- le noyau gaussien $k(x) = (1/\sqrt{2\pi})exp(-x^2/2)$.
- le noyau d'Epanechnikov $k(x) = (3/4)(1 - x^2)$ $-1 \leq x \leq 1$.
- Le noyau triangulaire $k(x) = (1 - |x|)$ $|x| \leq 1$.

Le paramètre $h > 0$ est appelée largeur de noyau (bandwidth) ou paramètre de lissage.

1.9.2 Estimation non paramétrique d'une fonction de répartition

Soient $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ les variables ordonnées des variables X_1, X_2, \dots, X_n .

L'estimation de la fonction de répartition empirique, notée F_n , est définie par:

$$\begin{aligned} F_n(x) &= \frac{\text{nombre d'observations } \leq x}{n} \\ &= \frac{\#\{i = X_i \leq x\}}{n} \\ &= \frac{1}{n} \sum_{i=1}^n I\{X_i \leq x\} \\ &= \begin{cases} 0 & \text{si } x < X_{(1)} \\ \frac{k}{n} & \text{si } X_{(k)} \leq x < X_{(k+1)} \quad k = 1, \dots, n-1 \\ 1 & \text{si } x \geq X_{(n)} \end{cases} \end{aligned}$$

On peut également utiliser la méthode des noyaux pour estimer une fonction de répartition F .

Cet estimateur est défini par :

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) \quad (1.25)$$

où

$$K(x) = \int_{-\infty}^x k(y) dy.$$

1.9.3 Estimation paramétrique d'une fonction de densité de probabilité

Les méthodes d'estimation paramétrique d'une fonction de densité de probabilité suppose que les données sont issues d'une loi de probabilité de forme connue dont seuls les paramètres sont inconnus. Dans ce cas, on utilise des tests d'adéquation.

Les tests d'adéquation ne permettent pas de trouver la loi d'une variable aléatoire, mais seulement de vérifier si une variable aléatoire donnée suit une loi de distribution connue à priori) dite théorique. Elle consiste à appliquer un test d'adéquation permettant d'accepter ou de rejeter une hypothèse simple émise a priori, généralement, notée H_0 ou H_1 telle que :

H_0 : la distribution de probabilité théorique a engendrée l'échantillon.

H_1 : la distribution de probabilité théorique n'a pas engendré l'échantillon.

L'estimation paramétrique consiste à comparer la distribution (fonction de répartition ou fonction de densité) empirique, estimée par l'approche non paramétrique présentée précédemment, à une distribution théorique dont les paramètres sont estimés par l'une des méthodes citées dans le paragraphe précédent. Parmi ces méthodes, on peut citer le test de Kolmogorov qui consiste à mesurer la distance entre la distribution empirique F_n et la distribution théorique F tel que:

$$K_n = \sup_x |F_n(X) - F(X)| \quad (1.26)$$

1.10 Conclusion

Dans ce chapitre nous avons donné un aperçu sur la théorie classique des lois de probabilités et des variables aléatoires continus et discrets. Nous avons également présenté quelques notions sur les lois de distributions monovariées, l'estimation de leurs paramètres, l'estimation de leurs fonctions de densité de probabilité et de répartition et leurs généralisations au cas multidimensionnel. Ainsi, nous pouvons noter à travers ce chapitre l'importance des lois de distributions dans la description de phénomènes aléatoires et le lien entre la fonction de densité de probabilité conjointe de plusieurs variables avec les lois conditionnelles et marginales.

Les notions présentées dans ce chapitre vont nous servir pour appréhender le concept de copules entre variables que nous présentons dans le prochain chapitre.

2.1 Introduction

Les copules constituent un outil statistique assez récent permettant de modéliser la dépendance entre des variables aléatoires. La fonction copule permet de relier la fonction de densité conjointe aux fonctions de densités marginales.

Dans ce chapitre, nous introduisons quelques définitions de base qui nous permettront de décrire le concept de copule et ces principales propriétés. Quelques copules paramétriques les plus utilisées ainsi que leurs propriétés sont également décrites.

2.2 Définition et théorème d'existence

2.2.1 Définition Une copule, notée C , est une fonction de répartition multivariable dont les lois marginales sont uniformes sur $[0,1]$ et qui vérifie les conditions suivantes :

$$1. C(u, 0) = C(0, u) = 0 \quad \forall u \in [0,1]. \quad (2.1)$$

$$2. C(u, 1) = C(1, u) = u \quad \forall u \in [0,1]. \quad (2.2)$$

$$3. C \text{ est 2-croissante : } C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0, \\ \text{ pour } (u_1, u_2) \in [0,1]^2, (v_1, v_2) \in [0,1]^2 \text{ tel que } u_1 \leq u_2 \text{ et } v_1 \leq v_2. \quad (2.3)$$

2.2.2 Théorème de Sklar

Ce théorème établi par Sklar en 1959 est fondamentale dans la théorie des copules, il permet de relier la loi conjointe d'un couple aléatoire et ses marginales.

Théorème 2.1 Soit F une fonction de répartition conjointe avec des marginales F_1 et F_2 .

Alors, il existe une copule C telle que pour tout $x = (x_1, x_2) \in \mathbb{R}^2$:

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \quad (2.4)$$

Réciproquement, si C est une copule et F_1 et F_2 sont des distributions univariées, alors la fonction F définie par (2.4) est la distribution conjointe dont les marginales sont F_1 et F_2 . Si F_1 et F_2 sont continues, alors C est unique

Preuve : On a :

$$\begin{aligned}
 F(x_1, x_2) &= \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \\
 &= \mathbb{P}(F_1(X_1) \leq F_1(x_1), F_2(X_2) \leq F_2(x_2)) \\
 &= C(F_1(x_1), F_2(x_2)).
 \end{aligned}$$

Ce théorème est très important car il permet d'associer à chaque distribution bidimensionnelle une copule. La relation (2.1) donne une représentation canonique de la fonction de distribution F en mettant en présence d'un côté les distributions F_1 et F_2 des distributions unidimensionnelles et d'un autre côté la copule qui permet de « cimenter » ces marginales. La structure de dépendance est représentée par la copule C .

Dans le cas de fonctions marginales continues, on peut extraire l'unique copule associées à F comme le montre le corollaire suivant.

Corollaire 2.1. Soit F une fonction de répartition conjointe avec les marginales F_1, F_2 .

Soient F_1^{-1}, F_2^{-1} les fonctions inverses de F_1, F_2 .

Alors pour tout $(u_1, u_2) \in [0, 1]^2$, on a :

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)) \quad (2.5)$$

Cette formule est utile pour l'estimation de la copule.

Preuve : F_1 et F_2 sont continues, alors $F_1(X_1)$ et $F_2(X_2)$ sont uniformes sur $[0, 1]$ et par la suite.

$$\begin{aligned}
 C(u_1, u_2) &= \mathbb{P}(F_1(X_1) \leq u_1, F_2(X_2) \leq u_2) \\
 &= \mathbb{P}(X_1 \leq F_1^{-1}(u_1), X_2 \leq F_2^{-1}(u_2)) \\
 &= F(F_1^{-1}(u_1), F_2^{-1}(u_2)).
 \end{aligned}$$

2.2.3 Fonction de densité de probabilité d'une copule

La fonction de densité de la copule C est :

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} \quad (2.6)$$

Cette fonction permet de relier la fonction de densité de probabilité conjointe f à celles de ces marginales f_1 et f_2 sans faire appel aux fonctions de densité de probabilité conditionnelle.

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2) \quad (2.7)$$

Dans le cas de N variables, on a :

C est une fonction définie de $[0,1]^N$ sur $[0,1]$

$$F(x_1, x_2, \dots, x_N) = C(F_1(x_1), \dots, F_N(x_N)). \quad (2.8)$$

et $f(x_1, x_2, \dots, x_N) = c(F_1(x_1), \dots, F_N(x_N))f_1(x_1), \dots, f_N(x_N)$

2.2.4 Bornes de Fréchet

Les bornes de Fréchet permettent de définir une région où se situent toutes les copules.

Théorème 2.2 soit F une fonction de répartition conjointe d'un couple aléatoire (X_1, X_2) de fonction de répartitions marginales F_1 et F_2 . Pour toute copule bivariable C associée à F et $(u, v) \in I^2$, on a :

$$W(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = M(u, v) \quad (2.9)$$

Preuve. On a :

$$(u, v) \in I \begin{cases} C(u, v) \\ C(u, v) \end{cases} \begin{cases} C(u, 1) = u \\ C(1, v) = v \end{cases} \quad C(u, v) \leq \min(u, v) = M \quad (2.10)$$

de la définition (2.3) et $(u, v) \in I$, on a :

$$C(u, v) \leq C(u, 1) + C(1, v) - C(1, 1) = C(u, v) \leq (u + v - 1) \text{ or}$$

$$(u, v) \in I, C(u, v) \geq 0$$

$$C(u, v) = \max(u + v - 1, 0) = W(u, v) \quad (2.11)$$

En combinant (2.10) et (2.11), on a :

$$W(u, v) \leq C(u, v) \leq M(u, v) \quad (2.12)$$

Les copules W et M sont appelées borne inférieure (respectivement borne supérieure) de Fréchet–Hoeffding ou copule minimale (respectivement copule maximale).

2.3 Propriétés d'une copule

2.3.1 Symétrie : Soient X_1, X_2 deux variables aléatoires continues, de fonction de répartition conjointe F et de marginales F_1 et F_2 , et soit C une copule.

On dit que X_1, X_2 sont échangeable si seulement si $F_1 = F_2$ et

$$C(u, v) = C(v, u) \text{ pour tout } (u, v) \in [0, 1]^2$$

Si $C(u, v) = C(v, u)$ pour tout $(u, v) \in [0, 1]^2$, On dit que C est symétrique.

2.3.2 Théorème d'invariance : Soient X_1, X_2 deux variables aléatoires continues, de marginales F_1 et F_2 , et de copule C_{X_1, X_2} . Si α et β sont deux fonctions strictement croissantes, alors

$$C_{\alpha(X_1), \beta(X_2)} = C_{X_1, X_2} \quad (2.13)$$

La copule C_{X_1, X_2} est invariante par transformation strictement croissante des variables aléatoires.

2.3.3 Ordre : Soit C_1, C_2 deux copules, On dit que C_1 est plus petite que C_2 ou C_2 est plus grand que C_1 et on note $C_1 < C_2$ Si $C_1(u, v) \leq C_2(u, v)$ pour tout $(u, v) \in [0, 1]^2$

Exemple : La copule $W = \max(u + v - 1, 0)$ est la plus petite copule et $M = \min(u, v)$ est la plus grande copule.

2.3.4 Convexité et concavité : Soit $(a, b), (c, d) \in [0,1]^2$ et $\lambda \in [0,1]$.

Une copule C est concave (convexe) si on a :

$$C(\lambda a + (1 - \lambda)c, \lambda b + (1 - \lambda)d) \geq \lambda C(a, b) + (1 - \lambda)C(c, d).$$

2.3.5 Copule harmonique : Soit C une copule dont les dérivées partielles de second ordre sont continues dans $[0,1]^2$.

C est harmonique dans $[0,1]^2$ si C satisfait l'équation de Laplace dans $[0,1]^2$

$$\frac{\partial^2 C(u, v)}{\partial u^2} + \frac{\partial^2 C(u, v)}{\partial v^2} = 0 \quad (2.14)$$

Exemple : La copule $\pi = uv$ est une copule harmonique : $\frac{\partial^2 \pi}{\partial u^2} = \frac{\partial^2 \pi}{\partial v^2} = 0$.

2.3.6 Copule homogène : Une copule C est homogène de degré k si $\forall k \in \mathbb{R}, \forall u, v, \lambda > 0, C(\lambda u, \lambda v) = \lambda^k C(u, v)$.

Exemple :

(a) La fonction $\pi = uv$ est homogène de degré 2, car $(\lambda u)(\lambda v) = \lambda^2 uv$.

(b) La fonction $M = \min(u, v)$ est homogène de degré 1 :

$$\min(\lambda u, \lambda v) = \lambda \min(u, v)$$

2.3.7 Dérivées partielles : Les dérivées partielles de $C(u, v)$ existent presque sûrement pour tout $u, v \in [0,1]$

$$0 \leq \frac{\partial C(u, v)}{\partial u} \leq 1 \text{ et } 0 \leq \frac{\partial C(u, v)}{\partial v} \leq 1.$$

2.4 Quelques copules

Il existe un grand nombre de fonctions copules, parmi elles, les plus importantes sont la copule Gaussienne, la copule de Student et les copules Achimediennes.

2.4.1 La Copule Gaussienne

La fonction de distribution de la copule Gaussienne N -dimensionnelle, s'écrit pour tout $(u_1, \dots, u_N) \in [0,1]^N$:

$$C(u_1, \dots, u_N) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \dots, (\Phi^{-1}(u_N))) \quad (2.15)$$

La fonction $\Phi_{\Sigma}(x) = \frac{(\exp -\frac{1}{2}x\Sigma^{-1}x)}{(2\pi)^{d/2} \det(\Sigma)^{1/2}}$ est la fonction de répartition de la loi normale centrée réduite et Σ sa matrice de variance covariance (égale à la matrice de corrélation dans ce cas).

La fonction de distribution Φ^{-1} est l'inverse de distribution normale centrée réduite monovariable. La fonction de densité de la copule Gaussienne de dimension N s'écrit :

$$f_c(u_1, \dots, u_d) = \frac{1}{\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}\beta(\Sigma^{-1} - I_d)\beta'\right) \quad (2.16)$$

Où I_d est la matrice unité et $\beta = (\Phi^{-1}(u_1), \dots, (\Phi^{-1}(u_N)))$.

La figure (2.1) montre l'allure de la fonction de densité d'une copule bidimensionnelle avec

une matrice de corrélation $= \begin{vmatrix} 1 & 0.5 \\ 0.5 & 1 \end{vmatrix}$

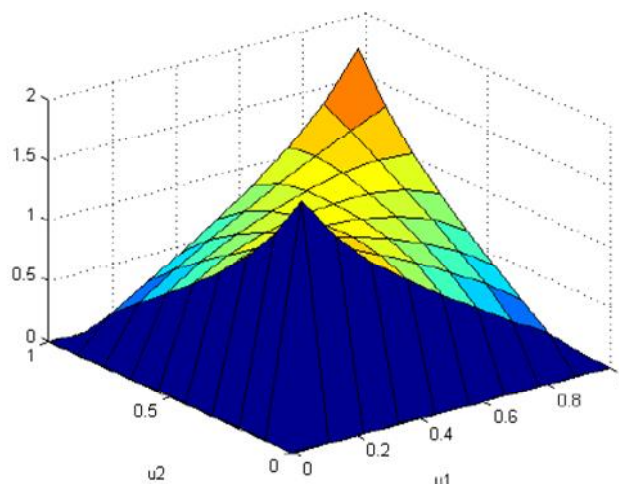


Figure (2.1) : Exemple de copule Gaussienne

2.4.2 La copule de Student

La copule de Student (t copula) est la copule sous-jacente à une distribution multi-variable de Student. Elle est définie de la même manière que la copule Gaussienne mais à partir de la distribution de Student centrée réduite. La fonction de densité de la copule de Student à N variables, s'écrit pour tout $(u_1, \dots, u_N) \in [0,1]^N$:

$$c(u_1, \dots, u_N) = \frac{f_{v\Sigma}(t_v^{-1}(u_1), \dots, t_v^{-1}(u_N))}{\prod_{i=1}^N f_v(t_v^{-1}(u_i))} \quad (2.17)$$

La fonction de distribution t_v^{-1} est l'inverse de la distribution de Student centrée réduite univariée à v degrés de liberté. La fonction $f_{v\Sigma}$ est la densité de probabilité de la loi de Student centrée réduite, Σ sa matrice de corrélation est f_v la densité univariée de la loi de Student centrée réduite ($\Sigma = 1$), qui s'écrit pour tout $x = (x_1, \dots, x_N) \in \mathbb{R}^N$:

$$f_{v\Sigma}(x) = \frac{\Gamma(\frac{v+N}{2})}{\Gamma(\frac{v}{2})\sqrt{(\pi v)^N \det(\Sigma)}} \left(1 + \frac{x\Sigma^{-1}x'}{v}\right)^{-(v+N)/2} \quad (2.18)$$

Où Γ la fonction gamma.

La figure (2.2) montre un exemple de la fonction de la densité de la copule de Student , obtenu avec $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$ et $v = 1$.

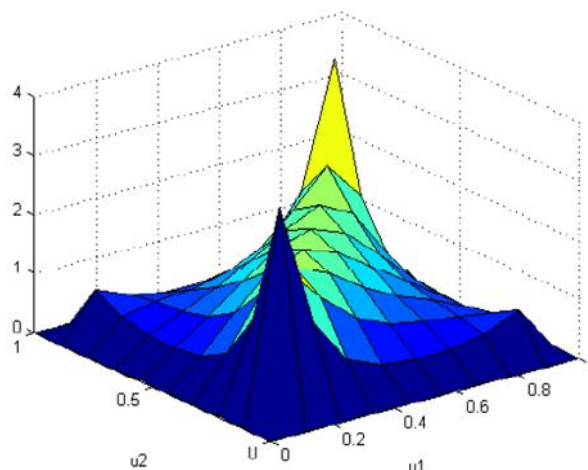


Figure (2.2) : Exemple de copule Student

2.4.3 Copules Archimédiennes

Les copules Archimédiennes sont très utilisées en pratique pour plusieurs raisons :

1. Grande variété de familles paramétriques
2. Les propriétés particulières et intéressantes que cette classe possède.
3. La facilité avec laquelle peuvent être construites et simulées.
4. La grande variété des différentes structures de dépendance.

Les copules Archimédiennes sont définies à partir d'une fonction génératrice φ telle que :

φ est une fonction continue, strictement décroissante définie de $[0,1]$ dans $[0, +\infty[$, avec $\varphi(1) = 0$.

Le pseudo-inverse (inverse généralisée) de φ est la fonction φ^{-1} de domaine de définition $[0, +\infty[$, a valeur dans $[0,1]$, donnée par

$$\varphi^{-1}(t) = \begin{cases} \varphi^{-1}(t) & \text{si } 0 \leq t \leq \varphi(0) \\ 0 & \text{si } t > \varphi(0) \end{cases}$$

φ^{-1} est continue, non croissante sur $[0, +\infty[$ et strictement décroissante sur $[0, \varphi(0)]$. En outre, $\varphi^{-1}(\varphi(u)) = u$ sur $[0,1]$ et

$$\begin{aligned} \varphi(\varphi^{-1}(t)) &= \begin{cases} t & \text{si } 0 \leq t \leq \varphi(0) \\ \varphi(0) & \text{si } t > \varphi(0) \end{cases} \\ &= \min(t, \varphi(0)). \end{aligned}$$

Si $\varphi(0) = +\infty$, alors $\varphi^{-1}(t) = \varphi^{-1}(t)$.

Définition

Soit φ une fonction continue, strictement décroissante définie de $[0,1]$ dans $[0, +\infty[$, telle que $\varphi(1) = 0$, et soit φ^{-1} son inverse généralisée. La copule C est Archimédienne Si :

$$C(u, v) = \begin{cases} \varphi^{-1}(\varphi(u) + \varphi(v)) & \text{si } \varphi(u) + \varphi(v) \leq \varphi(0) \\ 0 & \text{sinon} \end{cases}$$

Avec $\varphi(1) = 0$; $\varphi'(u) \leq 0$ et $\varphi'(u) > 0$ pour tout $0 < u < 1$.

C vérifie effectivement les propriétés fondamentales d'une copule

Preuve

$$C(u, 0) = \varphi^{-1}(\varphi(u) + \varphi(0)) = 0, \text{ et } C(u, 1) = \varphi^{-1}(\varphi(u) + \varphi(1)) = \varphi^{-1}(\varphi(u)) = u.$$

Par symétrie ; $C(0, v) = 0$ et $C(1, v) = v$.

On trouve principalement trois familles de copules Archimédienne.

a- La famille de Gumbel :

Elle est définie par la fonction génératrice $\varphi(t) = (-\ln t)^\theta$, $\theta \geq 1$ et $t \in [0,1]$.

$$\text{avec } \varphi'(t) = -\frac{\theta}{t}(-\ln t)^{\theta-1} \text{ et } \varphi'(t) < 0.$$

φ est une fonction continue strictement décroissante définie de $[0,1]$ dans $[0, +\infty[$, convexe.

$$\varphi''(t) = \frac{\theta}{t^2}(-\ln t)^{\theta-2}[\theta - 1 - \ln t] \text{ et } \varphi''(t) > 0.$$

La fonction génératrice inverse est donnée par :

$$\varphi^{-1}(t) = \exp(-t^{1/\theta}) \quad (2.19)$$

Dans ce cas, la fonction copule bidimensionnelle de Gumbel est déduite à partir de l'équation (2.17) :

$$C_\theta(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right) \quad (2.20)$$

La figure (2.3) montre la fonction de densité d'une copule de Gumbel avec $\theta = 1.5$.

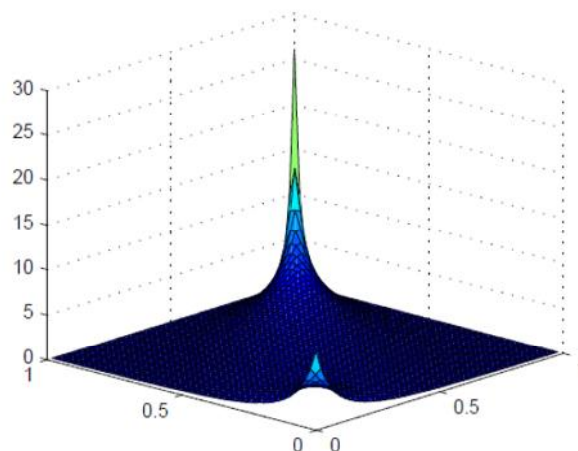


Figure (2.3) : fonction de densité d'une copule de Gumbel

b- La famille de Clayton

Elle est définie par la fonction génératrice $\varphi(t) = (t^{-\theta} - 1) / \theta$ telle que $\theta \in [-1, \infty) \setminus \{0\}$

A partir de l'expression (2.18), la famille des copules de Clayton prend la forme suivante:

$$C_{\theta}(u, v) = \max\left([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0\right) \quad (2.21)$$

Pour $\theta > 0$, cette expression devient simplement :

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

$$\frac{\varphi(t)}{\varphi'(t)} = \frac{t^{\theta+1} - t}{\theta}$$

Cette copule est différentiable et sa densité est donnée par :

$$\begin{aligned} c(u_1, u_2; \theta) &= \frac{\partial}{\partial u_2} \left(\frac{\partial C(u_1, u_2; \theta)}{\partial u_1} \right) \\ &= \frac{\partial}{\partial u_2} \left[\frac{-1}{\theta} (u_1^{-\theta} + u_2^{-\theta} - 1)^{\frac{-1}{\theta}-1} (-\theta u_1^{-\theta-1}) \right] \\ &= \frac{\partial}{\partial u_2} \left[(u_1^{-\theta} + u_2^{-\theta} - 1)^{\frac{-1-\theta}{\theta}} (-u_1^{-\theta-1}) \right] \\ &= \frac{-\theta - 1}{\theta} \left[(u_1^{-\theta} + u_2^{-\theta} - 1)^{\frac{-1-\theta}{\theta}-1} (-\theta u_2^{-\theta-1}) (-u_1^{-\theta-1}) \right] \\ &= (\theta + 1) (u_1 u_2)^{-\theta-1} (u_1^{-\theta} + u_2^{-\theta} - 1)^{\frac{-1}{\theta}-2} \end{aligned} \quad (2.22)$$

La figure (2.4) montre l'allure de la fonction de densité de cette copule avec $\theta = 2.5$.

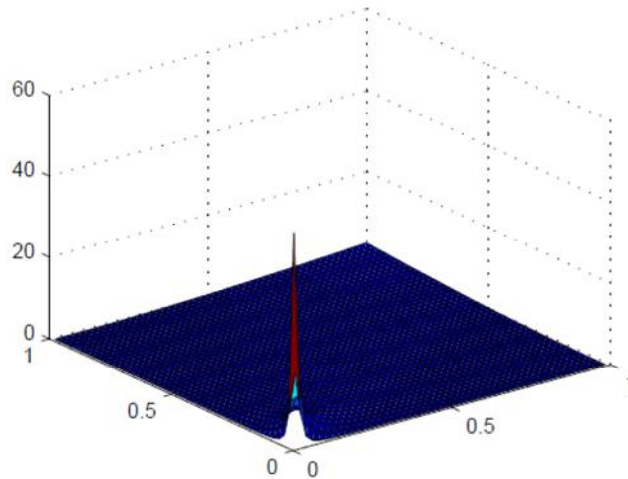


Figure (2.4) : Allure de la fonction de densité de copule de Clayton

c-La famille de Frank

Elle est définie par la fonction génératrice $\varphi(t) = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$, telle que $\theta \in \mathbb{R} \setminus \{0\}$.

avec

$$\varphi^{-1}(s) = -\frac{1}{\theta} \ln(1 - (1 - e^{-\theta})e^{-s}) \text{ et } \varphi^{-1}(s) = \frac{-\frac{1}{\theta}(1 - e^{-\theta})e^{-s}}{1 - (1 - e^{-\theta})e^{-s}}$$

La copule bidimensionnelle de la famille de Frank est donnée par

$$C_{\theta}(u, v) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right)$$

La fonction de densité de la copule de Frank est égale à :

$$\begin{aligned} c(u_1, u_2; \theta) &= \frac{\partial}{\partial u_2} \left(\frac{\partial C(u_1, u_2; \theta)}{\partial u_1} \right) \\ &= \frac{\partial}{\partial u_2} \left[\frac{-1}{\theta} \left(\frac{-\theta \exp^{-\theta u_1} (\exp^{-\theta u_2} - 1)}{(e^{-\theta} - 1) + (\exp^{-\theta u_1} - 1)(\exp^{-\theta u_2} - 1)} \right) \right] \\ &= \frac{\theta(1 - \exp^{-\theta}) \exp^{-\theta(u_1 + u_2)}}{[(1 - \exp^{-\theta}) - (\exp^{-\theta u_1})(\exp^{-\theta u_2} - 1)]^2} \end{aligned} \tag{2.23}$$

La figure (2.5) montre l'allure d'une fonction de densité de la copule de Frank avec $\theta = 0.6$.

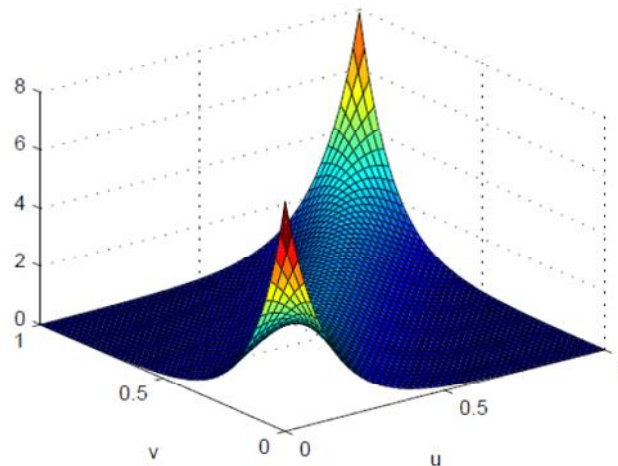


Figure (2.5) : Fonction de densité de la copule de Frank.

2.5 Conclusion

Dans ce chapitre, nous avons introduit quelques définitions de base qui nous a permis de décrire le concept de copule et ces principales propriétés. Quelques copules paramétriques les plus utilisées ainsi que leurs propriétés sont également décrites.

Nous pouvons retenir qu'une copule est une fonction de répartition qui permet de relier une fonction de densité de probabilité conjointe à plusieurs variables à celles de ces marginales sans faire appel aux fonctions de densité de probabilité conditionnelle. Cette caractéristique fera l'objet de son application à la segmentation d'une image couleur qui sera présentée en détails dans le prochain chapitre.

3.1 Introduction

La segmentation d'images constitue une étape essentielle en traitement d'images car elle conditionne la qualité des traitements effectués ultérieurement.

Dans ce dernier chapitre, nous présenterons d'abord les différentes approches de la segmentation d'images, ensuite, nous montrerons comment utiliser le concept des copules pour segmenter une image couleur. La méthode de segmentation ainsi élaborée sera testée sur des images couleur.

3.2 Segmentation d'images couleur

La segmentation d'images est une opération qui a pour but de regrouper les pixels d'une image suivant des critères prédéfinis. Les pixels sont ainsi regroupés en régions qui constituent un pavage ou une partition de l'image.

Contrairement aux images en niveaux de gris, une image couleur I est composée de pixels où chaque pixel est caractérisé par trois composantes au lieu d'une seule (niveau de gris).

Les trois composantes les plus courantes, car issues des caméras couleur, sont le rouge (R), le vert (V) et le bleu (B). Dans ce cas, on parle d'images RVB. D'autres composantes couleurs dérivant du système RVB ont été proposées. Parmi elles, on peut citer le système XYZ, La^*b^* , YUV, etc....

Les méthodes de segmentation d'images couleur sont une extension des méthodes de segmentation en niveaux de gris au cas tridimensionnel. Elles peuvent être abordées selon l'approche contour ou l'approche région. L'approche contour cherche à extraire les contours des objets présents dans l'image. Elle s'appuie sur la détection des changements abrupts de la couleur. Tandis que l'approche région cherche à regrouper les pixels ayant des attributs couleur identiques en régions homogènes. Ces régions sont construites en évaluant la similarité entre les pixels ou entre un pixel et ceux d'une même région d'un point de vue colorimétrique. Dans cette approche, on distingue les méthodes par croissance de régions, par division-fusion et par classification. Les méthodes basées sur la classification considèrent une classe comme l'ensemble des pixels ayant des attributs colorimétriques similaires. Elles supposent donc que les pixels qui appartiennent à une même ou plusieurs régions possèdent des caractéristiques similaires et forment un nuage de points dans l'espace colorimétrique. La classification consiste alors à regrouper les points proches en groupes

ou classes. Cette opération peut être réalisée de deux manières: la première suppose l'existence de certains pixels (prototypes) dont l'appartenance aux classes est connue à priori, elle nécessite l'intervention de l'utilisateur. La seconde dite non supervisée (clustering), vise à regrouper automatiquement des pixels de l'image en classes sans aucune connaissance préalable sur l'appartenance des pixels aux classes.

On s'est intéressé dans notre travail à la classification supervisée et plus exactement à la classification Bayésienne.

3.2.1 Segmentation d'images basée sur la classification Bayésienne

La classification Bayésienne est une méthode classification supervisée de type probabiliste. Elle fait appel à la règle de Bayes afin de classer des objets quelconques selon des probabilités conditionnelles.

Soit $\mathbf{X} = [X_1, X_2, \dots, X_Q]$ un ensemble de Q pixels d'une image couleur I où chaque pixel X_q est caractérisé par trois attributs correspondant, dans notre cas, aux trois composantes couleurs r, v, b , tel que $X_q = [r_q, v_q, b_q]^T$ ($q = 1, 2, \dots, Q$).

La segmentation supervisée sous l'hypothèse paramétrique, suppose qu'on dispose de pixels (prototypes) dont l'appartenance aux K classes C_k , $k = 1, 2, \dots, K$ est connue à priori. D'un point de vue statistique, l'ensemble des prototypes forme un échantillon composé d'observations tirées, indépendamment, de K classes C_k selon les lois de probabilité conditionnelles $p(X_q/C_k)$ connues ($X = [r, v, b]^T$). Par application du théorème de Bayes, il est possible de déterminer la probabilité $P(C_k/X_q)$ pour qu'un pixel X_q appartienne à la classe C_k . Cette probabilité, appelée aussi probabilité a posteriori de la classe C_k est donnée par la relation :

$$P(C_k/X_q) = \frac{p(X_q/C_k)P(C_k)}{p(X_q)} \quad (3.1)$$

$p(X_q) = \sum_{k=1}^K p(X_q/C_k)P(C_k)$ désigne la valeur de la fonction de densité sous-jacente à la distribution des observations au point X_q et $P(C_k)$ désigne la probabilité a priori de la classe C_k de telle sorte que :

$$\sum_{k=1}^K P(C_k) = 1 \quad (3.2)$$

$p(X_q/C_k)$ désigne la loi de probabilité conditionnelle de la classe C_k c'est-à-dire, la fonction de densité de probabilité sous-jacente à la distribution des observations provenant de la classe C_k . Dans ces conditions, le risque d'erreur lié à l'affectation de l'observation X_q à la classe C_k est donné par :

$$R(C_k/X_q) = 1 - P(C_k/X_q) \quad (3.3)$$

Pour obtenir un classement optimal, il faut affecter toute observation X_q à la classe C_k qui minimise ce risque, donc qui maximise la probabilité a posteriori. Le problème de classification d'une observation X_q peut être résolu en adoptant la règle de décision suivante :

Décider C_s si $P(C_s/X_q) > P(C_k/X_q), \forall k = 1, 2, \dots, K, k \neq s$.

L'application de la règle de Bayes revient à trouver des fonctions de décision $d_k(X_q), k = 1, 2, \dots, K$, tel que l'observation X_q soit affectée à la classe C_k si seulement si:

$$d_k(X_q) > d_i(X_q) \quad i = 1, 2, \dots, K, i \neq k \text{ avec } d_k(X_q) = p(X_q/C_k)P(C_k).$$

Mélange Gaussiens

Dans les conditions particulières où les échantillons ont la structure de mélanges gaussiens, la fonction de densité de probabilité au point X_q est donnée par :

$$p(X_q) = \sum_{k=1}^K p(X_q/C_k)P(C_k) \quad (3.4)$$

où la fonction de densité de probabilité conditionnelle multivariable de la classe C_k est définie par :

$$p(X_q/C_k) = \frac{1}{(2\pi)^{N/2} |\Sigma_k|^{1/2}} \exp\left\{-\frac{1}{2}(X_q - \bar{X}_k)^T \Sigma_k^{-1}(X_q - \bar{X}_k)\right\} \quad (3.5)$$

\bar{X}_k est le vecteur moyenne de la classe C_k et Σ_k la matrice de covariance de la classe C_k .

En considérant le logarithme, les fonctions de décision d_k peuvent être écrites sous la forme :

$$d_k(X_q) = -\frac{1}{2} \text{Log}|\Sigma_k| - \frac{1}{2}(X_q - \bar{X}_k)^T \Sigma_k^{-1}(X_q - \bar{X}_k) + \text{Log}[P(C_k)] \quad (3.6)$$

En pratique, le vecteur moyenne \bar{X}_k et la matrice de covariance σ_k de la classe C_k sont estimés à partir des prototypes de la classe C_k .

$$\bar{X}_k = \frac{1}{Q_k} \sum_{q=1}^{Q_k} X_q \quad (3.7)$$

$$\sigma_k = \frac{1}{Q_k} \sum_{q=1}^{Q_k} (X_q - \bar{X}_k)^2 \quad (3.8)$$

avec $X_q \in C_k$

L'algorithme ci-dessous donne les étapes de la segmentation d'une image couleur basée sur la décision Bayésienne et dans le cas d'un mélange Gaussien.

- 1 - Charger les images couleur et leurs images de référence
- 2 - Tirer un nombre de pixels prototypes pour chaque classe
- 3 - Calculer le vecteur moyenne \bar{X}_k et la matrice de covariance σ_k de chaque classe à partir des équations (3.7) et (3.8)
- 4 - Pour chaque pixel X_q de l'image
 - 4.1 - Pour chaque classe C_k de l'image
 - Calculer la fonction $d_k(X_q)$ selon l'équation (3.6)
 - Fin Pour
 - 4.2 - Affecter le pixel X_q à la classe C_{k^*} telle que $k^* = \max_{k=1, \dots, K} d_k(X_q)$
 - Fin Pour
- 5- Afficher l'image segmentée

La règle de classification Bayésienne est optimale, aucune autre procédure ne peut conduire à un taux d'erreur plus petit. C'est la raison pour laquelle cette règle est prise comme référence pour une étude comparative des résultats qu'elle fournit avec ceux obtenus avec d'autres méthodes de classification, dans le cas des distributions dont les caractéristiques statistiques sont connues.

3.3 Méthode de segmentation basée sur les copules

La méthode de segmentation d'images couleur que nous proposons est également basée sur la méthode de classification Bayésienne. L'idée de cette méthode est d'utiliser le concept des copules pour estimer la loi de probabilité conditionnelle multivariable $p(X_q/C_k)$ de la classe C_k , c'est-à-dire, la fonction de densité de probabilité sous-jacente à la distribution des pixels provenant de la classe C_k .

Dans le cas d'un image RVB, cette fonction peut s'écrire:

$$\begin{aligned}
p(X_q/C_k) &= f(r, v, b/C_k) \\
&= f_R(r/C_k) f_V(v/C_k) f_B(b/C_k) c(F_R(r/C_k), F_V(v/C_k), F_B(b/C_k))
\end{aligned} \tag{3.9}$$

avec $f_R(r/C_k)$, $f_V(v/C_k)$ et $f_B(b/C_k)$ les fonctions de densité marginale de la classe C_k .

$F_R(r/C_k)$, $F_V(v/C_k)$ et $F_B(b/C_k)$ les fonctions de répartition marginales de la classe C_k .

$c(F_R(r/C_k), F_V(v/C_k), F_B(b/C_k))$ est la fonction de densité de la copule par rapport à la classe C_k .

Ainsi, les fonctions de décision d_k peuvent être écrites comme :

$$\begin{aligned}
d_k(X_q) &= P(C_k) f(r_q, v_q, b_q/C_k) = \\
&P(C_k) f_R(r_q/C_k) f_V(v_q/C_k) f_B(b_q/C_k) c(F_R(r_q/C_k), F_V(v_q/C_k), F_B(b_q/C_k))
\end{aligned} \tag{3.10}$$

En pratique, les fonctions de densité marginale et les fonctions de répartition marginales de chaque classe sont estimées par la méthode non paramétrique des noyaux à partir des prototypes représentatifs des classes avec un noyau Gaussien (voir chap. 1).

$$f_{RouVouB}(x/C_k) = \frac{1}{Q_k h} \sum_{q=1}^{Q_k} k\left(\frac{x-X_q}{h}\right) \tag{3.11}$$

$$F_{RouVouB}(x/C_k) = \frac{1}{Q_k} \sum_{q=1}^{Q_k} K\left(\frac{x-X_q}{h}\right) \tag{3.12}$$

avec $X_q \in C_k$ et $x = r \text{ ou } v \text{ ou } b$, $x = 0, 1, \dots, 255$.

La largeur h du noyau est fixée à 1.

Le modèle de la fonction de densité de la copule est fixé a priori (copule Gaussienne ou Student dans notre cas). Seuls les paramètres du modèle sont alors estimés. Il s'agit de la matrice de corrélation ρ_k pour la copule Gaussienne (voir Eq. 2.13 du chap. 2) et la matrice de corrélation ρ_k ainsi le degré de liberté ν pour la copule de Student (voir Eq. 2.15 du chap. 2).

La probabilité a priori $P(C_k)$ de la classe C_k est égale au nombre de prototypes de la classe C_k sur le nombre de pixels de l'image.

L'algorithme ci-dessous donne les étapes de la segmentation d'une image couleur basée sur les copules.

- 1 - Charger les images couleur et leurs images de référence
- 2 - Tirer un nombre de pixels prototypes pour chaque classe
- 3 - Pour chaque classe C_k de l'image
 - Estimer en utilisant les prototypes de chaque classe
 - les fonctions de densité marginale $f_R(r/C_k)$, $f_V(v/C_k)$ et $f_B(b/C_k)$ par la méthode des noyaux
 - les fonctions de répartition marginale $F_R(r/C_k)$, $F_V(v/C_k)$ et $F_B(b/C_k)$ par la méthode des noyaux
 - les paramètres de la copule matrice de corrélation ρ_k et degré de liberté ν pour la copule de Student.
 - Fin Pour
- 4 - Pour chaque pixel X_q de l'image
 - 4.1 - Pour chaque classe C_k de l'image
 - Calculer la fonction $d_k(X_k)$ selon l'équation (3.10)
 - Fin Pour
 - 4.2 - Affecter le pixel X_q à la classe C_{k^*} telle que $k = \max_{k=1, \dots, K} d_k(X_q)$
 - Fin Pour
- 5- Afficher l'image segmentée

3.4 Tests et résultats

Pour évaluer la méthode de segmentation d'images couleur proposée, nous l'avons appliquée sur des images issues de la base de Berkeley. La figure (3.1) montre quatre images tests avec leurs images de référence. La première image contient deux classes, la deuxième et la troisième 3 classes et la quatrième 6 classes. Les images de référence permettent de tirer aléatoirement 50% de pixels prototypes pour chaque classe. Ces pixels prototypes sont alors utilisés pour estimer les probabilités à priori $P(C_k)$ et les fonctions de densité de probabilités conjointes $f(r, v, b/C_k)$ de chaque classe C_k . La méthode proposée est comparée avec la méthode basée sur le mélange Gaussien. Toutes les procédures ont été implémentées sous l'environnement Matlab.

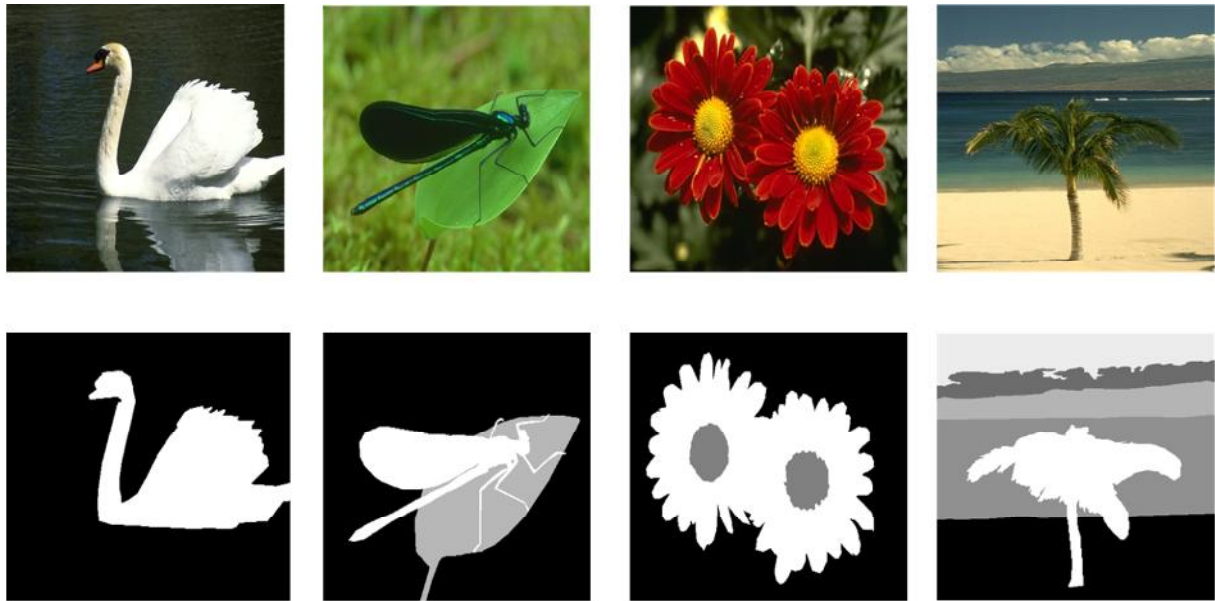


Fig. 3.1: Images couleur tests.

Les figures (3.2) à (3.5) montrent les composantes R,V, B sous forme d'images, leurs histogrammes ainsi que les fonctions de densité marginale estimées par la méthode des noyaux de chaque classe.

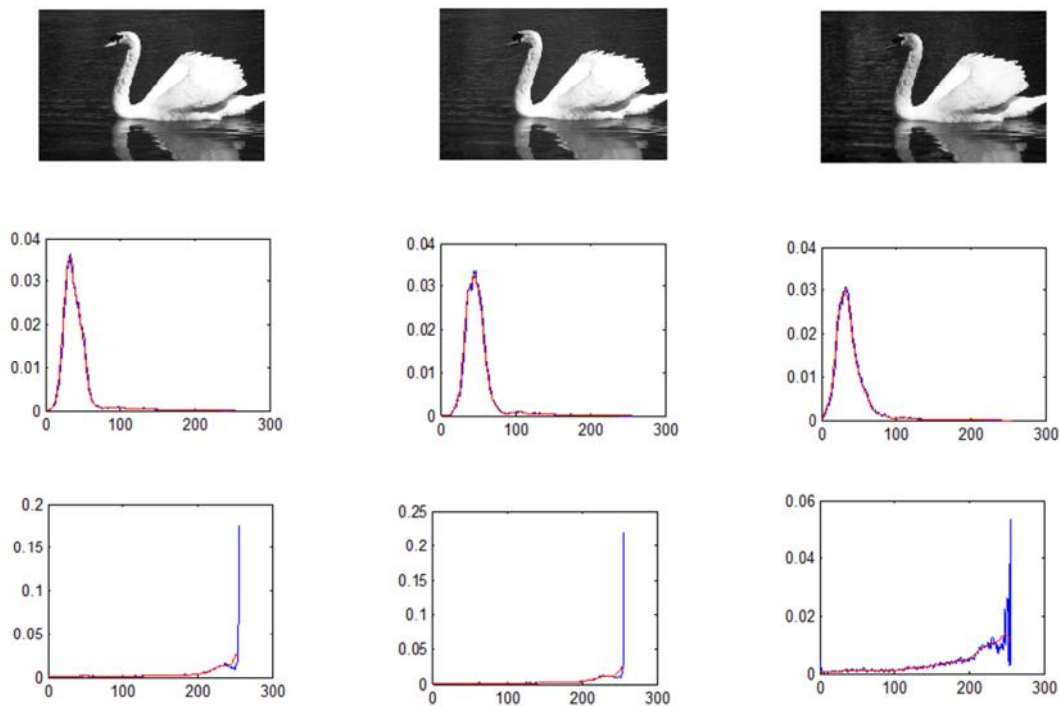


Fig. 3.2: Composantes R,V, B, leurs histogrammes ainsi que les fonctions de densité marginale de l'image test 1.

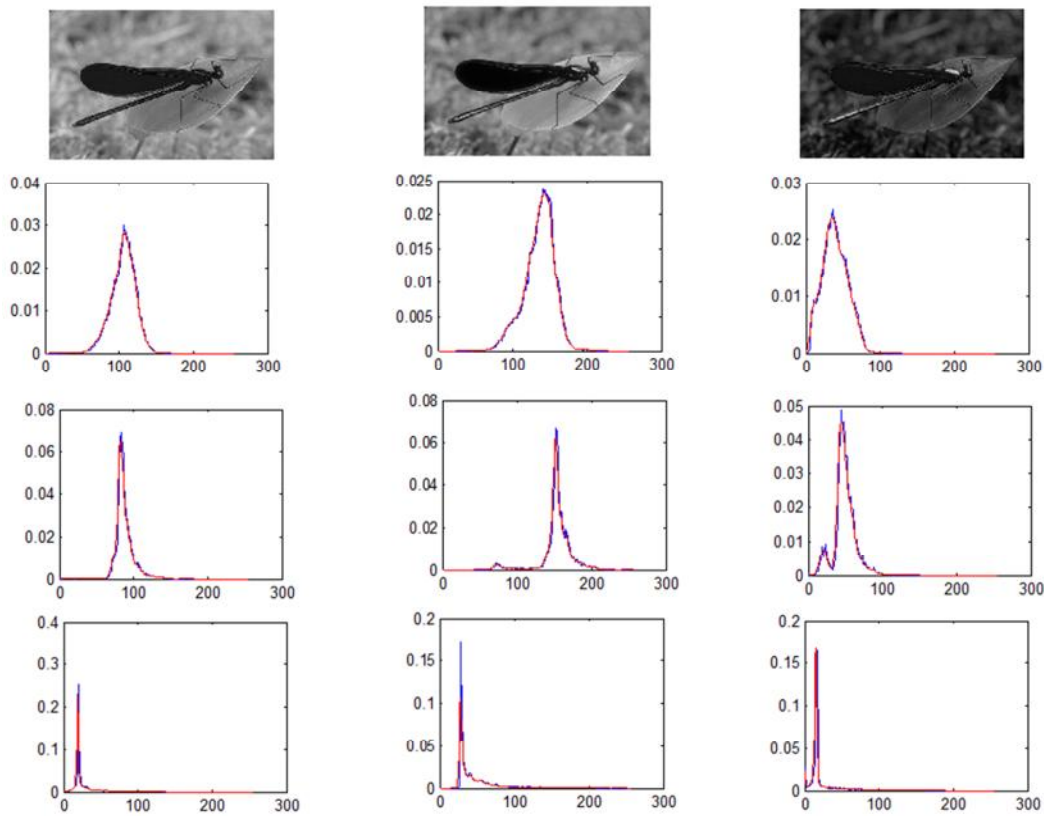


Fig. 3.3: Composantes R,V, B, leurs histogrammes ainsi que les fonctions de densité marginale de l'image test 2.

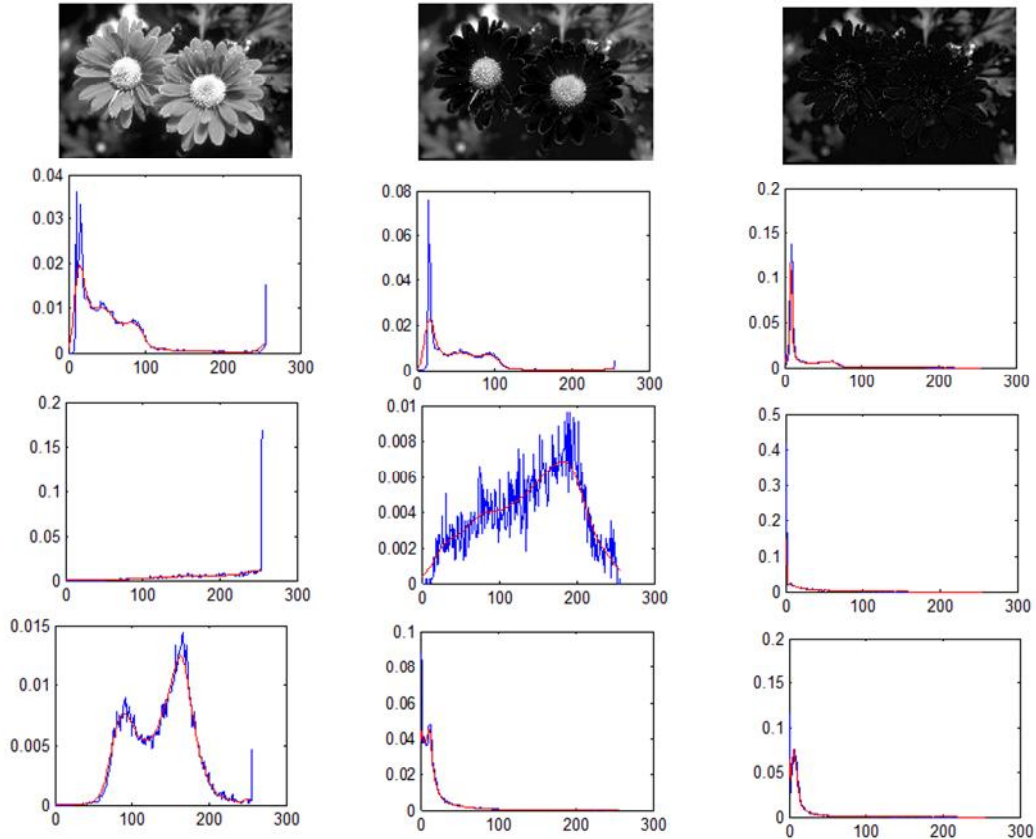


Fig. 3.4: Composantes R,V, B, leurs histogrammes ainsi que les fonctions de densité marginale de l'image test 3.

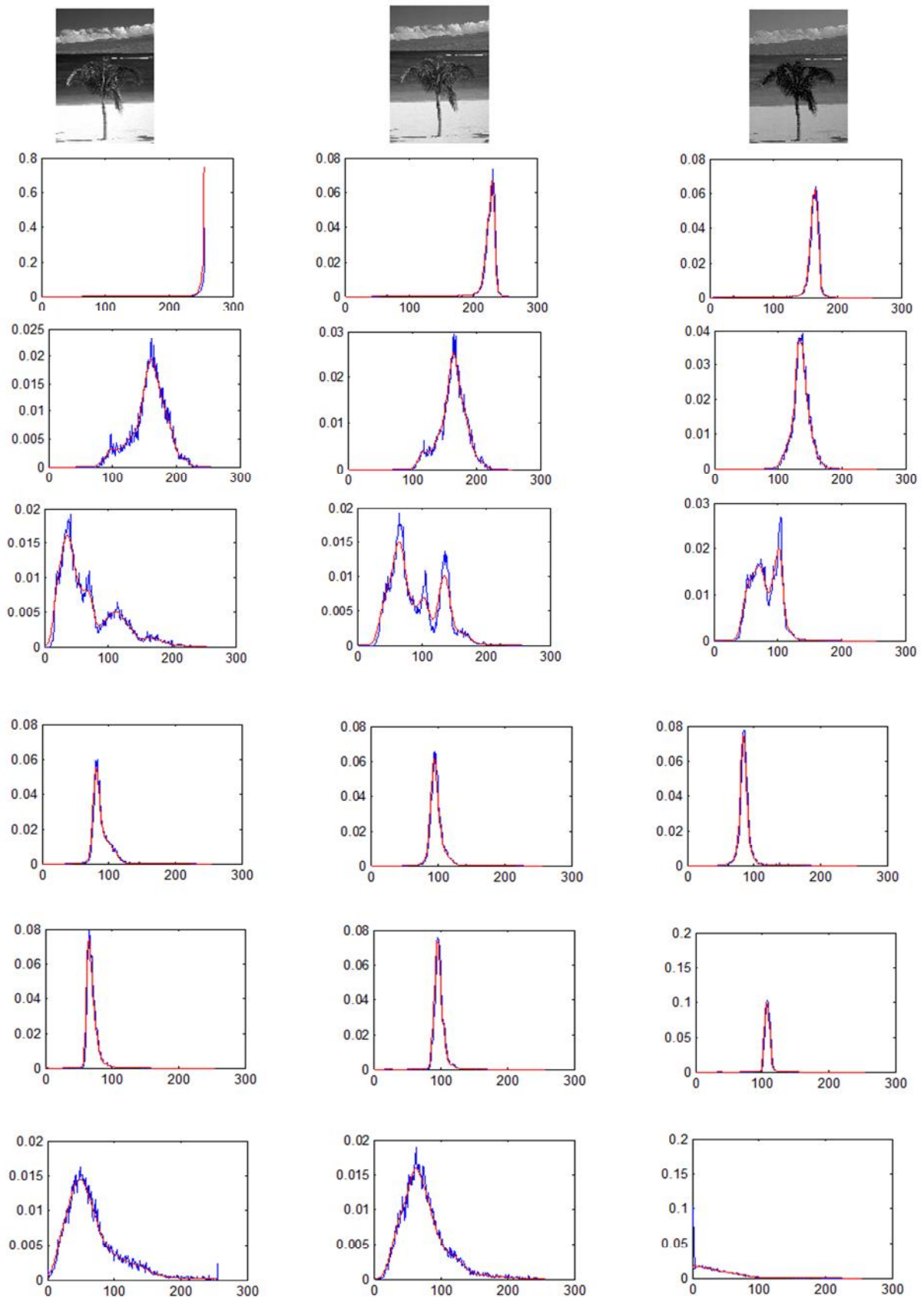
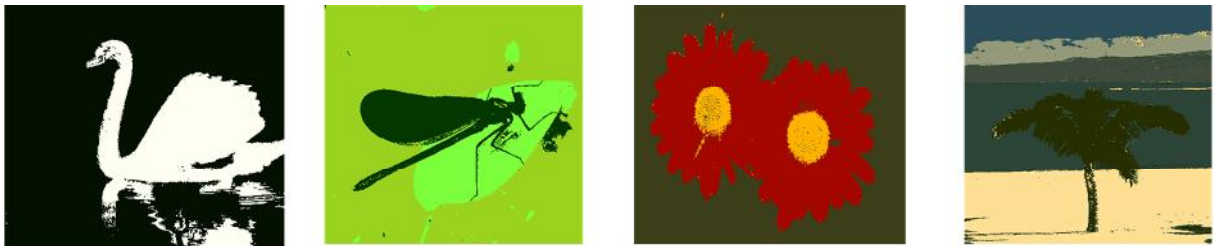


Fig. 3.5: Composantes R,V, B, leurs histogrammes ainsi que les fonctions de densité marginale de l'image test 4.

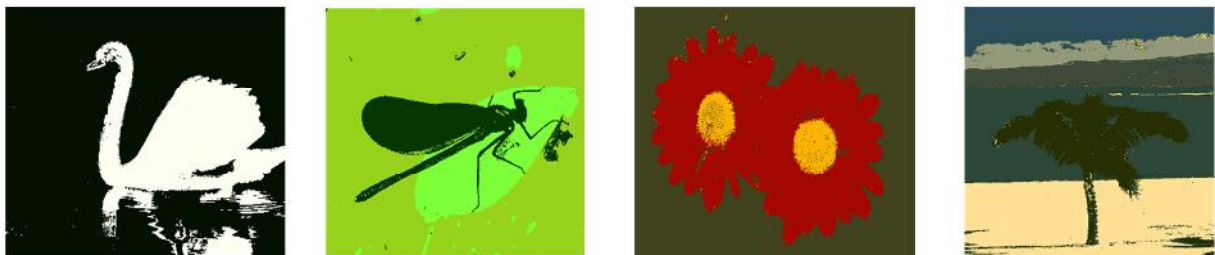
Ces figures montrent que les fonctions de densité marginale, estimées par la méthode des noyaux, décrivent correctement la distribution des niveaux des couleurs R, V et B puisque les allures de ces distributions sont des versions lissées des histogrammes correspondants.

Le tableau (3.1) regroupe les valeurs estimées des paramètres des copules (matrice de corrélation pour la copule Gaussienne et matrice de corrélation et le degré de liberté pour la copule de Student) de chaque classe et de chaque image.

La figure (3.6) regroupe les images segmentées obtenues par la méthode proposée et celle basées sur le mélange Gaussien. Elle montre clairement que la méthode proposée donne de meilleurs résultats puisque les régions homogènes sont mieux détectées. Ce constat est vérifié par le taux d'erreur de classification estimé par le nombre de pixels mal classés sur le nombre total de pixels (voir tableau (3.2) ci-dessous). Notons que ces taux sont évalués avec tous les pixels de l'image, c'est à dire les pixels prototypes et les pixels non étiquetés.



Méthode de Student.



Méthode de Gaussienne.



Mélange Gaussien.

Fig. 3.6: Images tests segmentées

Image test	Classe	Copule Gaussienne			Copule de Student			Degré de liberté
		Matrice de corrélation			Matrice de Corrélation			
1	1	1.0000	0.9695	0.8663	1.0000	0.9687	0.8616	4.4686
	2	0.9695	1.0000	0.9366	0.9687	1.0000	0.9345	
0.8663		0.9336	1.0000	0.8616	0.9345	1.0000		
2	2	1.0000	0.9859	0.9414	1.0000	0.9892	0.9400	2.7776
		0.9859	1.0000	0.9704	0.9892	1.0000	0.9674	
0.9414	0.9704	1.0000	0.9400	0.9674	1.0000			
2	3	1.0000	1.6059	0.3393	1.0000	0.6202	0.3946	4.3867
		0.6059	1.0000	0.3675	0.6202	1.0000	0.6223	
0.3393	0.3675	1.0000	0.3946	0.3623	1.0000			
3	1	1.0000	0.9473	0.9784	1.0000	0.9725	0.7893	0.0000
		0.9784	1.0000	0.7945	0.9725	1.0000	0.8306	
0.7184	0.7945	1.0000	0.7893	0.8306	1.0000			
3	2	1.0000	0.9220	0.5199	1.0000	0.9215	0.5164	3.4517
		0.9220	1.0000	0.5733	0.9215	1.0000	0.5720	
0.5199	0.5733	1.0000	0.5164	0.5720	1.0000			
3	3	1.0000	0.2058	0.5914	1.0000	0.1690	0.5972	0.0000
		0.2058	1.0000	0.5139	0.1690	1.0000	0.5031	
0.5914	0.5139	1.0000	0.5972	0.5031	1.0000			
4	1	1.0000	0.8132	0.7835	1.0000	0.8256	0.8169	2.3947
		0.8132	1.0000	0.9562	0.8256	1.0000	0.9675	
0.7835	0.9562	1.0000	0.8169	0.9675	1.0000			
4	2	1.0000	0.9890	0.8668	1.0000	0.9894	0.8656	10.3581
		0.9890	1.0000	0.9162	0.9894	1.0000	0.9140	
0.8668	0.9162	1.0000	0.8656	0.9140	1.0000			
4	3	1.0000	0.9864	0.9383	1.0000	0.9868	0.9482	4.2240
		0.9864	1.0000	0.9681	0.9868	1.0000	0.9743	
0.9383	0.9681	1.0000	0.9482	0.9743	1.0000			
4	4	1.0000	0.9078	0.7183	1.0000	0.9101	0.6993	3.1006
		0.9078	1.0000	0.8684	0.9101	1.0000	0.8517	
0.7183	0.8684	1.0000	0.6993	0.8517	1.0000			
4	5	1.0000	0.9803	0.8863	1.0000	0.9775	0.9223	4.5707
		0.9803	1.0000	0.9044	0.9775	1.0000	0.9318	
0.8863	0.9044	1.0000	0.9223	0.9318	1.0000			
4	6	1.0000	0.9588	0.7363	1.0000	0.9600	0.7332	18.1740
		0.9588	1.0000	0.8389	0.9600	1.0000	0.8359	
0.7363	0.8389	1.0000	0.7332	0.8359	1.0000			

Tableau 3.1: Valeurs estimées des paramètres des copules

D'après ce tableau, on voit que les valeurs de la matrice de corrélation sont entre 0 et 1 et sont presque les mêmes dans la copule Gaussienne et la copule de Student, ce qui représente une forte corrélation entre les composantes R, V et B.

Image test	Copule gaussien	Copule Student	Mélange Gaussien
1	0.0691	0.0658	0.0966
2	0.0413	0.0418	0.0855
3	0.0318	0.0276	0.0276
4	0.0664	0.0595	0.1673

Tableau 3.2: Taux d'erreur de classification

Ces résultats montrent également que la copule de Student est plus performante que la copule Gaussienne car elle aboutie à un taux d'erreur inférieur pour les images 1, 3 et 4.

Les résultats précédents ont été obtenus en considérant 50% de pixels de chaque classe comme prototypes. Nous avons fait varier le nombre de prototypes afin d'évaluer les performances de la méthode proposée. Le tableau (3.3) regroupe les valeurs du taux d'erreur de classification obtenus pour différents nombres de prototypes par classe sur la 2^{ème} image test. Il montre que la segmentation basée sur les copules est meilleurs que celle basée sur les Gaussienne. De plus, au-delà de 20% de prototypes par classe, le taux de classification reste presque constant.

Image test	5%	10%	20%	30%	50%	60%	80%
Mélange Gaussien	0.1392	0.2282	0.1330	0.0982	0.0855	0.0799	0.1035
Copule gaussien	0.0849	0.0760	0.0537	0.0537	0.0413	0.0393	0.0381
Copule Student	0.0844	0.0776	0.0535	0.0477	0.0418		0.0427

Tableau 3.3 : Variation du taux d'erreur de classification en fonction du nombre de prototypes par classe de la 2^{ème} image test.

3.5 Conclusion

Nous avons présenté dans ce chapitre une méthode de segmentation d'images couleur que basée sur la méthode de classification supervisée Bayésienne qui intègre le concept de copules. L'intérêt de cette méthode est de représenter la loi de probabilité conditionnelle multivariable d'une classe ou d'une région de l'image par une composition d'une fonction copule et de fonctions de densité marginale.

Les résultats expérimentaux portant sur différents images ont montré l'efficacité de cette méthode de segmentation par rapport à la technique, bien connue, basée sur le mélange gaussien.

Conclusion générale

Nous avons abordé dans ce mémoire la segmentation d'images couleur basée sur une approche de classification de type probabiliste dans un cadre supervisé.

Cette approche fait appel à la méthode de classification Bayésienne et plus exactement à la règle de Bayes afin de classer les pixels selon les fonctions de densité de probabilité conditionnelle de chaque classe.

La fonction de densité de probabilité conditionnelle de chaque classe est couramment décrite par une loi Gaussienne multivariable. Pour éviter cette hypothèse, restrictive, nous avons proposé dans ce mémoire de substituer la fonction de densité de probabilité conditionnelle multivariable de chaque classe par des fonctions copules afin de modéliser la dépendance entre les variables aléatoires. Dans ce cas, la fonction de densité de probabilité conjointe à plusieurs variables est redéfinie comme un produit d'une fonction copule multivariable et de fonctions de densité marginale relative à chaque composante couleur (variable) de l'image. Comme fonction copule, nous avons utilisé la copule Gaussienne et la copule de Student.

Les résultats expérimentaux portant sur la segmentation de différentes images couleur ont montré l'efficacité de l'utilisation des fonctions copules par rapport à celles des fonctions Gaussiennes. Ces résultats ont également montré que le choix de la copule Gaussienne ou celle de Student dépend du type de l'image.

Il serait intéressant de tester d'autres fonctions copules multivariées comme les Archimédiennes (Gumbel, Frank et Clayton).

Les variables considérées dans notre travail correspondent aux composantes R, V et B de l'image couleur. L'évaluation de la méthode de segmentation d'images couleur basée sur les copules en considérant d'autres variables (composantes couleur) est envisageable.

L'extension de l'utilisation des copules pour la segmentation d'images couleur dans un cadre non supervisé, ne disposant pas de prototypes des classes, constitue une perspective intéressante.

- 1- <http://www/-sop-inria.fr/membres/Olivier.Faugueras/MVA/.../probabilités.pdf>/12-10-2008
- 2- J. Kittler and J. Illingworth. Minimum error thresholding. *Pattern Recognition*, vol. 19, pp. 41-47, 1986.
- 3- M. Evans. *Statistical distributions*. A Wiley-Interscience Publication, 1993.
- 4- N. L. Johnson, S. Kotz and N. Balakrishnan. *Continuous univariate distributions*. Volume 1, Wiley-Interscience Publication, 1970.
- 5- T. T. Soong. *Fundamentals probability and statistics for engineers*. A Wiley-Interscience Publication, 2004.
- 6- P. Tassi. *Méthodes statistiques*. *Méthodes statistiques*. Collection : Economie et Statistiques Avancées. Economica, Paris, 1985.
- 7- F. Flitti. *Techniques de réduction de données et analyse d'images multispectrales astronomiques par arbres de Markov*. Thèse de doctorat, Université Louis Pasteur-Strasbourg I, 2005.
- 8- M. K. Varanasi and B. Aazhang. Parametric generalized gaussian density estimation. *J. Acoust. Soc. Amer*, 86, pp. 1404-1415, 1989.
- 9- J. N. Provost. *Classification bathymétrique en imagerie multispectrale SPOT*. Thèse de doctorat, Université de Bretagne Occidentale - Ecole Navale (Laboratoire GTS), 2001.
- 10- C. Walck. *Hand-book on statistical distributions for experimentalists*. Particle Physics Group Fysikum University of Stockholm, 2007.
- 11- E. Monfrini. *Identifiabilité et méthode des moments dans les mélanges généralisés de distributions du système de Pearson*. Thèse de doctorat, Université Paris, 2002.
- 12- Lounas fadhila. (2011)-*Modélisation de la dépendance par les copules et application-mémoire de magister (université Tizi-Ouzou)*.
- 13- Kadi Nabil.(2014)-*Estimation-non paramétrique de la distribution et densité de copules-mémoire recherche (université de Sherbrooke)*.
- 14- Slimani (2012)-*Estimation des paramètres d'une copule_mémoire de magister (université Mohamed khider Biskra)*.
- 15- Pravin K. Trivedi and David M. Zimmer (2005)-*Copula Modeling : An Introduction for Practitioners*.
- 16- R.B. Nielsen (2006)- *An Introduction to Copulas*, second edition, Springer,2005.
- 17- Frédérie Planchet (2011)- *dependance stochastique, Introduction à la théorie des copules*.