

Cours 2 :
Méthodes d'estimation des paramètres d'une copule

1. Introduction

L'estimation est une opération qui s'intéresse à quantifier l'inconnu dans le modèle . Si la loi de la variable aléatoire est connue mais paramétrée par un paramètre inconnu, ou plus on cherche donc à estimer ce paramètre et on parle alors de l'estimation paramétrique. En revanche, on parle de l'estimation non paramétrique si la loi de la variable aléatoire elle même est inconnue.

Estimer un paramètre c'est en chercher une valeur approchée en se basant sur les résultats obtenue dans un échantillon.

Dans ce cours, nous allons donner les méthodes principales pour estimer un paramètre θ . Puis nous allons donner quelques méthodes d simulation pour les copules.

2. Estimation paramétrique et semi paramétrique

Il existe beaucoup de méthodes d'estimation telle que la méthodes du maximum de vraisemblance(méthode classique), la méthodes des moments, ...

2.1. La méthode de maximum de vraisemblance :

D'après le théorème de Sklar $H(x,y)=C(F(x),G(y))$, on déduit par dérivation l'expression de la densité du vecteur (X_1, \dots, X_n) :

$$f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n)) \prod_{i=1}^n f_i(x_i)$$

où $c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}$ désigne la densité de la copule

L'expression log-vraisemblance $L(\theta)$ est définie comme suit :

$$\begin{aligned} L(\theta) &= \sum_{j=1}^n \ln f(x_j) \\ &= \sum_{j=1}^n \ln c(F_1(x_1), \dots, F_n(x_n)) \prod_{i=1}^n f_i(x_i) \\ &= \sum_{j=1}^n (\ln c(F_1(x_1), \dots, F_n(x_n)) + \ln \prod_{i=1}^n f_i(x_i)) \\ &= \sum_{j=1}^n \ln c(F_1(x_1), \dots, F_n(x_n)) + \sum_{j=1}^n \sum_{i=1}^n \ln f_i(x_i) \end{aligned}$$

On déduit l'estimateur de maximum de vraisemblance de θ , noté $\hat{\theta}_{ML}$, est donné par

$$\hat{\theta}_{ML} = \arg \max_{\theta} L(\theta)$$

$\hat{\theta}_{ML}$ est un estimateur sans biais, convergent et asymptotiquement normal sous certaines hypothèses que nous ne détaillerons par ici (Durrleman, et al .2000)

$$\sqrt{n}(\hat{\theta}_{ML} - \theta) \xrightarrow{loi} N(0, I^{-1}(\theta))$$

où $I(\theta)$ représente la matrice d'information de Fisher.

Notons que la maximisation de la fonction $L(\theta)$ est difficile lorsque le nombre de paramètre est très élevé, et peut alors engendrer des calculs très longs.

2.2. Méthodes des fonctions d'inférence des marginales (IFM) :

Cette méthode a été proposée par shib et louis(1995) dans le but d'éviter les erreurs pouvant découler de l'estimation des paramètres des lois marginales et ceux de la copules d'une façon simultanée.

L'idée de base de cette méthode est de faire cette estimation en deux étapes:

1. Estimer les paramètres γ_j des marginales
2. Etant données ces estimations, estimer ensuite le paramètre θ de la copule.

On estime d'abord les paramètres des distributions marginales univariées γ_j en maximisant $L_j(\gamma_j) = \sum_{i=1}^n \ln f_j(x_{ij})$, on obtient $\hat{\gamma}_j = \arg \max L_j(\gamma_j)$ où f_j est la densité de F_j , puis on estime θ , En utilisant les estimateurs $\hat{\gamma}_j$.

nous avons alors : $\hat{\theta}_n^{IFM} = \arg \max L(\theta)$ où $L(\theta) = \sum_{i=1}^n \ln C_\theta \{F_{\hat{\gamma}_1}(x_{i1}), \dots, F_{\hat{\gamma}_d}(x_{id})\}$

Là encore, l'estimateur $\hat{\theta}_n^{IFM}$ est également convergent, asymptotiquement normal et sans biais : l'estimateur $\hat{\theta}_n^{IFM}$ vérifie la propriété de normalité asymptotique démontré par Joe (1997): $\sqrt{n}(\hat{\theta}_n^{IFM} - \theta) \longrightarrow N(0, v^{-1}(\theta))$ avec $v(\theta)$ est la matrice d'information de Godambe, définie par : $v(\theta) = D^{-1}M(D^{-1})^t$ où $D = E \left[\frac{\partial(g(\theta))^t}{\partial \theta} \right]$, $M = E [g(\theta)^t g(\theta)]$ et $g(\theta) = \left(\frac{\partial L_1}{\partial \gamma_1}, \dots, \frac{\partial L_d}{\partial \gamma_d} \right)$

2.3. Méthode des moments

Cette méthode est notamment utilisée pour les mesures de dépendance, l'estimateur des moments de la mesure de dépendance considérée est alors simplement obtenu en égalant l'expression paramétrique (analytique) de la mesure avec un estimateur non paramétrique de cette même mesure.

Exemple 1 Pour la copule de Gumbel de paramètre θ , on a

$$\tau = 1 - \frac{1}{\theta}$$

Nous en déduisons que:

$$\theta = \frac{1}{1 - \tau}$$

Si nous avons une estimation $\hat{\tau}$ du tau de Kendall, nous pouvons obtenir une estimation du paramètre de la copule en posant :

$$\hat{\theta} = \frac{1}{1 - \hat{\tau}}$$

Dans le cas général, on a l'estimateur non paramétrique du tau de Kendall est donné par :

$$\hat{\tau} = \frac{c - d}{c + d}$$

où c et d sont respectivement le nombre de paires disjointes concordantes et discordantes

2.4. Méthode de pseudo-maximum de vraisemblance(maximum de vraisemblance canonique)

cette méthode a été proposé dans le cas ou les marginales F_1, \dots, F_d associées aux X_1, \dots, X_d sont inconnues, elle se compose de deux étapes:

-On remplace les marginales F_1, \dots, F_d par leurs estimations naturelles (estimateur empirique)définie par: $\hat{F}_{j,n}(x_{ij}) = \frac{1}{n} \sum_{i=1}^n P(X_{ij} \leq x)$

-On maximise la pseudo log-vraisemblance pour estimer θ , telle que :

$$L(\theta) = \sum_{i=1}^n \ln C_{\theta} \left\{ \hat{F}_{1,n}(x_{1i}), \dots, \hat{F}_{d,n}(x_{id}) \right\}$$

alors l'estimateur $\hat{\theta}_n^{PMV} = \operatorname{argmax} L(\theta)$

3. Simulation des copules

Nous abordons le problème de la simulation du vecteur aléatoire $X = (X_1, \dots, X_n)$ dont la distribution est $F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$: Ce problème revient à simuler le vecteur aléatoire $U = (U_1, \dots, U_n)$ dont la fonction de distribution est la copule C . Nous présentons trois méthodes pour obtenir des nombres aléatoires de C . Parfois, les marginales (F_1, \dots, F_n) ne sont pas connues analytiquement. Dans ce cas, nous utilisons la méthode dite "des quantiles empiriques".

3.1. La méthode des distributions

Cette méthode est intéressante lorsque la distribution générée par la copule est plus facile que la copule elle-même. C'est par ex. le cas de la copule gaussienne : un vecteur gaussien de dimension d est facile à simuler alors que la copule gaussienne n'est pas simple à simuler directement. Nous utilisons la transformation

$$X = (F_1^{-1}(U_1), \dots, F_n^{-1}(U_n))$$

Nous avons donc

$$F(U_1, \dots, U_n) = C(F_1^{-1}(U_1), \dots, F_n^{-1}(U_n)).$$

Pour simuler $U = (U_1, \dots, U_n)$, nous pouvons simuler le vecteur aléatoire $X = (X_1, \dots, X_n)$ de distribution F et appliquer la transformation $U = (F_1(X_1), \dots, F_n(X_n))$.

3.2. La méthode des distributions conditionnelles

Considérons le cas bivarié où $U = (U_1, U_2)$ est un vecteur aléatoire de distribution C. Nous savons que

$$P[U_1 \leq u_1] = u_1$$

et

$$P[U_2 \leq u_2 | U_1 = u_1] = C_{2|1}(u_1, u_2)$$

Comme $C(U_1, 1)$ et $C_{2|1}(u_1, U_2)$ sont deux variables aléatoires uniformes, nous obtenons l'algorithme suivant :

1. Simuler deux variables aléatoires uniformes v_1 et v_2 .
2. Prendre u_1 égal à v_1 .
3. Soit $C(u_2, u_1) = C_{2|1}(u_1, u_2)$.
4. Prendre u_2 égal à $C^{-1}(v_1, u_1)$.

Cet algorithme est suggéré par Genest et MacKay (1986). Il est utilisé par Genest (1987) pour simuler la copule Frank. Nous rappelons que la fonction de cette copule est

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

Nous en déduisons que

$$C_{2|1}(u_1, u_2, \theta) = \frac{(e^{-\theta u_2} - 1)e^{-\theta u_1}}{(e^{-\theta} - 1) + (e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}$$

Nous obtenons finalement

$$C^{-1}(u_1, u_2) = u_2 : C_{2|1}(u_1, u_2; \theta) = u = -\frac{1}{\theta} \ln \left(1 + \frac{u(e^{-\theta} - 1)}{u + (1 - u)e^{-\theta u_1}} \right)$$

La méthode bivariée s'étend sans difficultés (sur le plan mathématique) au cas multivarié. Prenons par expl le cas trivarié. Nous avons l'algorithme suivant :

1. Simuler trois variables aléatoires uniformes v_1, v_2 et v_3
2. Prendre u_1 égal à v_1
3. Soit $C(u_2, u_1) = C_{2|1}(u_1, u_2; 1)$. Prendre u_2 égal à $C^{-1}(v_2, u_1)$.
4. Soit $C(u_3, u_1, u_2) = C_{3|1,2}(u_1, u_2, u_3)$. Prendre u_3 égal à $C^{-1}(v_3, u_1, u_2)$.

Il existe plusieurs algorithmes pour simuler les copules archimédiennes. Genest et MacKay (1986) ont pris l'idée de simuler la distribution conjointe du vecteur aléatoire X en procédant par des simulations récursives des distributions conditionnelles de X_j sachant X_i .

1. Générer une variable aléatoire uniforme U_1 .
2. Poser $X_1 = F_1^{-1}(U_1)$
3. Pour $j = 2, \dots, n$, on calcule récursivement :

$$U_i = F_j(X_j | x_1, \dots, x_{j-1}) = \frac{\phi^{-1(j-1)}[c_{j-1} + \phi(F_j(x_j))]}{\phi^{-1(j-1)}(c_{j-1})}$$

où $c_j = \phi(F_1(x_1)) + \dots + \phi(F_j(x_j))$ et $\phi^{-1(j)}$ est la j^{ime} dérivée de la fonction inverse de la fonction ϕ .

4. Poser $X_j = F_j^{-1}(U_j)$

3.3. La méthode analytique

Dans ce paragraphe, la simulation est spécifique à chaque copule ou à un type de copule.

Pour simuler la copule Clayton (pour $\theta \neq 0$), nous pouvons employer l'algorithme donné par Devroye (1986) :

1. Simuler deux variables aléatoires exponentielles standards x_1 et x_2 .
2. Simuler une variable aléatoire x de fonction de distribution $\Gamma(1, \theta)$.
3. Prendre $u_1 = (1 + x_1/x)^{-\theta}$ et $u_2 = (1 + x_2/x)^{-\theta}$.

3.4. La méthode des quantiles empiriques

Le problème de cette section n'est plus la simulation du vecteur U dont la distribution est une copule C , mais concerne la simulation du vecteur X dont la copule est C et les marginales pas forcément uniformes. Dans les sections précédentes, X est simulé à partir de la transformation suivante :

$$\begin{array}{c} F_1^{-1}(U_1) \\ \vdots \\ F_n^{-1}(U_n) \end{array}$$

Cela implique la connaissance des distributions F_1, \dots, F_n . Ce qui n'est pas toujours le cas. Néanmoins, s'il est possible de simuler les marginales F_1, \dots, F_n , alors, nous pouvons simuler la distribution multidimensionnelle F grâce à la méthode des quantiles empiriques. Soit $F_{i,m}$ le processus de distribution empirique (non normalisé). Nous avons le résultat suivant :

$$\sup |F_{i,m} - F_i(x)| \xrightarrow{a.s} 0 \text{ lorsque } m \rightarrow \infty$$

Soient U_m et F_m les processus de distribution empirique correspondants aux distributions $C(u_1, \dots, u_n)$ et $F(x_1, \dots, x_n)$. En utilisant un argument de type Glivenko-Cantelli, nous avons

$$\sup |U_p(F_{1,m}(u_1), \dots, F_{n,m}(u_n))| \xrightarrow{a.s} 0, \text{ lorsque } m \wedge p \rightarrow \infty$$

avec \wedge indique le maximum.

3.5. Exemples

Copule de Clayton

```
> theta <- -5
> nsim = 2000
> U <- matrix(runif(nsim*3), nsim, 3, byrow = T)
> Theta <- -qgamma(V[,1], 1/theta, 1)
> Y <- sapply(1:2, function(t) qexp(V[,t+1], Theta))
> V <- (1+Y)**(-1/theta)
> plot(V)
```

Estimation

```
> library(copula)
> est <- fitCopula(claytonCopula(), data = V, methode = "ml")
> est
```

Copule de Frank

```
> theta <- -10
> nsim = 2000
> U <- matrix(runif(nsim*2), nsim, 2)
> V <- cbind(V[,1], -1/theta * log(1 + (V[,2] * (exp(-theta) - 1)) / (exp(-theta * V[,1]) * (1 - V[,2]) + V[,2])))
> plot(V)
```

```
Estimation > library(copula)
> est <- fitCopula(frankCopula(), data=V, method="mpl")
> est
```

Copule normale

```
> theta = 0.8
> library(MASS)
> U <- rnorm(nsim, c(0,0), matrix(c(1, theta, theta, 1), 2))
> V <- sapply(1:2, function(t) pnorm(V[,t], 0, 1))
> plot(V)
```

Estimation

```
> est <- fitCopula(normalCopula(), data=V, method="ml")
> est
```

Copule d'indépendance

```
> U <- matrix(runif(2*nsim), nsim, 2, byrow=T)
> plot(U)
```

Fonction de répartition

```
> c <- function(U1, U2) U1 * U2
```

Copule borne supérieure de Fréchet

```
> U <- matrix(runif(nsim), nsim, 2)
> plot(U)
```

Fonction de répartition

```
> c <- function(u1, u2)min(u1, u2)
```

Copule borne inférieure de Fréchet

```
> U <- matrix(runif(nsim), nsim, 2)
```

```
> V[,2] <- 1-V[,1]
```

```
> plot(V)
```

Fonction de répartition

```
> c <- function(u1, u2)max(u1 + u2 - 1, 0)
```