Estimation des paramètres par la méthode composite

Sous la supervision de
Hélène Cossette
Étienne Marceau

Préparé par

Alexandre Lepage,
Diamilatou N'diaye,

le 03 Juillet 2019

Faculté des sciences et de génie
École d'actuariat
Université Laval

1 Estimation par la Méthode Composite

Le but de cette section est d'estimer les paramètres d'une copule archimédienne hiérarchique par la méthode composite. Dans les exemples suivants $N \sim Binom(5, q),\ q = 0.2$ et $X \sim Exp(\beta),\ \beta = 1/100$. La copule hiérarchique étudié ici est de loi mère géométrique de paramètre $\alpha_0 = 0.6$ et de loi enfant géométrique de paramètre $\alpha_1 = 0.4$. Les estimations sont réalisées sur un échantillon simulé de 10 000 observations.

2 Première Approche

La première approche consiste à estimer séparément tous les paramètres. Dans un premier temps, les paramètres q et β sont estimés pas les estimateurs du maximum de vraisemblance de leur loi respective.

$$\hat{q} = \frac{\sum_{i=0}^{10\,000} n_i}{5 \times 10\,000}$$

$$\hat{\beta} = \frac{n_{tot}}{\sum_{i=0}^{n_{tot}} X_i}$$

où n_{tot} représente le nombre de X_i présents dans l'échantillon.

Ensuite l'estimateur de α_0 est obtenu en minimisant la fonction de vraisemblance :

$$L(\alpha_0|\hat{q}, \hat{\beta}) = (F_N(0))^{n_0} \prod_{i=1}^{n_{NX}} \left(\frac{\partial}{\partial x} C(F_N(n_i), F_X(x_i)) - \frac{\partial}{\partial x} C(F_N(n_i - 1), F_X(x_i)) \right)$$

où n_{NX} représente le nombre de couples $(N_i, X_{i,j})$ présents dans l'échantillon.

De plus, avec le modèle collectif du risque définit par $S = \sum_{k=1}^{\infty} X_k \times \mathbb{1}\{k \leq N\}$, la copule reliant la v.a. N avec n'importe lequel des X_i s'exprime comme

$$C(u_0, u_1) = \mathscr{L}_M(\mathscr{L}_M^{-1}(u_0) + \mathscr{L}_M^{-1}(u_1)). \tag{1}$$

À noter que la copule décrite en (1) est archimédienne.

Et enfin, de la même manière que pour α_0, l'estimateur de α_1 est obtenu à partir de la fonction de vraisemblance

$$L(\alpha_1|\hat{\beta}, \hat{\alpha_0}) = \prod_{i=0}^{n_{XX}} \frac{\partial^2}{\partial x_1 \partial x_2} C(F_X(x_1), F_X(x_2))$$

où n_{XX} représente le nombre de couples $\{(X_{i,k}, X_{i,j}) : k \neq j\}$ présents dans l'échantillon.

$$C(u_1, u_2) = \mathscr{L}_\Theta(\mathscr{L}_\Theta^{-1}(u_1) + \mathscr{L}_\Theta^{-1}(u_2))$$

avec

$$\mathscr{L}_\Theta(t) = \mathscr{L}_M(-\ln(\mathscr{L}_B(t)))$$

et

$$\mathscr{L}_\Theta^{-1}(u) = \mathscr{L}_B^{-1}\left(exp(-\mathscr{L}_M^{-1}(u))\right).$$

Le tableau 1 présente les valeurs des estimateurs ;

	q	β	α_0	α_1
Valeur	0.2000	0.0100	0.6000	0.4000
Estimateur	0.1996	0.0087	0.3603	0.8545
Temps	0.0000	0.0000	1.3900	0.7900

Tableau 1 – Valeurs estimées de q, β, α_0, α_1 ainsi que le temps de calcul par la première approche.

On voit en regardant le tableau 1 que les résultats de l'estimation sont loins des vrais paramètres.

2.1 Comportement Asympotique

La méthode précédente est répétée 100 fois afin d'observer le comportement asymptotique des estimateurs. Pour α_0, on obtient une moyenne de 0.3556106 et un intervalle de confiance de niveau $\alpha = 95\%$ de $[0.3117377, 0.3994836]$.

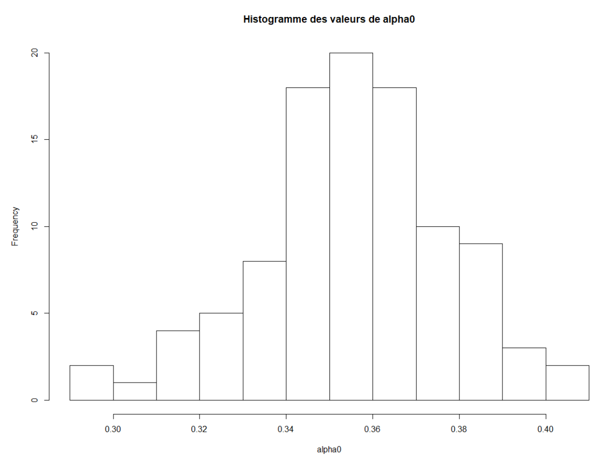

Illustration 1 – Histogramme des valeurs prises par α_0 sur 100 répétitions

Pour α_1 on obtient une moyenne de 0.8201792 et un intervalle de confiance de niveau $\alpha = 95\%$ de $[0.7704843, 0.8698740]$.

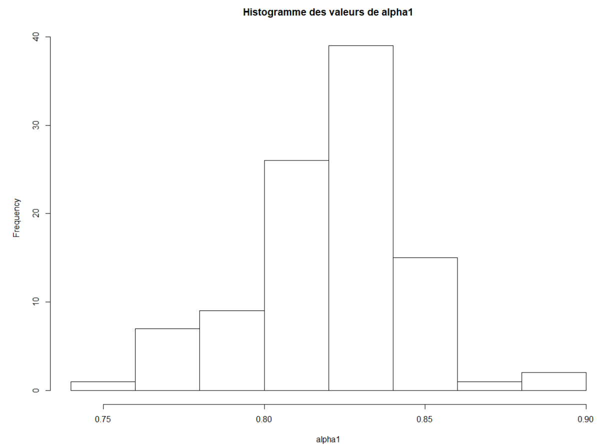

Illustration 2 – Histogramme des valeurs prises par α_1 sur 100 répétitions

2.1.1 Comportement asymptotique des estimateurs pour une copule archimédienne hiérarchique Logarithmique-Gamma de paramètres $(\alpha_0 = 6, \alpha_1 = 4)$

L'estimateur de α_0 produit une moyenne de 3.232702 et un intervalle de confiance de niveau $\alpha = 95\%$ de $[3.114684, 3.350721]$.

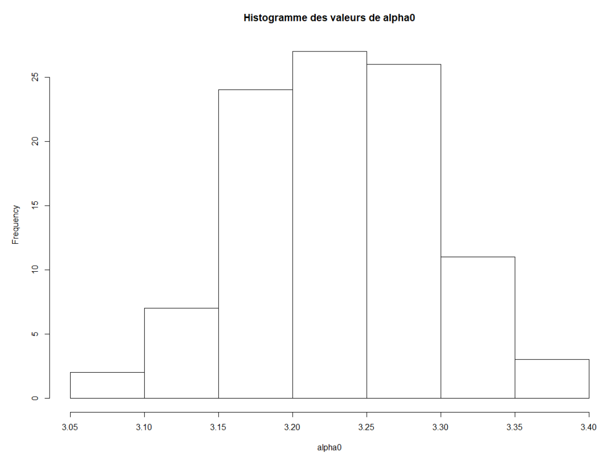

Illustration 3 – Histogramme des valeurs prises par α_0 sur 100 répétitions

L'estimateur de α_1 produit une moyenne de 6.646969 et un intervalle de confiance de niveau $\alpha = 95\%$ de $[6.094799, 7.199139]$.

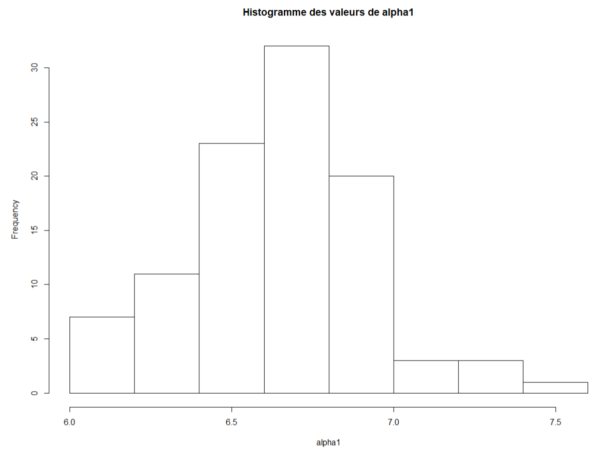

Illustration 4 – Histogramme des valeurs prises par α_1 sur 100 répétitions

Les estimations sont donc assez loin des vrais paramètres.

3 Seconde Approche

La seconde approche consiste à estimer q, β et α_0 en même temps puis α_1 par la suite. La fonction de vraisemblance à minimiser est :

$$L(\alpha_0, q, \beta) = (F_N(0))^{n_0} \prod_{i=0}^{n_{NX}} \left(\frac{\partial}{\partial x} C(F_N(n), F_X(x)) - \frac{\partial}{\partial x} C(F_N(n-1), F_X(x)) \right)$$

où n_{NX} représente le nombre de couples $(N_i, X_{i,j})$ présents dans l'échantillon.

$$C(u_0, u_1) = \mathscr{L}_M(\mathscr{L}_M^{-1}(u_0) + \mathscr{L}_M^{-1}(u_1))$$

Pour ce qui est de α_1, la même estimation que la section précédente est réalisée.

$$L(\alpha_1 | \hat{\beta}, \hat{\alpha_0}) = \prod_{i=0}^{n_{XX}} \frac{\partial^2}{\partial x_1 \partial x_2} C(F_X(x_1), F_X(x_2))$$

$$C(u_1, u_2) = \mathscr{L}_\Theta(\mathscr{L}_\Theta^{-1}(u_1) + \mathscr{L}_\Theta^{-1}(u_2))$$

avec

$$\mathscr{L}_\Theta(t) = \mathscr{L}_M(-\ln(\mathscr{L}_B(t)))$$

où n_{XX} représente le nombre de couples $(X_{i,k}, X_{i,j})$ présents dans l'échantillon

Le tableau 1 présente les valeurs des estimateurs ;

	q	β	α_0	α_1
Valeur	0.2000	0.0100	0.6000	0.4000
Estimateur	0.3589	0.0087	0.4972	0.8000
Temps	45.6000	45.6000	45.6000	1.4100

Tableau 2 – Valeurs estimées de q, β, α_0, α_1 ainsi que le temps de calcul par la deuxième approche. Le temps de calcul de q, β et α_0 est leur temps de calcul commun étant donné qu'ils sont estimés ensembles.

Dans le tableau 2, on voit que le niveau de précision est légèrement meilleur que pour la première approche lorsque l'on regarde α_0 et α_1, bien que ce ne soit toujours pas satisfaisant. Tandis que pour $\hat{q}$, les résultats se détériorent.

3.1 Comportement Asympotique

L'estimation est ensuite répétée 100 fois afin d'observer le comportement asymptotique des estimateurs. L'estimateur de α_0 produit une moyenne de 0.4856781 et un intervalle de confiance de niveau $\alpha = 95\%$ de $[0.3888275, 0.5825286]$.

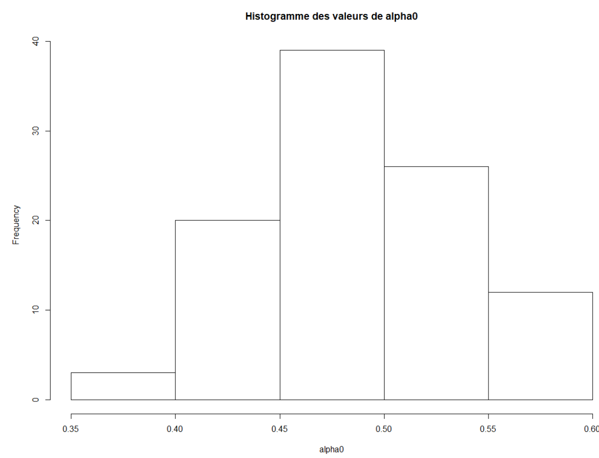

Illustration 5 – Histogramme des valeurs prises par α_0 sur 100 répétitions

L'estimateur de α_1 produit une moyenne de 0.7690707 et un intervalle de confiance de niveau $\alpha = 95\%$ de $[0.7103108, 0.8278306]$.

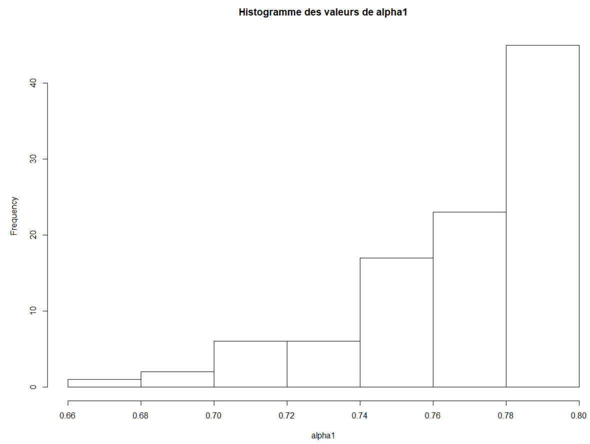

Illustration 6 – Histogramme des valeurs prises par α_1 sur 100 répétitions