Christophe Dutang christophe.dutang@grenoble-inp.fr

TP3 - Modélisation de la survie d'un couple et calcul de rente viagère

Le TP est à rendre sur Teide sous la forme d'un fichier Rmarkdown `NomTrinome.Rmd`. Merci de vous assurer que le fichier compile (en un temps raisonnable).

Nous analysons des données de survie issues de contrat d'assurance vie d'un assureur canadien entre 1988 et 1993. Ces contrats portent sur 14 889 rentes viagères de couple homme-femme. Ces données sont complètes : nous observons l'âge de décès de chaque individu. On note T^1 la durée de vie aléatoire de l'individu 1 et T^2 la durée de vie aléatoire de l'individu 2. Le jeu de données `canlifeins` comporte deux colonnes `DeathAgeM` et `DeathAgeF`. Pour ce TP, nous utiliserons les packages `copula`, `fitdistrplus` et `flexsurv`.

1. Chargez les données par `canlifeins <- read.csv("canlifins.csv")`.
2. Numériquement et graphiquement montrer la dépendance entre T^1 et T^2.

On considère le modèle paramétrique pour T^1 et T^2

$$H(t_1, t_2) = P(T_1 \leq t_1, T_2 \leq t_2) = C(H_1(t_1; \theta_1), H_2(t_2; \theta_2); \alpha)$$

où H_1 est la distribution marginale de T^1 et H_2 celle de T^2 de paramètre θ_1 et θ_2 respectivement; α le paramètre de la copule.

3. Utiliser la méthode de maximisation de la vraisemblance canonique sur ces données. On pourra considérer les copules suivantes : gaussienne, student, Clayton, Frank, Gumbel et Joe. Voir `archmCopula`, `ellipCopula` du package `copula`.
4. Utiliser la méthode d'inférence sur les marginales sur ces données. Pour les lois marginales, on peut essayer les lois lognormal, Weibull, Gompertz, gamma, loglogistique, voir package `flexsurv` et `fitdistrplus`. On pourra considérer les copules suivantes : gaussienne, student, Clayton, Frank, Gumbel et Joe.
5. Conclure sur la copule choisie.

Considérons une rente viagère jusqu'au dernier survivant définie par

$$\ddot{a}_{\overline{xy}} = \sum_{k=0}^{+\infty} v^k {}_k p_{\overline{xy}}, \text{ avec } {}_k p_{\overline{xy}} = P(T_{\overline{xy}} > k | T_{\overline{xy}} > 0)$$

où $T_{\overline{xy}} = \max(T_x^1, T_y^2)$ La distribution jointe $H_{x,y}$ de $T_x^1 = T^1 - x$ et $T_y^2 = T^2 - y$ conditionnellement à $T_x^1 > 0$ et $T_y^2 > 0$ s'exprime en fonction de la distribution jointe H de T^1 et T^2

$$H_{x,y}(t_1, t_2) = \frac{P(0 < T_x \leq t_1, 0 < T_y \leq t_2)}{P(T_x > 0, T_y > 0)} = \frac{H(x + t_1, y + t_2) - H(x, y + t_2) - H(x + t_1, y) + H(x, y)}{1 - H_1(x) - H_2(y) + H(x, y)}.$$

Ainsi

$${}_k p_{\overline{xy}} = 1 - H_{x,y}(k, k).$$

6. Dans le cas indépendance, déterminer $\ddot{a}_{\overline{xy}}$ pour différents x, y
7. Utilisez votre meilleur copule pour déterminer $\ddot{a}_{\overline{xy}}$ pour différents x, y.
8. Commenter