UNIVERSITÉ DE BRETAGNE OCCIDENTALE Année 2018 - 2019
EURIA - Master 2

Modèles de durée et tables de mortalité
Examen du 21 décembre 2018 - Durée : 2h
Les codes R commentés seront récupérés sur clé USB.
L'utilisation d'Internet est interdite durant l'épreuve.

Exercice 1. Les données du data.frame **jasa** du package **survival** concernent la survie de patients acceptés pour une transplantation cardiaque à l'hopital de Stanford entre fin 1967 et début 1974. Dans le fichier étudié, les 103 patients sont suivis jusqu'au 1$^{\text{er}}$ avril 1974. Une fois le patient accepté dans le programme, il est mis sur une liste d'attente jusqu'à ce qu'un donneur potentiel soit trouvé. Si le patient est toujours en vie à ce moment là, la transplantation est effectuée. Le data.frame **jasa** contient les variables suivantes:

birth.dt : Date de naissance
accept.dt : Date d'entrée dans le programme
tx.date : Date de transplantation
fu.date : Date de sortie du programme
fustat : Mort=1/vivant=0
surgery : Pontage antérieur oui=1/non=0
age : Age (en jours) à l'entrée du programme
futime : Temps de suivi
wait.time : Temps écoulé (en jours) entre l'entrée dans le programme et la transplantation
transplant : Transplantation effectuée oui=1/non=0
mismatch : Score d'incompatibilité
hla.a2 : Type particulier d'incompatibilité
mscore : Autre score d'incompatibilité
reject : Rejet de la greffe oui=1/non=0

La durée de vie considérée est ici le nombre de jours (*futime*) s'écoulant entre l'entrée dans le programme et la mort (ou la sortie) du patient.

1. Tracer (sur ordinateur) l'estimateur de Kaplan-Meier $\hat{S}$ de la fonction de survie des individus, en prenant en compte les censures éventuelles, selon qu'ils ont été transplantés ou non. Que remarquez-vous ?

   Comparer les survies des patients transplantés et des patients non transplantés à l'aide d'un test statistique.

2. Ajuster la durée de survie avec un modèle de Cox en considérant les variables explicatives *age*, *surgery* et *transplant*. L'influence de ces variables est-elle significative? (répondre par un test statistique)

3. Ajuster la durée de survie avec un modèle de Cox en considérant les variables explicatives *age* et *transplant*. Interpréter l'influence de ces variables dans le modèle, en fonction des valeurs des paramètres obtenues.

   L'influence de ces variables vous semble-t-elle constante dans le temps ?

4. Ajuster la durée de survie avec un modèle de Cox stratifié par rapport à la variable *transplant* en considérant la variable explicative *age* et tracer les deux courbes de survie correspondantes.

5. Sur l'ensemble des patients ayant subi une greffe (transplant=1), étudier le modèle de Cox de la durée de vie avec variables explicatives *age, surgery, wait.time, mismatch, mscore, hla.a2* et *reject*.

6. Analyser cette affirmation : "*Afin de mieux analyser si la transplantation cardiaque prolonge la vie d'un patient, il est préférable de créer un nouveau tableau de données dans lequel :*

   - chaque patient transplanté a deux lignes, une correspondant à la période depuis son entrée jusqu'à sa transplantation, l'autre correspondant à la période depuis sa transplantation jusqu'à sa sortie.

   - chaque patient non transplanté a une ligne, une correspondant à la période globale depuis son entrée dans le programme jusqu'à sa sortie."

Exercice 2. Le data.frame **EWMaleData** du package **StMomo** contient le nombre d'hommes vivants et le nombre d'hommes décédés par âge en Angleterre et Pays de Galle durant la période 1961-2011, pour les âges de 0 à 100 ans (Human Mortality Database, 2014).

Définissons le taux de mortalité $q_x(t)$ et le taux instantané de mortalité $\mu_x(t)$ à l'âge x et l'année t. Nous ferons l'approximation $q_x(t) \simeq \mu_x(t)$.

1. Calibrer les taux bruts de mortalité avec un modèle de Makeham-Gompertz de fonction de hasard
   $$h(x) = \alpha + \beta\ \gamma^x$$
   pour les données de l'année 2011 (en utilisant la fonction *fitGM* ou un autre code).

   Quelles valeurs obtenez-vous pour les paramètres estimés $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$?

   Comparer graphiquement les taux bruts observés et les taux de mortalité discrets du modèle de Gompertz-Makeham correspondant aux valeurs $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$.

2. A partir des données **EWMaleData**, calibrer un modèle de Lee-Carter de taux instantanés de mortalité $\mu_x(t) = \exp(\alpha_x + \beta_x k_t)$ pour les années de 1961 à 2011 et les âges de 0 à 100 ans.

   Tracer sur le graphique de la question précédente les valeurs de $\mu_x(2011)$ et analyser les résultats obtenus.

3. Nous souhaitons estimer des capitaux constitutifs d'une rente viagère de la forme :
   $$a_x(t) = \sum_{i \geq 0} (\frac{1}{1+r})^{i+1} \prod_{j=0}^{i} \exp(-\mu_{x+j}(t+j))$$
   à l'année $t = 2012$ et l'âge $x = 65$ ans. Nous considérons que l'âge maximal pris en compte dans la formule est 100 ans et que le taux d'actualisation est $r = 0.035$.

   - Projeter le modèle de Lee Carter et en déduire une estimation de $a_{65}(2012)$. Préciser le modèle de projection choisi.

   - Comparer avec l'estimation obtenue en estimant la mortalité avec le modèle de Gompertz-Makeman calibré à la question 1. Nous considérons dans ce cas que la mortalité dépend seulement de l'âge et pas des années considérées dans le futur.

   - Comparer et interpréter les résultats obtenus. Quelle méthode est la plus prudente ?