Modèles de durée / Examen du 15 janvier 2012

Durée 2h – aucun document n'est autorisé

Corrigé

Problème : construction d'un modèle continu à partir d'un modèle discret

La qualité de la rédaction, des justifications apportées et de la présentation de la copie seront prises en compte dans la notation.

On suppose données pour x entier des probabilités conditionnelles de décès q_x.

Question n°1 (1 point) : Rappelez le lien entre les probabilités q_x et la fonction de survie du modèle $S(x)$. La donnée de q_x, $x = 1, 2, \ldots$ suffit-elle à définir complètement S ?

> On observe que $q_x = 1 - \frac{S(x+1)}{S(x)}$; non, puisqu'on ne définit ainsi S qu'aux points entiers.

Question n°2 (2 points) : Rappelez la définition de la fonction de survie conditionnelle $S_x(u)$ et son expression en fonction de S.

> $S_x(u) = P(X > x + u | X > x) = \frac{S(x+u)}{S(x)}$

Question n°3 (2 points) : On suppose que pour x entier et $u \in [0,1[$, $S_x(u) = (1 - q_x)^u$. Quelle est la fonction de hasard du modèle sur $[x, x+1[$? Quelle propriété remarquable présente-t-elle ?

> $h(u) = -\frac{d}{du} \mathbf{ln}\, S_x(u) = -\mathbf{ln}(1 - q_x) = \lambda_x$. Elle est constante.

Question n°4 (2 points) : Rappelez l'expression de la vraisemblance d'un modèle de durée en présence de censure aléatoire droite. Quelles sont les hypothèses sur la censure pour que cette expression soit valable ?

> Avec les notations utilisées en cours, $\mathbf{ln}\, L(y_1, \ldots, y_n) \propto \sum_{i=1}^{n} \left(\mathbf{ln}\, S_x(t_i) + d_i\, \mathbf{ln}\, h(t_i) \right)$;
> on suppose la censure indépendante de X et sa loi sans paramètre commun avec celle de X.

Question n°5 (2 points) : Donnez un exemple simple de censure aléatoire droite informative.

> $S_C(x) = S_X(x)^{\beta}$ où β est un paramètre (positif) à estimer.

Question n°6 (2 points) : Comment adapter l'expression obtenue à la question 4 en présence de troncature gauche ? La formule sera justifiée avec soin.

$$\mathbf{ln}\, L(y_1,\ldots,y_n) \propto \sum_{i=1}^{n}\left(\mathbf{ln}\, S_x(t_i) - \mathbf{ln}\, S_x(e_i) + d_i\, \mathbf{ln}\, h(t_i)\right)$$

Question n°7 (3 points) : A l'aide des questions 3 et 4, écrire la vraisemblance du modèle en présence de censure aléatoire droite sur l'intervalle $[x, x+1[$. Quelle est la loi du nombre de sorties sur cet intervalle ?

$\mathbf{ln}\, L(y_1,\ldots,y_n) = cste + \sum_{i=1}^{n}\left(\mathbf{ln}\, S_x(t_i) + d_i\, \mathbf{ln}\, h(t_i)\right) \propto \sum_{i=1}^{n}\left(\lambda_x \times t_i + d_i\, \mathbf{ln}\, \lambda_x\right)$ et donc on peut considérer que la loi du nombre de sorties est une loi de Poisson de paramètre $\lambda_x \times \sum_{i=1}^{n} t_i$.

Question n°8 (3 points) : Donnez l'expression de la fonction de survie du modèle S en un point $x+u$ avec x entier et $u \in [0,1[$. S est-elle continue ?

En revenant à la définition de la fonction de survie conditionnelle, $S(x+u) = S(x) \times S_x(u) = \prod_{y=0}^{x-1}(1-q_y) \times (1-q_x)^u$. S est continue puisque $S(x+1) = \prod_{y=0}^{x}(1-q_y)$.

Question n°9 (3 points) : Proposez une autre méthode d'extrapolation de la fonction de survie qui vous semblerait justifiable en expliquant pourquoi.