Modèles de durée / Examen du 13 février 2006

Durée 1,5h – aucun document n'est autorisé

Problème (modèle logistique)

Dans le cadre de la construction d'une table d'expérience pour des contrats en cas de décès, on s'intéresse ici à l'ajustement de taux de mortalité bruts $\hat{q}_x$ par un modèle logistique. La fonction logistique est par définition $\mathbf{lg}(x) = \mathbf{ln}\left(\frac{x}{1-x}\right)$.

Question n°1

Donner l'intervalle de définition de $x \to \mathbf{lg}(x)$ et calculer les dérivées première et seconde, ainsi que l'inverse. Que peut-on en conclure ?

Question n°2

On fait l'hypothèse que l'on a estimé le taux de décès à l'âge *x* par $\hat{q}_x$ supposé dans biais ; on se propose d'effectuer un ajustement des $\hat{y}_x = \mathbf{lg}(\hat{q}_x)$. Prouvez que l'estimateur $\hat{y}_x$ est biaisé pour estimer $\mathbf{lg}(q_x)$ et indiquez le sens du biais (on pourra supposer que les taux à estimer sont inférieurs à ½) . Qu'en concluez-vous ?

On rappelle l'inégalité de Jensen pour une fonction convexe : $f(EX) \le Ef(X)$.

Question n°3

On suppose une relation affine entre l'âge et le logit du taux de décès ; en d'autres termes on suppose que $\mathbf{lg}(q_x) = a + bx$ pour des paramètres *a* et *b* inconnus. Proposez un modèle simple et « naturel » pour estimer les paramètres *a* et *b* et décrivez en les principales propriétés.

Question n°4

Montrez que la paramétrisation $\mathbf{lg}(q_x) = a + bx$ peut s'écrire $q_x = \frac{ce^{dx}}{1+ce^{dx}}$ avec des paramètres *c* et *d* que l'on précisera en fonction de *a* et *b*.

Déterminez la fonction de survie et la fonction de hasard de ce modèle.

Question n°5

En supposant que l'estimateur $\hat{q}_x$ est déterminé dans le cadre d'un modèle binômial et que les effectifs sont suffisants pour utiliser l'approximation normale, écrivez la log-vraissemblance du modèle.

On rappelle que la densité d'une variable gaussienne d'espérance m et de variance σ^2 s'écrit : $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathbf{exp}\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right)$.

Question n°6

En remplaçant la varaiance théorique par la variance estimée dans la formule précédente, proposez une approximation de la log-vraissemblance déterminée à la question précédente qui ramène la recherche du maximum de vraisemblance à un problème de moindres carrés non linéaires.

Question n°7

Proposez une méthode de détermination des valeurs initiales des paramètres c et d que l'on pourrait utiliser dans un algorithme numérique.