# Modèles de durée / Examen du 8 janvier 2008

**Durée 2h – aucun document n'est autorisé**

## Problème : un exemple de censure informative

On considère une situation de censure aléatoire droite :

$$T_i = X_i \wedge C_i \text{ et } D_i = \begin{cases} 1 \; si \; X_i \leq C_i \\ 0 \; si \; X_i > C_i \end{cases}$$

dans laquelle on suppose les censures $C_i$ indépendantes des durées $X_i$ et on rappelle que la vraisemblance de l'échantillon $(T_1, D_1), .., (T_n, D_n)$ s'écrit, avec des notations évidentes :

$$L(\theta) = \prod_{i=1}^{n} \left[ f_X(T_i, \theta) S_C(T_i, \theta) \right]^{D_i} \left[ f_C(T_i, \theta) S_X(T_i, \theta) \right]^{1-D_i} .$$

On fait l'hypothèse qu'il existe $\beta > 0$ tel que $S_C(x) = S_X(x)^\beta$, pour tout $x \geq 0$.

**Question n°1 (1 point)** : Calculer la densité de *C* en fonction de $f_X$ et $S_X$ ; en déduire la fonction de hasard de *C* en fonction de celle de *X*.

**Question n°2 (2 points)** : Déterminer la loi de $D = 1_{\{X \leq C\}}$ ; en déduire un estimateur simple de $\beta$.

**Question n°3 (3 points)** : Préciser la loi limite de l'estimateur défini ci-dessus ; donner un intervalle de confiance asymptotique au niveau $\alpha$ (on rappelle la méthode delta : $\mathbf{V}(f(X)) \approx \left( \frac{df}{dx}(\mathbf{E}(X)) \right)^2 \mathbf{V}(X)$).

**Question n°4 (1 points)** : Calculer la fonction de survie de *T* en fonction de $S_X$ et $\beta$ ; en déduire une alternative à l'estimateur de Kaplan-Meier de $S_X$.

**Question n°5 (1,5 points)** : Commenter les cas particuliers du modèle $\beta = 1$, $\beta \to 0$ et $\beta \to +\infty$.

A partir de maintenant on suppose que $S_X(t) = \mathbf{exp}(-\theta t)$, *ie X* suit une loi exponentielle de paramètre $\theta$.

**Question n°6 (4 points)** : Ecrire la log-vraisemblance de $(\theta,\beta)$ en fonction de $n$, $\overline{D}=n^{-1}\sum D_i$ et $\overline{T}=n^{-1}\sum T_i$. En déduire les estimateurs du maximum de vraisemblance de :

- $(\theta,\beta)$ lorsque les 2 paramètres sont inconnus ;
- $\theta$ lorsque $\beta$ est connu ;
- $\beta$ lorsque $\theta$ est connu.

**Question n°7 (2,5 points)** : Calculer l'espérance et la variance de $\hat{\theta}$ lorsque $\beta$ est connu ; on rappelle que lorsque $Z$ suit une loi $\gamma(r,\lambda)$, $\mathbf{E}\left(Z^p\right)=\lambda^{-p}\dfrac{\Gamma(r+p)}{\Gamma(r)}$, pour $p>-r$.

**Question n°8 (2,5 points)** : Calculer l'information de Fisher sur $(\theta,\beta)$ ; on rappelle que dans un modèle paramétrique, l'information de Fisher est définie par $I(\omega)=-\mathbf{E}\left(\dfrac{\partial^2 \mathbf{ln}\, L(\omega)}{\partial\omega\partial\omega'}\right)$.

**Question n°9 (2,5 points)** : En déduire la loi limite de la statistique

$$Q_n(\theta,\beta)=n\left(\left(\frac{\hat{\theta}}{\theta}-1\right)^2+\frac{\beta}{\beta+1}\left(\frac{\hat{\beta}}{\beta}-1\right)^2+\frac{2\beta}{\beta+1}\left(\frac{\hat{\theta}}{\theta}-1\right)\left(\frac{\hat{\beta}}{\beta}-1\right)\right)$$