Université de Strasbourg | Ségolen Geffray
M1 - Magistère | Année 2014/2015
Statistique - étude de cas | Lundi 11 mai 2015

Contrôle continu

La durée de l'épreuve est 2h.
Les téléphones portables doivent être éteints et rangés.
Les notes de cours (comme tout autre document) et les calculatrices ne sont PAS autorisées.
La qualité de la rédaction sera largement prise en compte.

Rappels sur la fonction Gamma :

$$\Gamma(u) = \int_0^\infty e^y y^{u-1} dy, \quad u > 0.$$

La fonction $\Gamma(.)$ est c^∞ sur $\mathbb{R}^{*+}$ et ses dérivées sont données par :

$$\Gamma^{(k)}(u) = \int_0^\infty (\log y)^k \, e^y y^{u-1} dy, \quad u > 0.$$

En voici quelques valeurs remarquables :

$$\Gamma(1) = 1, \qquad \Gamma'(1) = -\gamma, \qquad \Gamma''(1) = \gamma^2 + \frac{\pi^2}{6}$$
$$\Gamma(2) = 1, \qquad \Gamma'(2) = 1-\gamma, \qquad \Gamma''(2) = (1-\gamma)^2 + \frac{\pi^2}{6} - 1$$

exprimées au moyen de la constante d'Euler notée γ :

$$\gamma = -\int_0^\infty \log(y) \, e^y dy \approx 0.577\,215$$

Exercice 1.
Soit $(X_1, ..., X_n)$ un échantillon i.i.d. de variable parente X de loi de Gumbel définie par sa fonction de répartition $F_{\alpha,\beta}$, paramétrée par $\theta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ avec $\alpha \in \mathbb{R}$ et $\beta > 0$, de la façon suivante :

$$F_{\alpha,\beta}(x) = \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right), \quad x \in \mathbb{R}.$$

La loi de Gumbel est très fréquemment utilisée en hydrologie, par exemple pour modéliser le niveau maximal annuel de crue d'un fleuve.

1. Vérifier qu'il s'agit bien d'une fonction de répartition. Justifier l'existence d'une densité. Déterminer l'expression de cette densité. Le modèle est-il homogène ? Appartient-il à la famille exponentielle ?

2. Calculer les quantités suivantes :

$$\mathbb{E}_\theta\left[\left(\frac{X-\alpha}{\beta}\right)^k\right] \qquad \text{pour } k=1,2,3,4 \qquad \text{et} \qquad \mathrm{Var}_\theta\left(\frac{X-\alpha}{\beta}\right).$$

3. Déterminer l'estimateur de $\theta = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ par la méthode des moments. On notera $\widetilde{\theta}_n = \begin{pmatrix}\widetilde{\alpha}_n\\ \widetilde{\beta}_n\end{pmatrix}$ cet estimateur.
4. Etudier son comportement asymptotique.
5. Montrer que l'estimateur du maximum de vraisemblance de β satisfait l'équation suivante :

$$\sum_{i=1}^n (X_i - \overline{X}_n + \beta)\exp\left(-\frac{X_i}{\beta}\right) = 0$$

en notant $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Quelle est l'équation satisfaite par l'estimateur du maximum de vraisemblance de α ?
On notera $\widehat{\theta}_n = \begin{pmatrix}\widehat{\alpha}_n\\ \widehat{\beta}_n\end{pmatrix}$ l'estimateur du maximum de vraisemblance de θ.
6. Calculer pour $k=0,1,2$:

$$\mathbb{E}_\theta\left[\left(\frac{X-\alpha}{\beta}\right)^k \exp\left(-\frac{X-\alpha}{\beta}\right)\right].$$

7. En admettant que le modèle est régulier, calculer la matrice d'information de Fisher (dans le modèle probabiliste à une observation).
8. Que peut-on dire du comportement asymptotique de $\widehat{\theta}_n$ si l'on sait trouver une solution au problème d'optimisation du maximum de vraisemblance qui soit consistante pour θ (toujours en admettant que le modèle est régulier) ?
9. On se donne une probabilité $p \in]0,1[$ (par exemple 10^{-4}). Proposer un estimateur de x_p défini par la relation $\mathbb{P}(X > x_p) = p$. Justifier.
10. On considère dans cette question le cas particulier où $\beta = 1$.
    (a) Le modèle est-il exponentiel ?
    (b) Déterminer une statistique exhaustive pour α dans le modèle à n observations. Est-elle complète ? minimale ?
    (c) Déterminer l'estimateur du maximum de vraisemblance de α noté $\widehat{\alpha}_n$.
    (d) Etudier son comportement asymptotique.
    (e) Construire un intervalle de confiance bilatéral symétrique asymptotique pour α au niveau de confiance 95%.