TD 8

Exercice 1 Estimation de la variance
On considère un modèle linéaire non-gaussien. On veut déterminer l'espérance de $\hat{\sigma}^2$.
1. Montrer que $(n-k)\hat{\sigma}^2 = Tr(\varepsilon' \cdot P_{[X]^\perp}\varepsilon)$ où Tr désigne la trace.
2. En utilisant le fait que $Tr(A \cdot B) = Tr(B \cdot A)$, montrer que

$$(n-k)\mathbb{E}(\hat{\sigma}^2) = Tr(P_{[X]^\perp}\mathbb{E}(\varepsilon \cdot \varepsilon')).$$

3. Conclure.

Exercice 2 Maximum de vraisemblance
On rappelle que, dans le cadre d'un modèle statistique paramétrique (c'est-à-dire que la loi du modèle ne dépend que d'un nombre fini de paramètres), la densité des observations $Y_1, \ldots, Y_n$ vue comme une fonction des paramètres est appelée la vraisemblance. L'estimateur du maximum de vraisemblance de ces paramètres est la valeur des paramètre qui maximise la vraisemblance.

a) Considérons le modèle gaussien

$$Y_i = \mu + \beta X_i + \varepsilon_i, \quad i = 1, \ldots, n$$

et ses paramètres (μ, β, σ^2). Montrer que $-2 \times$ log-vraisemblance vaut

$$L(\mu, \beta, z) = n\log(2\pi) + n\log z + \frac{1}{z}\sum_{i=1}^{n}(Y_i - \mu - \beta X_i)^2,$$

en notant $z = \sigma^2$ pour dériver plus facilement.

b) Comparer les estimateurs du maximum de vraisemblance avec ceux des moindres carrés de μ et β.

c) Comparer (au sens de la vitesse de convergence) l'estimateur du maximum de vraisemblance de σ^2 avec $\hat{\sigma}^2$ défini dans le cours.

Exercice 3 Test du rapport de vraisemblance
On rappelle que dans le cadre de variables aléatoires $Y_1, \ldots, Y_n$ suivant un modèle statistique admettant une densité, la statistique U du rapport de vraisemblance est le rapport entre le maximum (en les paramètres) de la vraisemblance sous l'hypothèse H_0 et le maximum (en les paramètres) de la vraisemblance sous l'hypothèse H_1. La région de rejet du test est du type $U \leq \lambda$, avec $\lambda \in \mathbb{R}$ dépendant du niveau du test.

a) On considère le test du rapport de vraisemblance de $H_0 : \beta = 0$ contre $H_1 : \beta \neq 0$ dans le cadre du modèle gaussien de l'exercice 1. Montrer que si l'on pose

$$SC := \sum_{i=1}^{n}(Y_i - \hat{\mu} - \hat{\beta}X_i)^2,$$

alors le maximum sous H_1 de la log-vraisemblance vaut

$$-\frac{1}{2}(n\log(2\pi) + n\log(\frac{SC}{n}) + n).$$

b) Montrer un résultat équivalent sous H_0 en posant

$$\tilde{SC} := \sum_{i=1}^{n} (Y_i - \bar{Y})^2.$$

c) Donner l'expression de $\log(U)$ et comparer avec les tests de Student et Fisher.