Interrogation écrite n°2
Aucun document n'est autorisé
Durée: 1h20

Exercice [6 points]
On dispose d'un échantillon $(X_1,\ X_2,\ ...,\ X_N)$ de taille N d'une variable suivant une loi Normale de paramètres m et σ^2 avec $\sigma > 0$, de densité :

$$f(x\ ;m,\ \sigma^2) = \frac{1}{\sqrt{2\ \Pi\ \sigma^2}}\ \exp\left(-\frac{1}{2\ \sigma^2}(x-m)^2\right)$$

1- Ecrire la vraisemblance de l'échantillon ainsi que la log-vraisemblance.
2- On pose $s = \sigma^2$. Déterminer l'estimateur du maximum de vraisemblance des paramètres m et s.
3- Etudier le biais ainsi que l'efficacité de $\widehat{m}$ (l'estimateur de maximum de vraisemblance de m que vous avez déterminé en [2]).

4-***(Bonus) Etudier le biais de $\widehat{s}$ (l'estimateur de maximum de vraisemblance de s).
Si cet estimateur est biaisé, comment corriger le biais.

Problème [14 points]
On considère le modèle linéaire :

$$y_i = \beta_1 + \beta_2\ x_i + \varepsilon_i \qquad i = 1, ...N$$

avec les hypothèses : $E(\varepsilon_i) = 0$, $V(\varepsilon_i) = \sigma^2$, $Cov(\varepsilon_i, \varepsilon_j) = 0$ pour tout $i \neq j$
et les $\varepsilon_i \sim N(0, \sigma^2)$.
On note $\overline{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ la moyenne empirique de la variable endogène,
$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ celle de la variable explicative,
$S_{xy} = \sum_{i=1}^{N}(y_i - \overline{y})(x_i - \overline{x}) = \sum_{i=1}^{N} x_i y_i - N\overline{xy}$ la covariance empirique de x, y (à N près) et
$S_{xx} = \sum_{i=1}^{N}(x_i - \overline{x})^2 = \sum_{i=1}^{N} x_i^2 - N(\overline{x})^2$ la variance empirique de x (à N près).

1- Ecrire le modèle sous forme matricielle.
2- Ecrire les équations normales définissant l'estimateur MCO des coefficients β_1 et β_2. (Vous avez le choix d'utilisez soit la forme algébrique soit la forme matricielle).
3- En déduire $\widehat{\beta_2}$ en fonction de S_{xy} et de S_{xx}, et $\widehat{\beta_1}$ en fonction de $\overline{y}$,$\overline{x}$ et $\widehat{\beta_2}$.
4- Exprimer $\widehat{\beta_2}$ en fonction de ε_i.
5- Rappeler les propriétés des estimateurs des MCO. Vérifier ces propriétés pour $\widehat{\beta_2}$ puis pour $\widehat{\beta_1}$.
6-
6-1Quelle est la loi de $\widehat{\beta_2}$? Justifier.

6-2 Déterminer un intervalle bilatéral de confiance 95% pour β_2(sous l'hypothèse de σ^2 connue $\sigma^2 = 0,25$)

6-3 Calculer cet intervalle pour $N = 5$, $\sum_{i=1}^{5} y_i = 75$, $\sum_{i=1}^{5} x_i = 55$, $\sum_{i=1}^{5} x_i y_i = 880$, $\sum_{i=1}^{5} x_i^2 = 633$

7- On définit $z_i = y_i - x_i$ et on postule un nouveau modèle :

$$z_i = a + b\ x_i + v_i \qquad i = 1, ...N$$

7-1 Exprimer a, b et v_i en fonction de β_1 ,β_2 et ε_i.

7-2 Donner l'expression des estimateurs MCO de a et b.

7-3 Exprimer la covariance empirique entre z et x (à N près) (S_{xz}) en fonction de celle entre y et x (S_{xy}).

7-4 Montrer que $\widehat{a} = \widehat{\beta_1} - 1$ et que $\widehat{b} = \widehat{\beta_2}$.

Annexe

$F\,(2,33) = 0,975$
$F\,(2,05) = 0,98$
$F\,(1,96) = 0,975$
$F\,(1,65) = 0,95$
$F\,(1,28) = 0,90$

avec F la fonction de répartition de la loi normale standard $N(0,1)$