IFMA/MPE **Statistique de base**

Examen 2e session

Mardi 3 janvier 9h-11h

Tous les tests statistiques seront effectués au seuil de 5%.

Exercice 1. *On considère la fonction de production de Cobb-Douglas:*

$$Q = aK^bL^c$$

où Q est la production de l'entreprise, K la quantité de capital utilisée, L la quantité de travail utilisée. Les paramètres a, b et c sont des constantes inconnues. On recueille les données Q_i, K_i, L_i de n entreprises et on aimerait évaluer les paramètres a, b, c. Pour cela, on étudie le modèle linéaire associé:

$$\ln(Q_i) = \alpha + \beta \ln(K_i) + \gamma \ln(L_i) + \varepsilon_i.$$

1. *Quelle est la relation entre les paramètres (a, b, c) et (α, β, γ)?*
2. *Rappeler les hypothèses du modèle de régression linéaire.*
3. *Ecrire le modèle sous la forme $\mathbf{Y} = \mathbf{M}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$.*
4. *Rappeler l'expression de l'estimateur des moindres carrés $\widehat{\boldsymbol{\theta}}$. Donner sa matrice de covariance.*
5. *On note σ^2 la variance des ε_i et $(\hat{\varepsilon}_i)_{1\le i\le n}$ les estimateurs des résidus. Donner un estimateur sans biais de σ^2.*
6. *Quel test doit-on effectuer pour savoir si la quantité de capital a une influence positive sur la production? Vous donnerez la statistique de test et la règle de décision.*
7. *On veut tester l'hypothèse selon laquelle les rendements d'échelle sont constants, ce qui signifie que $Q(rK, rL) = rQ(K, L)$ pour tout réel $r > 0$. Quelles sont les contraintes vérifiées par le modèle dans ce cas? On note $\tilde{\varepsilon}_i$ les estimateurs des résidus de la régression linéaire suivante:*

   $$\ln(Q_i) - \ln(L_i) = \alpha + \beta(\ln(K_i) - \ln(L_i)) + \varepsilon_i.$$

   Donner en fonction de $(\hat{\varepsilon}_i)_{1\le i\le n}$ et $(\tilde{\varepsilon}_i)_{1\le i\le n}$ un test pour tester l'hypothèse des rendements d'échelle constants.

Exercice 2. *On considère* $(X_1, \ldots, X_n)$ *un* n*-échantillon de la loi d'une variable aléatoire* X *de densité*

$$f_\theta(x) = \theta^2 \frac{\ln(x)}{x^{\theta+1}} \mathbf{1}_{\{x>1\}}$$

où θ *est un paramètre réel inconnu strictement positif. Soit* $(x_1, \ldots, x_n)$ *une observation de cet échantillon. On rappelle que la densité de la loi du* χ^2 *à 4 degrés de libertés est* $\frac{1}{4} y e^{-\frac{1}{2}y} \mathbf{1}_{[0,+\infty[}(y)$.

On souhaite tester l'hypothèse nulle $(H_0) : \theta \leq 1$ *contre* $(H_1) : \theta > 1$.

1. *Montrer que la variable* $Y = 2\theta \ln(X)$ *suit une loi du* χ^2 *à 4 degrés de liberté. En déduire (sans calcul) que* $2\theta \sum_{i=1}^n \ln(X_i)$ *suit une loi du* χ^2 *à* $4n$ *degrés de liberté.*

2. *Déterminer l'estimateur du maximum de vraisemblance de* θ.

3. *Est-il fortement consistant?*

4. *Proposer un test de* (H_0) *contre* (H_1). *(On utilisera les quantiles de la loi* $\chi^2(4n)$*)*

Exercice 3. *Une marque de savon antibactérien prétend éliminer plus de bactéries qu'un savon ordinaire. Une association de consommation souhaite tester cette affirmation. On place une solution contenant le savon dans une boîte de Petri contenant la bactérie E.coli, puis on relève le nombre de bactéries après 24h. Des études précédentes ont montré que ce nombre est de 33 pour un savon ordinaire. Le savon antibactérien a été testé sur 35 boîtes: le nombre de bactéries après 24h est en moyenne de 31,2 avec un écart-type de 8,4. (On ne demande pas de faire d'applications numériques)*

1. *Enoncer le test d'hypothèses que l'on veut effectuer. On prendra soin de bien choisir les hypothèses nulle* (H_0) *et alternative* (H_1).

2. *Préciser la statistique de test utilisée et sa loi.*

3. *Quelles sont les hypothèses faites sur le modèle? Vous paraissent-elles crédibles?*

4. *Donner la règle de décision du test. (On supposera que l'application numérique conclut que les deux savons sont similaires)*

5. *La marque de savon affirme que d'après ses propres expériences, le nombre de bactéries après 24h est en moyenne de 29. Quel test effectuer pour savoir si la marque de savon dit bien la vérité? (On supposera que l'application numérique conclut qu'on ne peut pas rejeter que la marque de savon dit la vérité)*

6. *Donner l'expression d'un intervalle de confiance sur la vraie moyenne* μ *du nombre de bactéries après 24h.*

7. *On aimerait faire un nombre* N *de mesures afin de pouvoir départager les deux hypothèses* ($\mu = 33$ *et* $\mu = 29$)*. Proposer une méthode pour trouver une valeur de* N *adéquate.*