

## Instructions

- Veuillez répondre aux deux questions.
- Ecrivez lisiblement.
- Commencez chaque sous question sur une nouvelle page.
- Faites des phrases complètes.

### Q1. Vrai-Faux-Incertain [60%]

Dites si chacune des affirmations suivantes est vraie, fausse ou incertaine. Veuillez justifier votre réponse en moins d'une page par question.

- 1.1 Après avoir estimé un modèle bivarié à l'aide d'un échantillon de 10 observations, il convient d'enlever la 10e observation si le résidu associé à cette observation est le plus grand car cela veut dire que cette observation entraîne trop d'erreur. [10%]
- 1.2 Alanis est confrontée au modèle bivarié suivant:

$$\text{Revenus}_i = \beta_0 + \beta_1 \text{Dépenses policières} + \epsilon_i \quad (1)$$

Elle ne dispose pas des données brutes mais sait que le coefficient de corrélation entre la variable expliquée et les dépenses policières est égal à zéro. Elle conclut que la valeur de la statistique de student afin de tester si  $\beta_1 = 0$  est aussi égale à 0. [10%]

- 1.3 Supposez que le vrai modèle multivarié qui a généré la variable expliquée  $y_i$  est le suivant :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i \quad (2)$$

pour  $i = 1, \dots, N$  observations, et les  $x_j$ ,  $j = 1, \dots, 4$  sont les variables explicatives. Au lieu d'estimer le modèle (2), Yusuf estime par moindres carrés ordinaires le modèle

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \gamma z_i + \epsilon_i \quad (3)$$

où  $z_i = 0.01x_{4i}$ . Yusuf dit que la seule différence entre les deux modèles est que  $\hat{\gamma} = 0.01\hat{\beta}_4$ . [10%]

1.4 Les salaires sont générés par le vrai modèle suivant:

$$\text{logarithme des salaires}_i = \beta_0 + \beta_1 \text{ éducation}_i + \beta_2 \text{ intelligence}_i + \epsilon_i \quad (4)$$

Marthe sait que les variables explicatives éducation et intelligence sont corrélées et que  $\beta_1$  est positif. Elle ne dispose d'aucune donnée sur l'intelligence. Elle estime par moindres carrés ordinaires le modèle suivant :

$$\text{logarithme des salaires}_i = \beta_0 + \beta_1 \text{ éducation}_i + \epsilon_i \quad (5)$$

Marthe dit que l'omission de l'intelligence comme variable explicative n'aura aucun effet sur l'intervalle de confiance de  $\beta_1$ . [10%]

1.5 John analyse les déterminants de la demande de soins de santé en estimant le modèle économétrique suivant:

$$\text{Consommation de médicaments}_i = \beta_0 + \beta_1 \text{ Revenus}_i + \epsilon_i \quad (6)$$

pour  $i = 1, \dots, N$ . Denis dit à son collègue John que même si le modèle de consommation de médicaments est bien spécifié, l'estimateur des moindres carrés ordinaires sera biaisé car son modèle viole une des 4 hypothèses de base des MCO. [10%]

1.6 Soit les deux modèles suivants:

$$\begin{aligned} \text{Consommation}_i &= \beta_0 + \beta_1 \text{ Revenus}_i + \text{aleas}_i \\ \text{Épargne}_i &= \gamma_0 + \gamma_1 \text{ Revenus}_i + \text{aleas}_i \end{aligned}$$

Supposez que les revenus ne servent qu'à consommer et épargner. Dans ce cas:  $\beta_0 + \gamma_0 = 0$  et  $\beta_1 + \gamma_1 = 1$ . [10%]

## Q2. Problème [40%]

Répondez aux questions suivantes en montrant toutes les étapes de calcul. Soit le modèle multivarié :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i \quad (7)$$

pour  $i = 1, \dots, 100$ . Vous disposez des informations suivantes:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 0.0744 & -0.0711 & -0.0651 \\ -0.0711 & 0.1384 & 0.0155 \\ -0.0651 & 0.0155 & 0.1188 \end{bmatrix}, \quad \mathbf{X}'\mathbf{Y} = \begin{bmatrix} 566.4437 \\ 233.7753 \\ 294.3445 \end{bmatrix},$$

$$\text{SCR} = 88.4270, \quad \text{SCT} = 3416.8263.$$

où  $\mathbf{X}$  est la matrice de variables explicatives incluant la constante,  $\mathbf{Y}$  est le vecteur d'observations de la variable expliquée, SCR est la somme des carrés des résidus, et SCT est la somme des carrés totaux du modèle (7) estimé par moindres carrés ordinaires. Veuillez répondre aux questions suivantes. NB. vous n'avez besoin d'aucune autre information additionnelle pour répondre aux questions.

- 2.1 Estimez les paramètres du modèle (7) par moindres carrés ordinaires. [5%]
- 2.2 Supposez que les aléas ne sont pas auto-corrélés et ont la même variance (homoscédasticité). Estimez la variance de chaque aléa  $\epsilon_i$ . [5%]
- 2.3 Calculez la matrice de variance-covariance des paramètres estimés du modèle (7). [5%]
- 2.4 Calculez l'intervalle de confiance pour  $\beta_1$  au seuil de 5% en utilisant la valeur critique 1.96. Interprétez ce résultat. [3%]
- 2.5 Testez si  $\beta_2 = 1.75$ . Utilisez comme valeur critique 1.96. [3%]
- 2.6 Estimez les valeurs des paramètres si toutes les observations du modèle (7) sont divisées par 2.615. [5%]
- 2.7 Hector pense que les paramètres du modèle (7) diffèrent en fonction du sexe de l'individu. Modifiez la spécification (7) afin d'en tenir compte. [7%]
- 2.8 Calculez la statistique de test pour :

$$\begin{aligned} H_0 : & \quad \beta_0 = \beta_1 = \beta_2 = 0 \\ H_1 : & \quad \beta_0 \neq 0, \quad \text{ou} \quad \beta_1 \neq 0, \quad \text{ou} \quad \beta_2 \neq 0. \end{aligned}$$

Que concluez vous. (NB. La valeur trouvée devrait être tel que vous n'avez pas besoin de valeur critique pour conclure) [7%]