

<p><b>Épreuve : Méthodes quantitatives</b></p> <p><b>Durée : 02 Heures</b></p> <p><b>Date : 16/05/2023</b></p> <p><b>Nbre de pages : 02 pages</b></p>	<p><b>Université de Sousse</b></p>  <p><b>Institut des Hautes Études Commerciales de Sousse</b></p>	<p><b>Niveau : M1 BAGP</b></p> <p><b>Enseignant : Hamrita Mohamed Essaied</b></p> <p><b>Session principale</b></p>
---	--	--

*La présentation, la lisibilité, l'orthographe, la qualité de la rédaction et la clarté entreront pour une part importante dans l'appréciation des copies. Téléphone portable et documents interdits.*

### Exercice 1 (9 points)

On considère la relation linéaire entre le taux d'endettement public en pourcentage du PIB (noté  $Y$ ) et le taux d'investissement public en pourcentage du PIB (noté  $X$ ) sur une période de 40 années :

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t, \quad t = 1, 2, \dots, 40. \quad (1)$$

avec  $\varepsilon_t$  un terme d'erreur vérifiant les hypothèses de la régression classique.

- 1) Rappeler les hypothèses de la régression classique.
- 2) Estimer par la méthode des moindres carrés ordinaires les paramètres  $\beta_0$  et  $\beta_1$ , notés respectivement  $\hat{\beta}_0$  et  $\hat{\beta}_1$ . Commenter.

$$\sum_{t=1}^{40} X_t = 320, \quad \sum_{t=1}^{40} Y_t = -12, \quad \sum_{t=1}^{40} (X_t - \bar{X})^2 = 160, \quad \sum_{t=1}^{40} (Y_t - \bar{Y})^2 = 160, \quad \sum_{t=1}^{40} (X_t - \bar{X})(Y_t - \bar{Y}) = 80$$

- 3) Montrer que les estimateurs de  $\beta_0$  et  $\beta_1$  sont des estimateurs BLUE.
- 4) Écrire l'équation d'analyse de la variance et calculer ses composantes. En déduire le coefficient de détermination linéaire  $R^2$ . Commenter.
- 5) Tester la significativité individuelle des  $\beta_i$  au seuil  $\alpha = 5\%$ . (On donne  $Z_{0.975} = 1.96$ )
- 6) On suppose que  $\beta_0 = 0$ . Re-estimer le modèle sous cette hypothèse. L'estimateur de  $\beta_1$  est-il sans biais? Convergent?

### Exercice 2 (11 points)

Un assureur souhaite se baser sur les caractéristiques de ses 50 clients pour fixer le montant des indemnités en dinars ( $Y$ ) qui doit être renouvelé. Les caractéristiques dont il dispose pour chacun de ses clients sont : l'âge ( $X_1$ ), le temps passé depuis la dernière intervention de l'assurance ( $X_2$ ) et le genre de l'assuré ( $F$ ). Cette base de données est stockée dans un fichier appelé data enregistré dans le répertoire "D :/data.txt". L'estimation du modèle

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \quad (2)$$

a donnée les résultats suivants :

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.3463944	1.37678495	0.2515966	8.024498e-01
x1	35.4480127	0.02810995	1261.0484231	4.199451e-108
x2	-5.7124118	0.15496071	-36.8636138	2.442532e-36

- 1) Donner les lignes du code R permettant d'importer la base de données et d'effectuer l'estimation du modèle (2).
- 2) Interpréter les paramètres estimés et étudier la significativité individuelle de chacun des paramètres.
- 3) Écrire le code R permettant d'afficher le coefficient de détermination  $R^2$ . Interpréter

```
[1] 0.9999705
```

- 4) Écrire le code R qui donne le résultat suivant. Interpréter.

```

      Df  Sum Sq Mean Sq F value Pr(>F)
x1      1 2382262 2382262 1589409 <2e-16 ***
x2      1   2037    2037    1359 <2e-16 ***
Residuals 47      70      1
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

- 5) Tester, au seuil  $\alpha = 5\%$ ,  $H_0 : \beta_1 = 35$ .

L'assureur veut tester si le genre de l'assuré a un effet sur la prime d'indemnité. Il a introduit la variable  $F$  ( $F_i = 1$  si l'assuré est une femme, 0 sinon)

$$Y_i = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_3 F_i + \varepsilon_i \quad (3)$$

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.33628291	1.41421079	0.23778839	8.131012e-01
x1	35.44785347	0.02868835	1235.61862245	1.215212e-105
x2	-5.71373148	0.16003905	-35.70210925	3.475016e-35
F	0.01471364	0.36615204	0.04018452	9.681199e-01

- 6) Interpréter le paramètre  $\alpha_3$ .
- 7) Décrire le test à effectuer pour tester s'il y a une différence significative de la tarification entre les femmes et les hommes. (On précisera les hypothèses, la statistique à utiliser, sa loi sous l'hypothèse nulle, la règle de décision et les codes R nécessaires).

**Bon travail**