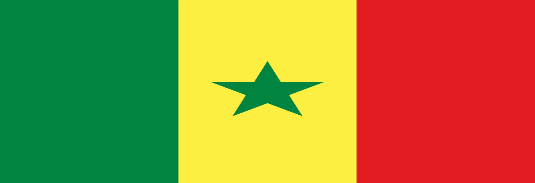
**REPUBLIQUE DU SENEGAL**



**Un peuple-un but-une foi**

**MINISTERE DE L’ENSEIGNEMENT SUPERIEUR DE LA RECHERCHE ET DE L’INNOVATION**



**UNIVERSITE DE THIES**

**UFR DES SCIENCES ET TECHNOLOGIES&**

**UFR DES SCIENCES ECONOMIQUES ET SOCIALES**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

Présenté par : **KHOUDIA MBODJI**

Niveau : MASTER 1 EN SCIENCES DES DONNEES ET APPLICATIONS ;OPTION :INTELLIGENCE ARTIFICIELLE

**APPRENTISSAGE STATISTIQUE**

**Travail individuel.**

**À rendre avant 31 janvier 2021 à 23h59**

**EXERCICE 1 :** On considère le modèle de régression yi = β0 + β1xi,1 + β2xi,2 + εi 1 ≤ i ≤ n, que l’on écrit sous la forme Y = Xβ+ε. Les xi, j sont des variables exogènes du modèle, les εi sont des variables aléatoires indépendantes, de loi normale centrée admettant la même variance σ¨2. On a observé :

X’X = (30,20,0 ; 20,20,0 ; 0,0,10), X’Y = (15,20,10), Y ‘ Y = 59.5

1. Déterminer n, la moyenne des xi,2, le coefficient de corrélation des xi,1 et des xi,2.

* n=n(X’X)=30
* xi,2=1/n(xi,2)= (X′X)1,2/30 = 2/3
* r1,2=(X’X)1,2/√∑i=1,30(xi,1-xi,1‾)²√∑i=1,30(xi,2)²=0

1. Estimer β0, β1, β2, σ2 par la méthode des moindres carrés ordinaires.

La méthode des MCO donne pour β = [β0, β1, β2] ′ l’estimateur suivant :

Βˆ = (X′X) ¨ −1\*X′Y = ??

X‘X= (30 ,20,0 ;20,20,0 ;0,0,10)

-Calculons le déterminant :

Det(X’X)=30\*[20,0 ;0,10] -20\*[20,0 ;0,10] +0=6000-4000=2000

**Det(X’X)=2000**

-Calculons l’inverse de la matrice

(X’X)T= (30,20,0 ;20,20,0 ;0,0,10)

[20,0 ;0,10] =200

[20 ,0 ;0,10] =-200

[20,20 ;0,0] =0

[20,0 ;0,10] =-200

[30 ,0 ;0,10] =300

[30,20 ;0,0] =0

[20,0 ;20,0] =0

[30,0 ;20,0] =0

[30 ,20 ;20,20] =200

Adj(X’X)= [200,200,0 ; 200,300,0 ;200,0,200]

(X’X) −1=1/det(X‘X) \*Adj(X’X)= 1/2000\*[200,200,0 ; 200,300,0 ;0,0,200]

(X’X)= [0.1,-0.1,0 ; - 0.1,0.15,0 ;0.1,0,0.1]

Βˆ = (X′X) ¨ −1\*X′Y) = [0.1,-0.1,0 ; - 0.1,0.15,0 ;0,0,0.1] \* [15,20,10] = [ −0.5,1.5,1]

**Βˆ = [ −0.5,1.5,1]**

Un estimateur non biaisé σˆ 2 de σ 2 s’écrit :

σˆ 2 = ⎟⎟Y − Xβˆ⎟⎟²/( n – 3) = (⎟⎟Y⎟⎟² − ⎟⎟Xβˆ⎟⎟²) / 27 =( Y ′Y − Y ′X(X′X) −1X′Y) /27 = 1

1. Calculer pour β1 un intervalle de confiance à 95% et tester β3 = 0.8 à niveau 10%

–Calcule de β1 un intervalle de confiance

(βˆ 1 – β1)/σˆ2 =( βˆ 1 – β1)/ σˆ (√(X′X) −1 2,2 )∼ Tn−3 = T27

On en déduit :

I(β1) = [βˆ 1 ± t27(0.95)ˆσ √(X′X)−1 1,1 ]

Ce qui signifie :

I(β1) ≈ [1.5 ± 2.05√ 0.15]

**I(β1) ≈ [0.71; 2.29**].

-Pour tester l’hypothèse:

β3 = 0.8 contre H1 : β3 6= 0.8 au niveau 10%, on calcule de même un intervalle de confiance à 90% de β3 :

I(β3) = [βˆ 3 ± t27(0.95)ˆσ (√ (X′X) −1 3,3)]

I(β3) ≈ [1 ± 1.70√ 0.1]

Ce qui donne :

I(β3) ≈ [0.46; 1.54]

Donc on accepte au niveau 10% l’hypothèse selon laquelle β3 = 0.8.

1. Tester β0 + β1 = 3 contre β0 + β1 6= 3 au niveau 5%

(βˆ ² + βˆ 3) − (β2 + β3) /σˆβˆ ²+βˆ 3 ∼ T27

Or on : σˆβˆ2 +βˆ3 = √(σˆ² 2 + 2Cov(βˆ 2, βˆ 3) + ˆσ²3 )= ˆσ √( (X′X) −1 2,2 + 2(X′X) −1 2,3 + (X′X) −1 3,3)

AVEC σˆβˆ2+βˆ3 = 0.5

Donc on a : I (β2 + β3) = [2.5 ± 0.5t27(0.95)]

Ce qui donne : **I (β2 + β3) = [1.47; 3.53**]

Par conséquent, au niveau 5%, on accepte H0 : β2 + β3 = 3 contre H1 : β2 + β3 ! = 3

1. Calculer y et déduire le coefficient de détermination ajusté R²

**y¯ = 15/30 = 0.5**

**-** R² :

R²a =1- (n – 1)/( n – p) \*⎟Y − Yˆ ⎟⎟²/⎟⎟Y − y¯1⎟⎟² = 1 − (n − 1) \*σˆ 2 /⎟⎟Y − y¯1⎟⎟²

Ce qui donne :

R²a = 1 – (29 / (Y ′Y − 30¯y²)) ≈ 0.44

6.Construire un intervalle de prévision à 95% de yn+1 si xn+1,1 = 3 et xn+1,2 = 0.5.

Posons : x ′ n+1 = [1, 3, 0.5]

La valeur prédite pour yn+1 est : yˆn+1 = x ′ n+1\*βˆ = 9/2

L’intervalle de prévision à 95% pour yn+1 est :

IC(yn+1) = [yˆn+1 ± t27(0.975)ˆσ (√(1 + x ′ n+1(X′X)−1\*xn+1))

Ce qui donne :

**IC(yn+1) ≈ [1.69; 7.31]**