**Examen de statistique  
1h30**

**Nom : ………………………………………………………**

**Filière : ……………………………………………………**

NB : - Cette examen comporte deux parties : la partie théorique et la partie pratique ; les questions de la partie pratique portent le signe (\*).  
 - Ecrire dans la dernière page les détails des calculs des questions qui porte le signe (\*\*).

**Exercice 1.**

On veut analyser s’il existe une relation de corrélation linéaire entre la densité de population dans les régions métropolitaines et le taux de criminalité correspondant dans ces régions. Le taux de criminalité (Y) est indiqué en nombre de crimes par 10 000 habitants et la densité de population (X) est mesurée en milliers d’habitants par km2.

Les données statistiques sont représentées dans tableau suivant



1. Préciser la variable statistique explicative

…la variable statistique explicative est le nombre de crimes par 10000 habitants (noté Y)…..

1. Tracer le nuage de point de cette distribution statistique bi-variée.



1. (\*) Donner le**s** commande**s** en R qui permet de réaliser la question 2 (depuis la création des vecteurs X et Y

> X= c(7.7 , 5.8 , 11.5 , 2.1 , 3.7 , 3.6 , 7.5 , 4.2 , 3.8 , 10.3 , 8.6 , 7.2)

> Y=c(12 , 9 , 15 , 4 , 4 , 2 , 10 , 3 , 5 , 11 , 10 , 11)

>plot(X,Y)……………………………………………………………………………………

1. Donner le coefficient de corrélation de régression linéaire de Y en X
   1. Formule : ………………………………..…………………………..…………………..
   2. (\*\*) Résultat de l’Application numérique : 0,9172042 (j’ai accepté tous les résultats proches de 0.91)
   3. (\*) Commande en R : … cor(X,Y)……………………………………………………..………….
2. Donner les coefficients de la droite de régression linéaire de Y en X (Y= a + bX).
   1. Formule : a= …ou … , b=……ou…
   2. (\*\*) Application numérique : a=…-0,2956………………………… , b=… 1,3098…….……
   3. (\*) Commande en R : lm(Y~X)……………………………………………………….……………….
3. Estimer le taux de criminalité pour une densité de population de 75 000 habitants par km2.

Utiliser le modèle linéaire de la question précédente : Y= -0,2956 + 1,3098 X

pour 75 000 habitants X= 75

on remplace X dans le modèle et on trouve : Y= -0,2956 + 1,3098 \* 75 🡪 Y= 97.9394

1. ~~A l’aide des calculs préliminaires, calculer la variation totale du taux de criminalité.~~
2. (\*) Implémenter en logiciel R une fonction TCrimByDPop qui permet de prédire le taux de criminalité pour une densité de population passée en paramètre

TCrimByDPop = function (x) {return (-0,2956 + 1,3098 x) }

Ou bien

TCrimByDPop = function (x) {

linModel= lm(Y~X)

return (coef(linModel)[1] + coef(linModel)[2] \* x) }

Ou bien

TCrimByDPop = function (x) {

linModel= lm(Y~X)

y=coef(linModel)[1] + coef(linModel)[2] \* x

return (y) }

**Exercice 2.**

L’entreprise CITRON fabrique un matériau en matière plastique qui est utilisé dans la fabrication de jouets. Le département de contrôle de qualité de l’entreprise a effectué une étude qui a pour but d’établir dans quelle mesure la résistance à la rupture (en kg/cm2) de cette matière plastique pouvait être affectée par l’épaisseur du matériau ainsi que la densité de ce matériau. Douze essais ont été effectués et les résultats sont présentés dans le tableau de la figure 1.

1. Citer les différentes variables statistiques étudiées ainsi que leur type, en précisant celles explicatives et celles à expliquer

… résistance à la rupture : variable à expliquer, variable quantitative continue

… l’épaisseur du matériau : variable explicative, variable quantitative discrète

… la densité de ce matériau : variable explicative, variable quantitative continue

…………………………………………………………………………………………………

1. (\*) Donner l’expression en logiciel R qui permet d’avoir le model linière modelLS : Y = A\*X1 + B\*X2 +C

modelLS = lm(Y ~ X1 + X2)

1. (\*) Même question pour le model  « ModelP1 »: Y = A\*X12 + B\*X22 + C\*X1 + D\*X2 +E\*X1 \*X2+C

ModelP1  = lm(Y ~ X1 + X2 + I(X1^2) + I(X2^2) + I(X1\*X2))

1. Donner la valeur estimer de Y pour X1 = 5 et X2 = 6 selon le model modelLS:

Y = 4.082\*5 + 14.106\*6 -37.709 = 67.337

1. Donner la valeur estimer de Y pour X1 = 5 et X2 = 6 selon le model ModelP1:

Y = -0.03351\* 5^2 + 7.94638\*6^2 + 1.23201\*5 - 42.03203\*6 + 0.86939\*5 \*6 + 59.78088

Y=125.0624

Expliquer en quelques lignes la signification des étoiles qu’on trouve dans le summary de modelLS

Les étoiles qu’on trouve dans le summary de modelLS indiquent les paramètres ayant un effet significatif, plus qu’on trouve des étoiles plus que le paramètre et significatif

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1