

Licence 1 mention Economie parcours économie-gestion
Licence 1 mention Economie parcours économie-droit
Epreuve : STATISTIQUE DESCRIPTIVE
A. RUIZ-GAZEN/S. CASANOVA/E. LECONTE/C. VIGNES/A. BOUDIAF/S. DINAR
Code : L1-S2-4

*Seules les calculatrices non programmables sont autorisées.
Toutes vos réponses doivent être justifiées.*

I- Questions de cours

On considère les valeurs x_1, \dots, x_n et y_1, \dots, y_n de deux variables quantitatives continues X et Y pour n observations. Répondre aux questions de cours suivantes.

1. Donner les formules de la moyenne, de la variance et de l'écart-type de X ainsi que les formules de la covariance et du coefficient de corrélation entre X et Y . Préciser les valeurs que peut prendre un coefficient de corrélation et son interprétation.
2. Enoncer la formule de König-Huyghens pour la variance et la démontrer.
3. On note $y = \hat{a}x + \hat{b}$ l'équation de la droite de régression par moindres carrés de Y sur X . Donner les formules pour calculer \hat{a} et \hat{b} .

On décide de découper la variable X en I classes et la variable Y en J classes.

4. Donner la formule de la distance du χ^2 associée au couple (X, Y) en précisant toutes les notations nécessaires.
5. Donner la définition du C de Cramer, ses valeurs possibles et son interprétation.

II- Exercice

On dispose du salaire annuel en milliers d'euros et de la catégorie d'emploi (employé, cadre stagiaire ou cadre) de 461 salariés d'une entreprise. Utiliser les commandes et résultats obtenus avec le logiciel R pour répondre aux questions ci-dessous.

```
> summary(salaire)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 15.75  24.00   28.50   33.98  35.70  135.00
> table(cat_emploi)
      Cadre Cadre_stagiaire      Employe
         81             27         353
> tapply(salaire, cat_emploi, mean)
      Cadre Cadre_stagiaire      Employe
    64.98             30.94         27.10
> tapply(salaire^2, cat_emploi, sum)
      Cadre Cadre_stagiaire      Employe
 367343.11         25961.06  271755.58
```

1. Préciser les individus, leur nombre, les deux variables étudiées et leur type.
2. Donner le nom et la valeur numérique de deux indicateurs de position et de deux indicateurs de dispersion pour la variable salaire.
3. Représenter graphiquement la distribution de la variable catégorie d'emploi et donner le nom du graphique.
4. Donner le nom d'un graphique qui permet de visualiser la liaison entre les deux variables. Pouvez-vous le tracer? Justifier votre réponse.
5. Rappeler la définition du rapport de corrélation en précisant toutes les notations. Le calculer et l'interpréter.