

Licence 1 MIASHS – Examen de Statistique

Coller ici votre étiquette

Répondre directement sur la présente copie aux endroits indiqués. Justifier et détailler les calculs. La note tiendra compte de la qualité de la rédaction. On arrondira les calculs à 2 décimales après la virgule (sauf pour l'exercice 4 pour le calcul des variances où 4 décimales sont demandées). Barème indicatif : 2,5 - 2,5- 7,5 - 7,5

Exercice 1 Soit $(x_i, y_i)_{i=1}^n$ une série statistique double correspondant aux n observations individuelles d'un couple (X, Y) de variables quantitatives.

1. Donner la formule de la covariance des variables X et Y .

$$\text{Cov}(X, Y) =$$

2. Rappeler et démontrer la formule de König-Huyghens pour la covariance.

Formule :

Démonstration :

Exercice 2 Pour N biens, soient $p_{i,0}$ le prix du bien i à $t = 0$, $p_{i,n}$ le prix du bien i à $t = n$, $q_{i,0}$ la quantité du bien i vendue à $t = 0$, $q_{i,n}$ la quantité du bien i vendue à $t = n$. On donne :

$$\sum_{i=1}^N p_{i,0} q_{i,0} = 400, \quad \sum_{i=1}^N p_{i,0} q_{i,n} = 410, \quad \sum_{i=1}^N p_{i,n} q_{i,n} = 422, \quad \sum_{i=1}^N p_{i,n} q_{i,0} = 416.$$

1. Calculer l'indice des prix de Laspeyre.

Formule :

Application numérique :

2. Calculer l'indice des prix de Paasche.

Formule :

Application numérique :

3. En déduire l'indice des prix de Fischer.

Formule :

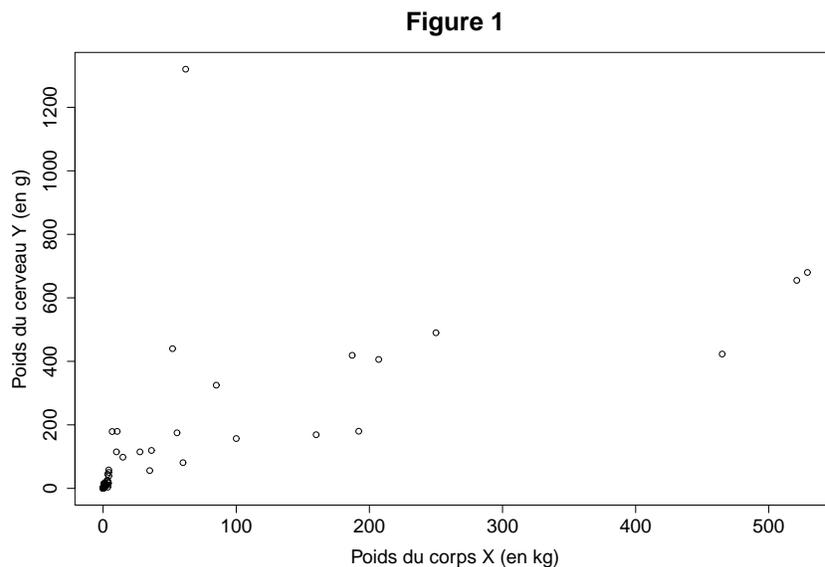
Application numérique :

4. Commenter brièvement pour chaque indice l'évolution des prix.

5. En quel indice avez-vous le plus confiance ? Expliquer pourquoi.

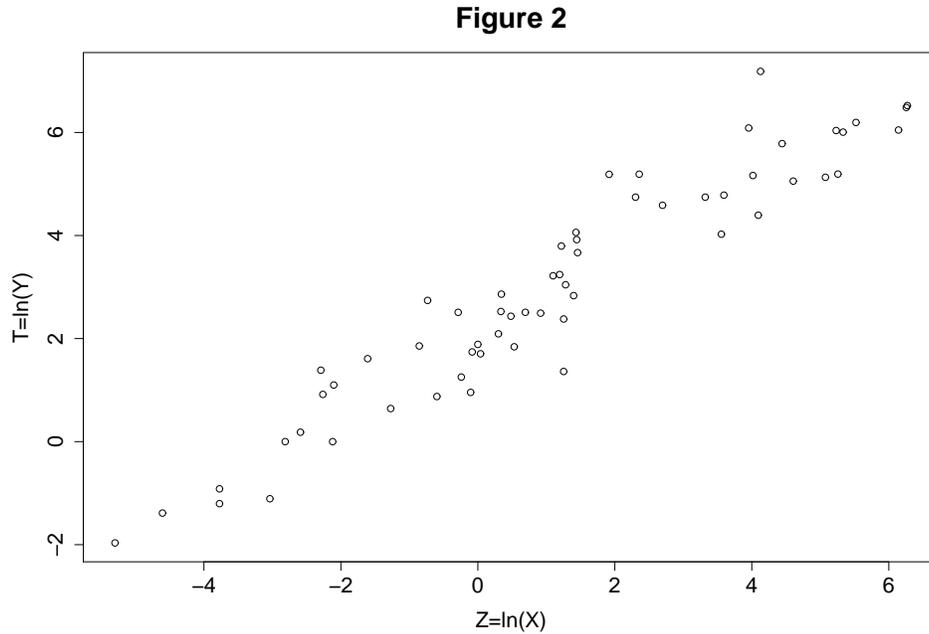
Exercice 3 On dispose d'un fichier de données contenant, pour 60 espèces de mammifères, le poids moyen du corps de l'espèce (en kg), variable notée X et le poids moyen du cerveau (en g), variable notée Y .

1. Le nuage de points associé à ces données est donné ci-dessous (figure 1).



Commenter l'allure de ce graphique.

2. On passe au logarithme et on note $Z=\ln(X)$ et $T=\ln(Y)$. Le nuage de points des observations $t_i = \ln(y_i)$ en fonction des observations $z_i = \ln(x_i)$ est donné ci-dessous (figure 2).



Commenter l'allure de ce graphique.

On donne :

$$\sum_{i=1}^n z_i = 66,28, \quad \sum_{i=1}^n t_i = 177,61, \quad \sum_{i=1}^n z_i^2 = 566,91, \quad \sum_{i=1}^n t_i^2 = 830,51, \quad \sum_{i=1}^n z_i t_i = 565,36.$$

3. Calculer les moyennes et variances des variables Z et T .
4. Calculer la covariance et le coefficient de corrélation linéaire des variables Z et T . Commenter sa valeur.

Covariance :

Coefficient de corrélation :

Commentaire :

On considère le modèle linéaire : $T = aZ + b$.

5. A quelle relation entre les variables X et Y de départ correspond ce modèle ?

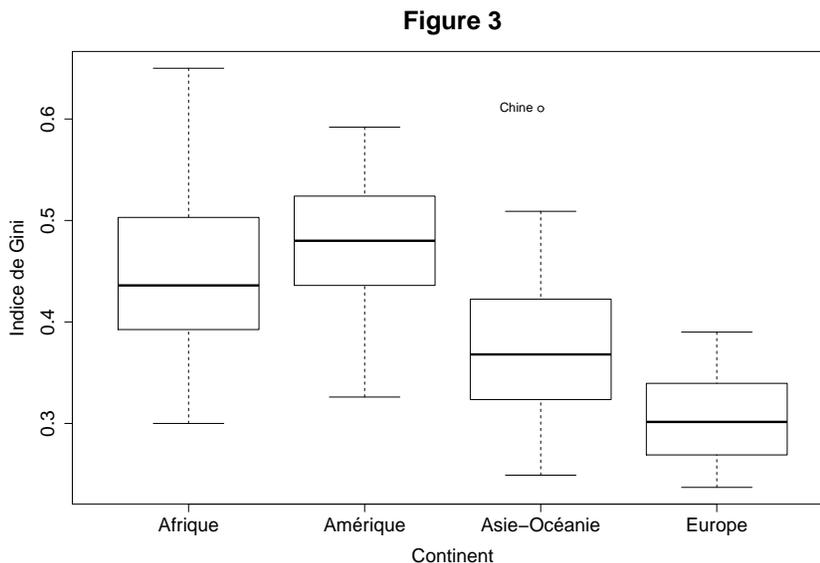
6. Calculer les estimateurs des moindres carrés \hat{a} et \hat{b} du modèle linéaire $T = aZ + b$ et écrire l'équation de la droite obtenue.

7. Tracer la droite de régression obtenue sur le nuage de points de la figure 2.
8. L'individu numéro 34 du fichier de données a un poids moyen du corps de 62 kg et un poids moyen du cerveau de 1320 g.
 - (a) Entourer cet individu sur les figures 1 et 2.
 - (b) Calculer la valeur du poids du cerveau que le modèle $T = aZ + b$ prédit pour cet individu.

 - (c) Calculer le résidu $\hat{\varepsilon}_{34} = t_{34} - \hat{t}_{34}$ de cet individu et le faire figurer sur la figure 2.

 - (d) Deviner de quel mammifère il s'agit (donner le nom de l'espèce) et commenter.

Exercice 4 On dispose, pour plusieurs pays du monde, de l'indice de Gini des revenus du pays ainsi que du continent où le pays est situé. On donne le graphique suivant (figure 3) ainsi que des instructions R et les sorties obtenues (l'instruction `round` permet d'arrondir la sortie au nombre de décimales désiré).



```
> round(var(gini), digits=3)
[1] 0.0096
> table(continent)
  Afrique  Amérique Asie-Océanie  Europe
      36      21         35         32
> round(tapply(gini, continent, mean), digits=2)
  Afrique  Amérique Asie-Océanie  Europe
    0.46    0.48         0.38    0.31
> round(tapply(Gini, continent, var), digits=4)
  Afrique  Amérique Asie-Océanie  Europe
 0.0087    0.0041         0.0058 0.0016
```

1. Quelle est la population considérée ?
2. Quelle est la taille de l'échantillon dont on dispose ?
3. Quelles sont les variables étudiées et quel est leur type ?
4. De quelles distributions le graphique de la figure 3 est-il la représentation ? Comparer ces distributions et commenter la valeur observée pour la Chine (vous avez de la place pour écrire à la page suivante).

5. Rappeler la formule qui relie la moyenne marginale aux moyennes conditionnelles et en déduire la moyenne marginale de la variable “gini”.

Formule :

Application numérique :

6. Calculer la variance marginale de la variable “gini” (arrondir à 4 décimales).

7. Calculer les variances conditionnelles de la variable “gini” sachant le continent (arrondir à 4 décimales)

8. Donner la formule de la variance intra-groupes et la calculer.

Formule :

Application numérique :

9. En déduire la valeur de la variance inter-groupes.

10. Calculer le quotient de corrélation η et qualifier l’intensité de la liaison entre les deux variables.

Formule :

Application numérique :

Commentaire :