



Université de Poitiers Département de Mathématiques

Statistique descriptive, 1er semestre, année univ. 2009-2010

Fiche 5

Statistiques descriptives bivariées

Exercice 1

Le relevé du revenu mensuel en centaines d'euros (caractère noté X) et du nombre de voitures (caractère noté Y) de chaque famille d'une petite population a permis d'établir le tableau de contingence suivant :

| |]7; 11] |]11; 15] |]15; 23] |]23; 39] |]39; 59] |
|---|---------|----------|----------|----------|----------|
| 0 | 5 | 4 | 4 | 0 | 0 |
| 1 | 9 | 12 | 15 | 3 | 5 |
| 2 | 1 | 8 | 10 | 12 | 8 |
| 3 | 0 | 1 | 0 | 1 | 2 |

- 1) Donner les indicateurs de position, dispersion des caractères X et Y .
- 2) Tracer l'histogramme de la loi de X .
- 3) Déterminer la loi de X conditionnée par la modalité 1 de Y notée X_1 .
- 4) Déterminer la loi de Y conditionnée par la modalité]15; 23] de X notée Y_3 .
- 5) Calculer la covariance empirique et le coefficient de corrélation linéaire $\rho_{X,Y}$. Les caractères X et Y sont-ils linéairement indépendants ?
- 6) Déterminer les droites de régression linéaire de Y par rapport à X et de X par rapport à Y .
- 7) Calculer le coefficient du χ^2 .

Exercice 2

Le tableau ci-dessous donne le nombre de contribuables en France soumis à l'impôt sur le revenu en 1980 (revenus de 1979) par tranches de revenus : $N(x)$ représente le nombre de revenus supérieurs au revenu x , x_i est la borne de gauche de chaque classe,

| Tranche de Revenu (kFF) | Nb de contribuables(*10 ³) | $N(x_i)$ | $y_i := \log x_i$ | $z_i := \log(N(x_i))$ |
|-------------------------|----------------------------------------|----------|-------------------|-----------------------|
| 15 – 20 | 1384 | 15000 | 2,71 | 9,62 |
| 20 – 25 | 1663 | 13616 | 3,00 | 9,52 |
| 25 – 30 | 1835 | 11953 | 3,22 | 9,39 |
| 30 – 35 | 1636 | ? | 3,40 | 9,22 |
| 35 – 40 | 1277 | 8482 | 3,56 | 9,05 |
| 40 – 50 | 2202 | 7205 | 3,69 | 8,88 |
| 50 – 60 | 1576 | 5003 | 3,91 | 8,52 |
| 60 – 70 | 1031 | 3427 | ? | 8,14 |
| 70 – 80 | 668 | 2396 | 4,25 | 7,78 |
| 80 – 100 | 734 | 1728 | 4,38 | 7,45 |
| 100 – 200 | 790 | 994 | 4,61 | ? |
| 200 – 400 | 163 | 204 | 5,30 | 5,32 |
| 400 – 600 | 41 | 41 | 5,99 | 3,71 |

- 1) Calculer le revenu médian.

- 2) Compléter le tableau donné et en déduire, par la méthode des moindres carrés, un ajustement linéaire de z en fonction de y .
- 3) En déduire une estimation des constantes réelles positives C et α dans la relation :

$$N(x) = \frac{C}{x^\alpha}.$$

- 4) Retrouver la valeur du revenu médian en utilisant la loi ajustée.

Exercice 3

Lors de l'analyse de la durée d'une période de chômage, une variable explicative potentielle est le niveau de formation. Le tableau suivant résume les données issues d'un panel socio-économique représentatif de 447 personnes.

| Chomage | ≤ 6 mois | 7 – 12 mois | ≥ 12 mois | Total |
|-------------------------|---------------|-------------|----------------|-------|
| Pas de formation | 86 | 19 | 18 | 123 |
| Formation de base | 170 | 43 | 20 | 233 |
| Formation spécialisée | 40 | 11 | 5 | 56 |
| Formation universitaire | 28 | 4 | 3 | 35 |
| Total | 324 | 77 | 46 | 447 |

- 1) Étudiez cette table de contingence : fréquences conjointes, profils conditionnels.
- 2) Calculez le coefficient du χ^2 (mesure de l'écart à l'indépendance).