

TD de Statistique L2 N°1

11 octobre 2007

Exercice 1

On s'intéresse à la distribution des salaires hebdomadaires des 70 employés d'une société. On a classé ces informations dans les tableaux suivants.

TAB. 1 –

| Salaires (€) | Fréquence |
|--------------|-----------|
| 250 - 259 | 8 |
| 260 - 269 | 10 |
| 270 - 279 | 16 |
| 280 - 289 | 15 |
| 290 - 299 | 10 |
| 300 - 309 | 5 |
| 310 - 319 | 3 |
| 320 - 329 | 0 |
| 330 - 339 | 1 |
| 340 - 349 | 0 |
| 350 - 359 | 1 |
| 360 - 369 | 0 |
| 370 - 379 | 1 |

TAB. 2 –

| Salaires (€) | Fréquence |
|--------------|-----------|
| 250 - 259 | 8 |
| 260 - 269 | 10 |
| 270 - 279 | 16 |
| 280 - 289 | 15 |
| 290 - 299 | 10 |
| 300 - 309 | 5 |
| 310 - 319 | 3 |
| 320 et plus | 3 |

TAB. 3 –

| Salaires (€) | Fréquence |
|--------------|-----------|
| 250 - 269 | 18 |
| 270 - 289 | 31 |
| 290 - 309 | 15 |
| 310 - 329 | 3 |
| 330 - 349 | 1 |
| 350 - 369 | 1 |
| 370 - 389 | 1 |

TAB. 4 –

| Salaires (€) | Fréquence |
|--------------|-----------|
| 250 - 259 | 8 |
| 260 - 269 | 10 |
| 270 - 279 | 16 |
| 280 - 289 | 15 |
| 290 - 299 | 10 |
| 300 - 319 | 8 |
| 320 - 379 | 3 |

1) Tracer les histogrammes et les polygones des fréquences pour chaque tableau. On rappelle qu'un polygone de fréquence est une ligne reliant les fréquences de classe en fonction des milieux de classes.

Commenter les avantages et inconvénients de chaque jeu de données.

2) Pour des intervalles de classes constants, démontrer que l'aire totale limitée par les rectangles d'un histogramme est égale à celle qui est limitée par le polygone de fréquence. Vérifier que ce n'est pas le cas pour des intervalles de classes inégaux.

3) Construire sur la même figure la distribution des fréquences cumulées et des pourcentages cumulés.

Exercice 2

1) Construire une échelle logarithmique dans l'intervalle $[0.1;100]$.

$$\log(2) = 0.30 \qquad \log(3) = 0.47$$

3) Donner les limites de classes d'une échelle logarithmique, les paramètres étant le nombre de classes n_c et les extrema de la distribution V_{min} et V_{max} .

3) A partir des données présentées dans Tabs. 5 et 6, exprimer une relation entre la concentration et la température, et entre la période du fil et sa longueur.

TAB. 5 – Concentration de X en fonction de la température.

| Concentration | Temp. (°C) |
|---------------|------------|
| 0.0280 | 28.5 |
| 0.0470 | 28.0 |
| 0.085 | 27.5 |
| 0.150 | 27.0 |
| 0.250 | 26.5 |
| 0.440 | 26.0 |
| 0.800 | 25.5 |

TAB. 6 – Période d'un pendule en fonction de sa longueur.

| Period (s) | Long. (m) |
|------------|-----------|
| 0.63 | 0.08 |
| 0.80 | 0.130 |
| 1.28 | 0.345 |
| 1.86 | 0.830 |
| 2.55 | 1.65 |
| 4.00 | 4.25 |
| 5.50 | 8.90 |

Exercice 3

1) Comment évaluer la médiane pour un échantillon de (a) 85 données et (b) 150 données?

2) A partir du Tab. 7, trouver la médiane et le mode du poids des porcs de cet élevage.

TAB. 7 – Poids des individus dans une porcherie.

| Poids (kg) | Fréquence |
|------------|-----------|
| 118-126 | 3 |
| 127-135 | 5 |
| 136-144 | 9 |
| 145-153 | 12 |
| 154-162 | 5 |
| 163-171 | 4 |
| 172-180 | 2 |

TAB. 8 – Fréquence des salaires chez Dupont & Fils.

| Salaires (€) | Fréquence |
|--------------|-----------|
| 250 - 259 | 8 |
| 260 - 269 | 10 |
| 270 - 279 | 16 |
| 280 - 289 | 14 |
| 290 - 299 | 10 |
| 300 - 309 | 5 |
| 310 - 319 | 2 |

Exercice 4

1) Trouver les quartiles $Q_{1,2,3}$ et les déciles $D_{1,2,\dots,9}$ pour les revenus des 65 employés de Dupont & Fils (Tab. 8).

2) Montrer comment ces résultats peuvent être obtenus graphiquement à partir de la distribution cumulative. Déterminer le 35^e et le 60^e centile de la distribution graphiquement et numériquement.

3) Construire et discuter une courbe indiquant le nombre d'employés de Dupont & Fils gagnant plus qu'un salaire donné.

TAB. 9 – Durée de la connection des internautes en minutes.

| | | | | | | | | | |
|---|----|----|----|----|----|----|----|----|----|
| 3 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 6 |
| 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 8 | 8 |
| 9 | 10 | 10 | 10 | 10 | 10 | 10 | 12 | 55 | 60 |

Exercice 5

A partir des données du Tab. 9, tracer un histogramme. Discuter la valeur moyenne des durées de connection. Montrer comment évolue cette moyenne lorsque l'on retire certains internautes.