

1. Questions de cours

a. Comment se traduit en terme de fréquence le fait que deux évènements E et F sont pratiquement indépendants ?

b. X est un caractère quantitatif ; $a < b$ sont des nombres réels et on sait $f_{X \in [a,b]} > 0.1$. Quelles sont les affirmations vraies parmi les affirmations suivantes :

- (1) “La fréquence de X est entre a et b avec une probabilité supérieure à 0.1” ;
- (2) “Les valeurs prises par X sont supérieures à 0.1” ;
- (3) “ X prend des valeurs entre a et b ” ;
- (4) “au moins 10% des valeurs prises par X sont inférieures à b ” ;
- (6) “ $a > 0.1$ ”.

c. X, Y sont des caractères quantitatifs ; $a < b$ et $c < d$ sont des nombres réels. Interpréter en terme d'évènements les énoncés suivants :

$$f_{X \in [a,b]} = 1, \quad \frac{f_{X \in [a,b] | Y \in [c,d]}}{f_{X \in [a,b]}} \gg 1, \quad f_{X \in [a,b] | Y \in [c,d]} \simeq 1.$$

2. Pour chacun des extraits de texte suivants indiquer la population, les évènements en jeu, la relation entre les probabilités ou fréquences liées à ces évènements :

- “Un an après la sortie de l’université, la probabilité d’être sans emploi est inférieure à 30%. Elle est deux fois plus grande après des études longues.”
- “Le risque de développer la maladie est 20% moins grand chez un patient de plus de 20 ans que chez un patient de moins de 20 ans.”
- “Fumer double le risque de maladie cardio-vasculaire.”

3. Voici ci-dessous un tableau extrait d’une publication de l’Insee

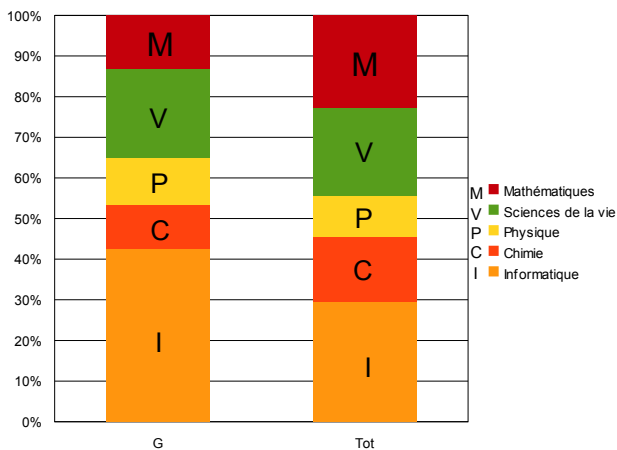
Tableau 2 : Taux d'activité et recours au temps partiel des hommes et des femmes de 25 à 49 ans selon le nombre et l'âge des enfants

| | Taux d'activité | | | Temps partiel parmi les personnes ayant un emploi | | |
|-------------------|-----------------|-------------|-------------|---|------------|-------------|
| | Femmes | Hommes | Ensemble | Femmes | Hommes | Ensemble |
| Ensemble | 83,1 | 96,2 | 89,4 | 29,2 | 3,7 | 15,9 |
| Aucun enfant | 91,0 | 94,2 | 92,8 | 15,8 | 5,1 | 9,6 |
| 1 enfant | 87,0 | 97,6 | 91,5 | 26,1 | 3,2 | 15,4 |
| 2 enfants | 82,7 | 97,5 | 89,3 | 37,7 | 2,5 | 20,2 |
| 3 enfants ou plus | 63,7 | 96,6 | 78,1 | 47,6 | 3,3 | 22,8 |

a. Quelle est la population étudiée ? Quels sont les caractères retenus dans le tableau ? Que représentent les nombres : des effectifs conjoints ? des fréquences conjointes, marginales, conditionnelles ? (si oui, de quoi ?) ou autre chose ?

b. Peut on déduire de la première ligne du tableau qu’il y a plus d’hommes que de femmes, ou l’inverse, au sein de la population étudiée ?

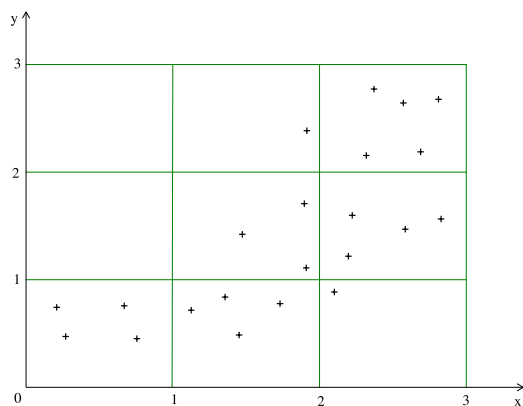
4. (Examen 1ère session mai 2010) On fait une étude statistique des filières (mathématiques, sciences de la vie, etc.) choisies par les étudiants du campus sciences.



Le graphique montre la fréquence des différentes filières parmi les étudiants garçons du campus (première barre) et parmi la population totale du campus (seconde barre).

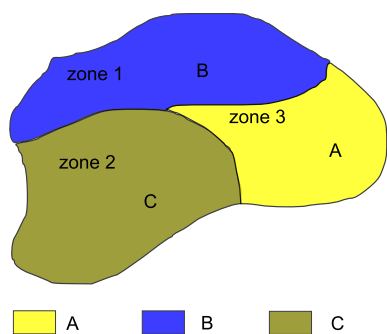
- Quelle est approximativement la fréquence de la filière "Chimie" chez les garçons ? Et quelle est approximativement la fréquence de cette filière dans la population totale ?
- Quelles sont les filières dont on observe pratiquement l'indépendance par rapport au sexe de l'étudiant ? Expliquez.

5. (Examen 1ère session avril 2012) Une population de 22 individus est étudiée via deux caractères quantitatifs X et Y . Le dessin ci-dessus montre le nuage de points $(X(s), Y(s))$ associé.



- Quel est approximativement l'étendue du caractère Y ?
- On découpe les étendues des caractères X et Y en intervalles $[0, 1[$, $[1, 2[$, $[2, 3[$. Donnez le tableau des effectifs conjoints des intervalles de X et Y .
- Quelle est la fréquence de l'évènement " $X \in [0, 1[$ " ?
Quelle est la fréquence de l'évènement " $X \in [0, 1[$ " conditionné à l'évènement " $Y \in [0, 1[$ " ?
- Peut-on affirmer que les caractères X et Y sont liés (ne sont pas pratiquement indépendants) ? Justifiez votre réponse.

6. Trois candidats A, B et C se disputent une élection. La carte ci-dessous montre le candidat arrivé en tête de l'élection dans chacune des trois zones géographiques 1, 2, 3 :



- Quelle population a-t-on raisonnablement étudiée et par quels caractères pour obtenir cette carte ?
Quelle information donne la carte sur les caractères (dans les termes de la statistique descriptive) ?
- L'information que donne cette carte suffit-elle pour observer une liaison au sens statistique entre le suffrage des électeurs et la zone géographique ?
- Il y a 2000 votants en zone 1, 4000 en zone 2 et 3500 en zone 3. Le dépouillement dans chacune des zones a donné les résultats suivants :
zone 1 : 29% pour A, 40% pour B, 31% pour C
zone 2 : 33% pour A, 33% pour B, 34% pour C
zone 3 : 40% pour A, 39% pour B, 21% pour C

Calculer le nombre $\chi^2(X, Y)$ où X, Y sont les deux caractères étudiés. Quel couple de valeurs contribue le plus à une liaison entre X et Y ?