Université Claude Bernard - Lyon 1 — **Semestre automne 2023-2024**

L2 - Informatique — UE : Statistiques pour l'Informatique

Corrigé du CC1 du 18 octobre 2023, durée 60 min

Les documents et calculatrices ne sont pas permises. Le barème est indicatif.

Question de cours (*2 pts.*) Que veut dire que trois événements A, B, C sont indépendants ?

Corrigé : A, B, C sont indépendants si

$$P(A\cap B)=P(A)P(B),\quad P(A\cap C)=P(A)P(C),\quad P(B\cap C)=P(B)P(C)$$

et

$$P(A\cap B\cap C)=P(A)P(B)P(C)$$

Exercice 1 (*6 pts.*) Dans un groupe de personnes choisis au hasard, on a relevé la taille x et le poids y de chaque personne. Les valeurs suivantes ont été obtenues :

Taille (m)	1.60	1.60	1.70	1.80	1.80
Poids (kg)	55	60	75	85	90

On admet que l'écart type (sans biais) de x vaut 0.1 m, l'écart type (sans biais) de y vaut $15,25$ kg et la covariance (sans biais) de x et y vaut 1.5 kg m.

1. Construisez un nuage de points (diagramme de dispersion) avec la taille sur l'axe x et le poids sur l'axe y.
2. Déterminez la moyenne et la médiane de x et de y.
3. Déterminez l'équation de la droite de régression permettant d'estimer le poids en fonction de la taille.
4. Quel devrait être le poids d'une personne de taille 1.75 m ?

Corrigé :

1.
2. La moyenne de x est $\bar{x}=\frac{1}{5}(1.6+1.6+1.7+1.8+1.8)\,\mathrm{m}=1.7$ m.
   La médiane de x est 1.7 m, car 5 est impair.
   La moyenne de y est $\bar{y}=\frac{1}{5}(55+60+75+85+90)\,\mathrm{kg}=\frac{365}{5}$ kg $=73$ kg.
   La médiane de y est 75 kg, car 5 est impair.
3. L'équation de la droite est $y=ax+b$ avec
   $a=\frac{V(x,y)}{V(x,x)}=\frac{1.5\,\mathrm{kg\,m}}{0{,}01^2\,\mathrm{m}^2}=150$ kg/m
   $b=\bar{y}-a\bar{x}=73\,\mathrm{kg}-150\,\mathrm{kg/m}\times 1.7\,\mathrm{m}=-182$ kg.
4. Le poids devait être $a\times 1.75\,\mathrm{m}+b=150\times 1.75\,\mathrm{kg}-182\,\mathrm{kg}=80.5$ kg.

Exercice 2 (*4 pts.*) Dans une entreprise, 60% des employés sont des hommes. Parmi les employés de l'entreprise, 10% fument. Parmi les femmes employées dans l'entreprise, 15% fument.

1. Quelle est la proportion de fumeurs féminins parmi les employés de l'entreprise ?
2. Quelle est la probabilité qu'un employé de l'entreprise qui fume soit un homme ?
3. Quelle est la probabilité qu'un employé de l'entreprise qui est un homme soit un fumeur ?

Corrigé : On pose H = employé de l'entreprise masculin, F = employé de l'entreprise feminin, f = employé de l'entreprise qui fume. L'énoncé nous indique que $P(H) = 0.6$, $P(f) = 0.1$ et $P(f|F) = 0.15$. Clairement, $P(H) + P(F) = 1$ et aussi $P(H|f) + P(F|f) = 1$.

1. Ici, il faut déterminer $P(F \cap f)$. Or, $P(F \cap f) = P(f|F)P(F) = P(f|F)(1 - P(H)) = 0.15 \times 0.4 = 0.06$.
2. Ici, il faut déterminer $P(H|f)$. Or, $P(H|f) = 1 - P(F|f) = 1 - \frac{P(F \cap f)}{P(f)} = 1 - 0.06/0.1 = 0.4$.
3. Ici, il faut déterminer $P(f|H)$. Or, $P(f|H) = \frac{P(f \cap H)}{P(H)} = \frac{P(H|f)P(f)}{P(H)} = 0.4 \times 0.1/0.6 = 2/30 \cong 0.067$.

Exercice 3 (*4 pts.*) Dans un jeu de cartes il y a 4 cartes marquées avec les chiffres $1, \cdots, 4$. Alice tire une carte au hasard qu'elle garde. Puis, après Alice, Bob tire une carte au hasard qu'il garde. Soit A l'événement "Alice tire une carte avec un chiffre pair". Soit B l'événement "Bob tire une carte avec un chiffre impair".

1. Proposez un espace des événements qui contient les deux événements A et B.
2. Quelle est la probabilité que Alice tire une carte avec un chiffre pair ?
3. Quelle est la probabilité que Bob tire une carte avec un chiffre impair ?
4. Est-ce que les événements A et B sont indépendantes ? Justifiez votre reponse.

Corrigé :

1. Le jeu contient 4 cartes, il y a 2 tirages et on prend en compte leur ordre. Donc il y a $4 \times 3 = 12$ événements simples, notament

$$\Omega = \{12, 13, 14, 21, 23, 24, 31, 32, 34, 41, 42, 43\}$$

   Lévénement A est que la première carte est un chiffre pair, donc

$$A = \{21, 23, 24, 41, 42, 43\}$$

   Lévénement B est que la deuxième carte est un chiffre impair, donc

$$B = \{13, 21, 23, 31, 41, 43\}$$

2. Ici, il faut déterminer $P(A)$. Or, $P(A) = \frac{card\,A}{card\,\Omega} = \frac{1}{2}$.
3. Ici, il faut déterminer $P(B)$. Or, $P(B) = \frac{card\,B}{card\,\Omega} = \frac{1}{2}$.
4. Ici, il faut vérifier si $P(A \cap B) = P(A)P(B)$. Or,

$$A \cap B = \{21, 23, 41, 43\}$$

   D'ou $P(A \cap B) = \frac{card\,A \cap B}{card\,\Omega} = \frac{1}{3}$. Comme $\frac{1}{3} \neq \frac{1}{2} \times \frac{1}{2}$ les deux événements ne sont pas indépendants.

Exercice 4 *(4 pts.)* Soient $X = \{x_1, \cdots, x_n\}$ et $Y = \{y_1, \cdots, y_n\}$ deux échantillons numériques de même taille.

1. Donnez la formule pour le coefficient de corrélation de ces deux variables.
2. On multiplie toutes les valeurs x_i de X par 10. Comment cela affecte le coefficient de corrélation ?

Corrigé :

1. $cor(X,Y) = \frac{V(X,Y)}{\sqrt{V(X,X)V(Y,Y)}}$ où $V(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ et $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ et $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. (On pourrait aussi prendre la covariance sans biais dans cette formule, c.à.d. remplacer $\frac{1}{n}$ par $\frac{1}{n-1}$ devant les expressions pour $V(X,Y)$, $V(X,X)$ et $V(Y,Y)$, mais pas pour $\bar{x}$ ou $\bar{y}$).
2. $V(10X,Y) = 10V(X,Y)$, $V(10X,10X) = 10^2V(X,X)$, pendant que $V(Y,Y)$ reste inchangé.

   Donc $cor(10X,Y) = \frac{10}{\sqrt{10^2}}cor(X,Y) = cor(X,Y)$. Le coefficient de corrélation reste inchangé.