

Correction CC1

Exercice 1 : Lors d'une campagne de mesures de la taille de manchots en antarctique, on a mesuré la taille de 6 mères et de leurs enfants de six mois (pour les manchots, il y a une seule naissance par année). Dans le tableau ci-dessous X la taille de la mère et Y la taille de son enfant, exprimées en pouces.

X	48.6	44.1	42.3	49.2	44.1	47.7
Y	27.9	22.8	20.7	28.2	24.6	28.8

1. Calculer $\bar{X}$, σ_X . Donner la médiane et les quartiles de X . Dessiner ensuite le diagramme de Tukey.
2. Calculer $\bar{Y}$, σ_Y .
3. Calculer $\text{Cov}(X, Y)$, puis le coefficient de corrélation $r(x, y)$ entre X et Y . Que pouvez-vous dire de la valeur du coefficient de corrélation?
4. Déterminer la droite de regression de Y par rapport à X . A l'aide de cette droite de regression, donner une estimation de la taille d'un enfant de la colonie dont la mère a une taille de 46 pouces.

1. On calcule, ici on a un tableau de donnée avec 6 individus. On calcule alors:

$$\bar{X} = \frac{1}{6} \sum_{i=1}^6 X_i = 46.$$

Également $\sigma_X = \sqrt{\text{Var}(X)}$. On calcule la variance:

$$\text{Var}(X) = \frac{1}{6} \sum_{i=1}^6 (X_i - \bar{X})^2 = 6.7$$

Ce qui nous donne $\sigma_X = 2.6$. Pour calculer la médiane et les quartiles, on ordonne nos données, on note $X^{(i)}$, la i -ème donnée ordonnée:

42.3	44.1	44.1	47.7	48.6	49.2
------	------	------	------	------	------

Puisqu'on a 6 observation, la médiane est la moyenne de la 3-ème et de la 4-ème valeur.

$$m = \frac{44.1 + 47.7}{2} = 45.9$$

.

Pour le premier quartile, on remarque que $6 \times \frac{1}{4} = 1.5$ et $1 \leq 1.5 \leq 2$, on peut donc prendre $X^{(2)}$ comme premier quartile de X . De même, $6 \times \frac{3}{4} = 4.5$, on peut donc prendre $X^{(5)}$.

Une fois toutes ces quantités calculées, on peut faire le diagramme de Tukey.

2. On procède comme dans la question précédente :

$$\bar{Y} = \frac{1}{6} \sum_{i=1}^6 Y_i = 25.$$

Et

$$\sigma_Y = \sqrt{\text{Var}(Y)} = 3.02$$

3. On peut calculer la covariance:

$$\text{Cov}(X, Y) = \frac{1}{6} \sum_{i=1}^6 (X_i - \bar{X})(Y_i - \bar{Y}) = 7.5$$

Puisque $r(x, y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$, on en déduit:

$$r(x, y) = 0.95$$

On remarque que $r(X, Y) \simeq 1$, X et Y sont donc fortement corrélé positivement.

4. On détermine la droite de régression, on peut poser $y = ax + b$, avec $a = \frac{\text{Cov}(x, y)}{\text{Var}(x)} = 1.32$ et $b = \bar{Y} - a\bar{X} = -35.5$. Une fois cette droite obtenue, on l'utilise pour prévoir la taille d'un enfant donc la mère a une taille de 46 pouce via la droite: on obtient $a \times 46 + b = 25.5$.

Exercice 2: Une maladie M affecte un français sur mille. On dispose d'un test sanguin T qui détecte M avec une fiabilité de 95% lorsque cette maladie est effectivement présente. Cependant, on obtient un résultat faussement positif pour 1% des personnes saines testées. On note M l'évènement "la personne choisie au hasard est malade" et T l'évènement "le test de la personne choisie au hasard est positif".

1. Calculer $\mathbb{P}(M|T)$ et $\mathbb{P}(\bar{M}|T)$.
2. Calculer $\mathbb{P}(M|\bar{T})$ et $\mathbb{P}(\bar{M}|\bar{T})$.

Avant de commencer l'exercice, on note mathématiquement les évènements qui nous sont donnés dans l'énoncé:

$$\mathbb{P}(M) = \frac{1}{1000}, \quad \mathbb{P}(T|M) = 0.95, \quad \mathbb{P}(T|\bar{M}) = 0.01$$

1. La formule de Bayes nous donne:

$$\mathbb{P}(M|T) = \frac{\mathbb{P}(M \cap T)}{\mathbb{P}(T)}$$

Or, on ne connaît pas $\mathbb{P}(T)$, on peut donc appliquer la formule des probabilités totales:

$$\begin{aligned} \mathbb{P}(T) &= \mathbb{P}(T \cap M) + \mathbb{P}(T \cap \bar{M}) \\ &= \mathbb{P}(M)\mathbb{P}(T|M) + \mathbb{P}(\bar{M})\mathbb{P}(T|\bar{M}) \\ &= 0.01 \end{aligned}$$

On a alors :

$$\begin{aligned}
 \mathbb{P}(M|T) &= \frac{\mathbb{P}(M \cap T)}{\mathbb{P}(T)} \\
 &= \frac{\mathbb{P}(M)\mathbb{P}(T|M)}{\mathbb{P}(T)} \\
 &= 0.095
 \end{aligned}$$

On peut aussi calculer $\mathbb{P}(\bar{M}|T) = 1 - \mathbb{P}(M|T) = 0.905$

2. Pour cette question, il fallait faire la même chose, avec $\mathbb{P}(\bar{T}) = 1 - \mathbb{P}(T)$. C'est-à-dire, on utilise la loi des probabilités totales puis la loi de Bayes.

Exercice 3 : Deux lemmings effectuent leur grande migration. Le lemming 1 a trois chances sur quatre de réussir. Le lemming 2 a deux chances sur trois de réussir si le lemming 1 réussit, mais seulement une chance sur trois si le lemming 1 échoue. Pour $i = 1, 2$, on note B_i la variable valant 1 si le lemming i réussit et 0 sinon.

1. Calculer $\mathbb{P}(B_1 = 1, B_2 = 1)$.
2. Calculer $\mathbb{P}(B_2 = 0|B_1 = 1)$, puis $\mathbb{P}(B_1 = 1, B_2 = 0)$.
3. Calculer, par des méthodes similaires, $\mathbb{P}(B_1 = 0, B_2 = 1)$ ainsi que $\mathbb{P}(B_1 = 0, B_2 = 0)$.
4. Soit $N = B_1 + B_2$. Donner la loi de N , puis calculer $\mathbb{E}(N)$. En déduire $\mathbb{E}(B_2)$ (on donnera $\mathbb{E}(B_1)$ avant). Déterminer alors la loi de B_2 .

Comme dans l'exercice, regardons les informations qui nous sont données par l'énoncé:

$$\mathbb{P}(B_1 = 1) = \frac{3}{4}, \quad \mathbb{P}(B_2 = 1|B_1 = 1) = \frac{2}{3}, \quad \mathbb{P}(B_2 = 1|B_1 = 0) = \frac{1}{3}$$

1. Pour cela, on utilise la formule de Bayes:

$$\mathbb{P}(B_1 = 1, B_2 = 1) = \mathbb{P}(B_1 = 1) \times \mathbb{P}(B_2 = 1|B_1 = 1)$$

2. $\mathbb{P}(B_2 = 0|B_1 = 1) = 1 - \mathbb{P}(B_2 = 1|B_1 = 1),$

on peut alors réutiliser la formule de Bayes, comme à la question 1. Ce qui donne:

$$\mathbb{P}(B_1 = 1, B_2 = 0) = \mathbb{P}(B_1 = 1) \times \mathbb{P}(B_2 = 0|B_1 = 1)$$

3. Pour cette question, il fallait faire comme pour la précédente: on calcule une proba conditionnelle qui nous permet ensuite de calculer la proba de l'intersection grâce à la loi de Bayes. On connaît déjà $\mathbb{P}(B_2 = 1|B_1 = 0)$.

4. $N = B_1 + B_2$, pour trouver la loi de N , on utilise les questions précédentes, il vient:

$$\begin{aligned}
 \mathbb{P}(N = 0) &= \mathbb{P}(B_1 = 0, B_2 = 0) = \\
 \mathbb{P}(N = 1) &= \mathbb{P}(B_1 = 1, B_2 = 0) + \mathbb{P}(B_1 = 0, B_2 = 1) = \\
 \mathbb{P}(N = 2) &= \mathbb{P}(B_1 = 1, B_2 = 1) =
 \end{aligned}$$

On peut alors calculer $\mathbb{E}[N]$:

$$\mathbb{E}[N] = 0 \times \mathbb{P}(N = 0) + 1 \times \mathbb{P}(N = 1) + 2 \times \mathbb{P}(N = 2) = \frac{5}{4}.$$

On a $\mathbb{E}[B_1] = 1 \times \mathbb{P}(B_1 = 1) = \frac{3}{4}$ Alors,

$$\mathbb{E}[B_2] = \mathbb{E}[N - B_1] = \mathbb{E}[N] - \mathbb{E}[B_1] = \frac{2}{4}$$

On a alors $B_2 \sim \mathcal{B}(\frac{1}{2})$.

Exercice 4 : On lance une pièce équilibrée autant de fois qu'il faut pour obtenir deux fois face. Soit N le nombre de lancers nécessaires pour cela ($N > 2$).

1. Soit B_k la variable aléatoire égale à 1 si le lancer k donne face et 0 sinon. Soit aussi S_m le nombre de face obtenus lors des m premiers lancers. Quelle est la loi de B_k ? Quelle est la loi de S_m ?

2. Montrer que, pour tout

$$n > 2, \{N = n\} = \{S_{n-1} = 1 \text{ et } B_n = 1\}.$$

En déduire que

$$P(N = n) = P(S_{n-1} = 1)P(B_n = 1).$$

3. En partant de 2. déterminer la loi de N .

1. Pour tout $k \in \mathbb{N}$, $B_k \sim \mathcal{B}(\frac{1}{2})$, $S_m \sim \mathcal{B}(m, \frac{1}{2})$.

2. $\{N = n\} = \{\text{On a obtenu face une deuxième fois au } n\text{-ème essai}\}$. Ce qui veut dire, qu'au rang $n - 1$, on avait obtenu face déjà une fois ($S_{n-1} = 1$), et qu'au rang n , on obtient aussi face $B_n = 1$. D'où $\{N = n\} = \{S_{n-1} = 1 \text{ et } B_n = 1\}$ Puisque S_{n-1} et B_n sont supposé indépendants, on obtient que:

$$\mathbb{P}(N = n) = \mathbb{P}(S_{n-1} = 1) \times \mathbb{P}(B_n = 1)$$

3. On peut alors déterminer la loi de N :

$$\begin{aligned} \mathbb{P}(N = n) &= \binom{n-1}{1} \left(\frac{1}{2}\right)^{n-1} \times \frac{1}{2} \\ &= \frac{n-1}{2^n} \end{aligned}$$

Car $\binom{n-1}{1} = \frac{(n-1)!}{(n-2)!1!} = n-1$, et $\mathbb{P}(S_{n-1} = 1) = \binom{n-1}{1} \left(\frac{1}{2}\right)^{n-2} \times \frac{1}{2}$.