TD- Licence 3 MASS

Exercice 1 *Deux machines industrielles $\mathcal{M}_1$ et $\mathcal{M}_2$ ont des taux de production quotidienne d'objets défectueux égaux à $p_1 = 5\%$ et $p_2 = 10\%$. Chacune de ces machines $\mathcal{M}_1$ et $\mathcal{M}_2$ produit respectivement $m_1 = 100$, et $m_2 = 200$ objets. Qu'elle est la probabilité pour qu'un object pris au hasard soit défectueux ? Qu'elle est la probabilité pour que ce soit la première machine $\mathcal{M}_1$ qui l'ait produit ?*

Solution :
Pour comprendre ce problème de statistique, il convient de noter que l'on a supposé implicitement que les taux d'objets défectueux p_1, p_2 ont été calculés sur la même base de production. Autrement dit, $\mathcal{M}_1$ et $\mathcal{M}_2$ produisent chaque jours

$$d_1 = m_1 p_1 = 5 \quad \text{et} \quad d_2 = m_2 p_2 = 20$$

objets défecteux. On représente le choix au hasard d'un objet parmis les m_1+m_2 objet par l'ensemble des indices $\Omega = \{1, \ldots, (m_1+m_2)\}$, muni de la tribu discrète $\mathcal{F} = \mathcal{P}(\Omega)$, et de la probabilité uniforme

$$\forall \omega \in \Omega \qquad \mathbb{P}(\omega) = \frac{1}{m_1 + m_2} = \frac{1}{300}$$

Sans perte de généralité, et à une permutation d'indices près, on peut supposer que les évenements $A_1 = \{1, \ldots, m_1\}$, et $A_2 = m_1 + \{1, \ldots, m_2\}$, représentent respectivement les objets produits par $\mathcal{M}_1$, et par $\mathcal{M}_2$. Le couple $\{A_1, A_2\}$ forme clairement une partition de Ω, et l'on a

$$\mathbb{P}(A_1) = 1 - \mathbb{P}(A_2) = \frac{|A_1|}{|\Omega|} = \frac{m_1}{m_1 + m_2} = \frac{1}{3}$$

L'ensemble des objets défectueux quotidien B est formé de $d_1 + d_2$ objets, dont d_1 dans A_1, et d_2 dans A_2. On retrouve ainsi le fait que

$$\mathbb{P}(B|A_i) = |B \cap A_i|/|A_i| = d_i/m_i = p_i$$

On obtient alors

$$\begin{aligned} \mathbb{P}(B) &= \sum_{i \in \{1,2\}} \mathbb{P}(B|A_i)\mathbb{P}(A_i) \\ &= \frac{m_1 p_1 + m_2 p_2}{m_1 + m_2} = \frac{d_1 + d_2}{m_1 + m_2} = \frac{1}{60} + \frac{1}{15} = \frac{1}{12} \end{aligned}$$

Ce résultat peut être retrouvé de façon immédiate, en notant que tout simplement que $\mathbb{P}(B) = |B|/|\Omega|$. Ceci répond à la première question. Les probabilités pour que l'une des machines $\mathcal{M}_i$, $i = 1, 2$, est produit cet objet défecteux, correspondent aux probabilités $\mathbb{P}(A_i|B)$, $i = 1, 2$, de choisir cet objet dans A_i, $i = 1, 2$, sachant que ce dernier est dans B. D'après la formule de Bayes, on a

$$\begin{aligned}
\mathbb{P}(A_i|B) &= \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{j\in\{1,2\}} \mathbb{P}(B|A_j)\mathbb{P}(A_j)} \\
&= \frac{p_i m_i}{p_1 m_1 + p_2 m_2} = \frac{d_i}{d_1 + d_2} \quad \left(= \frac{1}{5}\ 1_{i=1} + \frac{4}{5}\ 1_{i=2}\right)
\end{aligned}$$

On remarque que cette formule est bien cohérente avec notre intuition. En effet, si l'on restreint notre choix aux pièces défecteuses, on a bien évidemment $d_i/(d_1+d_2) = |A_i \cap B|/|B|$ "chances" pour qu'elle provienne de la machine $\mathcal{M}_i$! ■

Exercice 2 *Un sac contient* 2 *pièces de monnaie. L'une équitable, et ayant une probabilité* 1/2 *de donner "pile" ou "face" ; la seconde ayant une probabilité* 1/3 *de donner "face". On lance au hasard l'une des pièces, et l'on observe un résultat "face". Quelle est la probabilité d'avoir choisi la pièce équitable ?*

Solution :
Cette expérience aléatoire se déroule en deux temps. On choisit tout d'abord une pièce $\omega_1 = \{0, 1\}$ dans un sac Ω_1 contenant deux pièces de monnaie ; la valeur $\omega_1 = 1$ correspond au choix de la pièce équitable, et la valeur $\omega_1 = 0$ représente le choix de la pièce ayant une probabilité 1/3 de donner le coté"face". Une fois la pièce choisie, on la lance et l'on observe un résultat $\omega_2 \in \Omega_2 = \{0, 1\}$. Dans ce cas, la valeur $\omega_2 = 1$ correspond au coté "face", et la valeur $\omega_2 = 0$ représente le coté "pile". L'exercice peut être résolu selon deux approches connexes. La première consiste à décrire avec précision l'espace des évènements possibles. La seconde est liée à une interprétation dynamique naturelle de l'expérience en question.

1. L'espace des épreuves associé à l'expérience complète est donc donné par $\Omega = \{0, 1\}^2$. Chaque évènement élémentaire $\omega = (\omega_1, \omega_2) \in \{0, 1\}^2$, represente le choix de la pièce ω_1, et le resultat du lancer ω_2. D'après l'énoncé, nous avons
$$\mathbb{P}(\{(1, 1)\}) = \frac{1}{2} \times \frac{1}{2} = \mathbb{P}(\{(1, 0)\})$$
et
$$\mathbb{P}(\{(0, 1)\}) = \frac{1}{2} \times \frac{1}{3}, \qquad \mathbb{P}(\{(0, 0)\}) = \frac{1}{2} \times \frac{2}{3}$$
Les évènements
$$\begin{aligned}
F &= \{(0, 1), (1, 1)\}, \quad F^c = \{(0, 0), (1, 0)\} \\
E &= \{(1, 0), (1, 1)\}, \quad E^c = \{(0, 1), (0, 0)\}
\end{aligned}$$
représentent respectivement la situation où le résultat du lancer est "face" (F), celle où le lancer est "pile"(F^c), le fait d'avoir choisi la pièce équitable

(E), et celui d'avoir choisi la pièce non équitable (E^c). On a donc

$$\mathbb{P}(F) = \frac{1}{6} + \frac{1}{4} \quad \text{et} \quad \mathbb{P}(E) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

De même on montre que

$$\mathbb{P}(F|E) = \frac{1/4}{1/4 + 1/4} = \frac{1}{2} = 1 - \mathbb{P}(F^c|E) \quad \text{et} \quad \mathbb{P}(F|E^c) = \frac{1/6}{1/2} = \frac{1}{3}$$

On en conclut que

$$\begin{aligned}\mathbb{P}(E|F) &= \frac{\mathbb{P}(F|E)\mathbb{P}(E)}{\mathbb{P}(F|E)\mathbb{P}(E) + \mathbb{P}(F|E^c)\mathbb{P}(E^c)} = \frac{1/4}{1/4 + 1/6} \\ &= \frac{3}{5} = 1 - \mathbb{P}(E^c|F)\end{aligned}$$

2. La seconde méthode consiste à exploiter l'aspect temporel de l'expérience, de sorte à éviter une représentation complète de l'espace des épreuves. On note X_1 et X_2 les variables aléatoires à valeurs dans $\{0, 1\}$ représentant le choix initial X_1 de la pièce, et le résultat du lancer. D'après l'énoncé, le choix de la pièce dans le sac est uniforme

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = 0) = \frac{1}{2}$$

D'autre part, sachant que la pièce est equitable, le résultat du lancer sera encore uniforme

$$\mathbb{P}(X_2 = 1|X_1 = 1) = \mathbb{P}(X_2 = 0|X_1 = 1) = \frac{1}{2}$$

Si la pièce choisie n'est pas équitable, alors nous aurons une probabilité 1/3 de voir le coté"face" représenté par la valeur 1

$$\mathbb{P}(X_2 = 1|X_1 = 0) = \frac{1}{3} = 1 - \mathbb{P}(X_2 = 0|X_1 = 0)$$

On retrouve ainsi le résultat recherché

$$\begin{aligned}&\mathbb{P}(X_1 = 1|X_2 = 1) \\ &= \tfrac{\mathbb{P}(X_2=1|X_1=1)\mathbb{P}(X_1=1)}{\mathbb{P}(X_2=1|X_1=1)\mathbb{P}(X_1=1)+\mathbb{P}(X_2=1|X_1=0)\mathbb{P}(X_1=0)} \\ &= \tfrac{1/4}{1/4+1/6} = \tfrac{3}{5} = 1 - \mathbb{P}(X_1 = 0|X_2 = 1)\end{aligned}$$

On remarquera que cette méthode est une reformulation dynamique de la précédente. Plus précisément, si l'on considère les variables canoniques

$$X_1(\omega_1, \omega_2) = \omega_1 \quad \text{et} \quad X_2(\omega_1, \omega_2) = \omega_2$$

les équations précédentes sont une reformulation de celles décrites dans la première méthode.

$\blacksquare$

Exercice 3 (Formule de Bayes séquentielle) *Montrer que la probabilité conditionnelle d'une réalisation conjointe de n évènements $(A_p)_{1\leq p\leq n}$ par rapport à un évènement B (de probabilité non nulle) est donnée par la formule multiplicative*

$$\mathbb{P}(\cap_{p=1}^n A_n|B)$$

$$= \mathbb{P}(A_n|B\cap[\cap_{p=1}^{n-1}A_p])\ \mathbb{P}(A_{n-1}|B\cap[\cap_{p=1}^{n-2}A_p])\ldots\mathbb{P}(A_2|B\cap A_1)\mathbb{P}(A_1|B)$$

$$= \prod_{p=1}^n \mathbb{P}(A_p|B\cap[\cap_{q=1}^{p-1}A_q])$$

avec la convention $\prod_\emptyset = \Omega$, lorsque $p=1$.

Solution :
La formule multiplicative recherchée est une conséquence directe de la formule récursive suivante

$$\mathbb{P}(A_n\cap[\cap_{p=1}^{n-1}A_n]\mid B) = \mathbb{P}(A_n\mid B\cap[\cap_{p=1}^{n-1}A_p])\ \mathbb{P}(\cap_{p=1}^{n-1}A_n\mid B)$$

$\blacksquare$

Exercice 4 *Soient $X_1,\ldots,X_n$ une suite de n copies indépendantes d'une v.a. de Bernoulli X sur $\{0,1\}$*

$$\mathbb{P}(X=1) = 1-\mathbb{P}(X=0) = p\in[0,1]$$

On note $S_k=\sum_{i=1}^k X_i$, les variables de comptage du nombre de succés. Montrer que pour tout $m\leq n$, et $0\leq k\leq l\leq n$, on a

$$\mathbb{P}^{S_m|S_n}(k|l) = \frac{C_{n-m}^{l-k}\ C_m^k}{C_n^l} \quad et \quad \mathbb{P}^{X_1|S_n}(1|l) = \frac{l}{n}$$

Solution :
Par définition des v.a. $S_k=\sum_{i=1}^k X_i$, nous avons pour tout $m\leq n$, et $0\leq k\leq l\leq n$

$$\begin{aligned}
\mathbb{P}^{S_m|S_n}(k|l) &= \frac{\mathbb{P}(S_n=l,\ S_m=k)}{\mathbb{P}(S_n=l)} = \frac{\mathbb{P}(S_n-S_m=(l-k),\ S_m=k)}{\mathbb{P}(S_n=l)}\\
&= \frac{\mathbb{P}(\sum_{i=m+1}^n X_i=(l-k),\ \sum_{i=1}^m X_i=k)}{\mathbb{P}(S_n=l)}\\
&= \frac{\mathbb{P}(S_{n-m}=(l-k))\times\mathbb{P}(S_m=k)}{\mathbb{P}(S_n=l)}\\
&= \frac{\left[C_{n-m}^{l-k}\ p^{l-k}\ (1-p)^{(n-m)-(l-k)}\right]\left[C_m^k\ p^k\ (1-p)^{m-k}\right]}{C_n^l\ p^l\ (1-p)^{n-l}}\\
&= \frac{C_{n-m}^{l-k}\ C_m^k}{C_n^l}
\end{aligned}$$

Lorsque $m = 1$, et $k = 1$, on obtient clairement

$$\mathbb{P}^{X_1|S_n}(1|l) = \frac{C_{n-1}^{l-1}\ C_1^1}{C_n^l} = l/n$$

$\blacksquare$

Exercice 5 *Soit* $(X_p)_{0\leq p\leq n}$ *une suite de v.a. définies sur un même espace de probabilités* $(\Omega, \mathcal{F}, \mathbb{P})$*, et à valeurs dans des espaces au plus dénombrables* $(E_p, \mathcal{P}(E_p))_{0\leq p\leq n}$*. Supposons que la loi conditionnelle de la séquence* $(X_1, \ldots, X_n)$ *en l'évènement* $\{X_0 = x_0\}$ *soit de la forme*

$$\begin{aligned}&\mathbb{P}^{(X_1\ldots,X_n)|X_0}(x_1,\ldots,x_n|x_0)\\ &= p_n(x_n|x_0,\ldots,x_{n-1})\ p_{n-1}(x_{n-1}|x_0,\ldots,x_{n-2})\ldots p_2(x_2|x_0,x_1)p_1(x_1|x_0)\end{aligned} \tag{1}$$

Dans la formule précédente, p_k *désigne une suite de fonctions boréliennes positives, et telles que* $\sum_{x_k} p_k(x_k|x_0,\ldots,x_{k-1}) = 1$*, pour tout* $k \geq 1$*, et pour toute séquence* $(x_0,\ldots,x_{k-1}) \in (E_0 \times \ldots E_{k-1})$*. Montrer que l'on a nécessairement*

$$\mathbb{P}^{X_k|(X_0,\ldots,X_{k-1})}(x_k|x_0,\ldots,x_{k-1}) = p_k(x_k|x_0,\ldots,x_{k-1})$$

Solution :
Tout d'abord, on note que

$$\mathbb{P}^{X_n|(X_0\ldots,X_{n-1})} = \frac{\mathbb{P}^{(X_1\ldots,X_n)|X_0}(x_1,\ldots,x_n|x_0)}{\mathbb{P}^{(X_1\ldots,X_{n-1})|X_0}(x_1,\ldots,x_{n-1}|x_0)}$$

D'après l'équation (1), on a nécessairement

$$\mathbb{P}^{X_n|(X_0\ldots,X_{n-1})} = p_n(x_n|x_0,\ldots,x_{n-1})$$

$\blacksquare$

Exercice 6 *Soit* (X_1, X_2) *un couple de v.a. à valeurs dans un espace produit fini* $(E_1 \times E_2)$*, muni de l'algèbre discrète* $\mathcal{P}(E_1 \times E_2)$*. Montrer que pour toute fonction* $f : E_2 \to \mathbb{R}$*, on a la formule*

$$\begin{aligned}
&\mathbb{E}((f(X_2) - \mathbb{E}(f(X_2) \mid X_1))^2) \\
&= \inf \{\mathbb{E}((f(X_2) - Y_1)^2) \ : \ Y_1 \text{ v.a. } \sigma(X_1) - \text{mesurable}\}
\end{aligned}$$

Cette description variationnelle montre l'espérance conditionnelle est un estimateur optimal, en ce sens où elle minimise la variance d'erreur.

Solution :
En utilisant la formule

$$\forall Y_1 \underline{\in} \sigma(X_1) \qquad \mathbb{E}(Y_1 \ f_2(X_2)) = \mathbb{E}(Y_1 \ \mathbb{E}(f_2(X_2) \mid X_1))$$

on montre que pour toute v.a. Y_1 mesurable par rapport à $\sigma(X_1)$ on a

$$\mathbb{E}(Y_1 \ [f_2(X_2) - \mathbb{E}(f_2(X_2) \mid X_1)]) = 0$$

Si l'on remplace Y_1 par la v.a. $[\mathbb{E}(f_2(X_2) \mid X_1) - Y_1]$, on trouve

$$\mathbb{E}([\mathbb{E}(f_2(X_2) \mid X_1) - Y_1] \ [f_2(X_2) - \mathbb{E}(f_2(X_2) \mid X_1)]) = 0$$

On démontre ainsi que pour toute v.a. Y_1 mesurable par rapport à $\sigma(X_1)$

$$\begin{aligned}
&\mathbb{E}((f_2(X_2) - Y_1)^2) \\
&= \mathbb{E}(([f_2(X_2) - \mathbb{E}(f_2(X_2) \mid X_1)] + [\mathbb{E}(f_2(X_2) \mid X_1) - Y_1])^2) \\
&= \mathbb{E}([f_2(X_2) - \mathbb{E}(f_2(X_2) \mid X_1)]^2) + \mathbb{E}([\mathbb{E}(f_2(X_2) \mid X_1) - Y_1]^2)
\end{aligned}$$

La preuve du résultat recherché est désormais claire. ■

Exercice 7 *Soit* (X_1, X_2, X_3) *un triplet de v.a. à valeurs un espace produit fini* $(E_1 \times E_2 \times E_3)$*, muni de l'algèbre discrète. Montrer que pour toute fonction* $f : E_3 \to \mathbb{R}$*, on a la formule de conditionnements emboités*

$$\mathbb{E}(\mathbb{E}(f(X_3) \mid X_2, X_1) \mid X_1) = \mathbb{E}(f(X_3) \mid X_1)$$

Solution :
On remarque tout d'abord que

$$\mathbb{P}^{X_3, X_2 | X_1}(x_3, x_2 | x_1) = \mathbb{P}^{X_3 | X_1, X_2}(x_3 | x_1, x_2) \ \mathbb{P}^{X_2 | X_1}(x_2 | x_1)$$

En sommant sur la seconde coordonnée, on obtient

$$\mathbb{P}^{X_3 | X_1}(x_3 | x_1) = \sum_{x_2 \in E_2} \mathbb{P}^{X_3 | X_1, X_2}(x_3 | x_1, x_2) \ \mathbb{P}^{X_2 | X_1}(x_2 | x_1)$$

On en conclut que

$$
\begin{aligned}
&\mathbb{E}(\mathbb{E}(f(X_3) \mid X_2, X_1) \mid X_1 = x_1) \\
&= \sum_{x_2} \sum_{x_3} \ f(x_3) \\
&\qquad\qquad \times \mathbb{P}^{X_3|X_2,X_1}(x_3 \mid x_2,\ x_1)\mathbb{P}^{X_2|X_1}(x_2|x_1) \\
&= \sum_{x_3} \ f(x_3) \\
&\qquad\qquad \times \left[\sum_{x_2} \mathbb{P}^{X_3|X_2,X_1}(x_3 \mid x_2,\ x_1)\mathbb{P}^{X_2|X_1}(x_2|x_1)\right] \\
&= \mathbb{E}(f(X_3) \mid X_1 = x_1)
\end{aligned}
$$

■