Les probabilités conditionnelles

Qu'est-ce qu'une probabilité conditionnelle ? Lorsque l'on s'intéresse à un évènement (et à sa probabilité), il arrive qu'en obtenant des informations supplémentaires on soit amené à faire évoluer la valeur que l'on attribue à sa probabilité. On définit la probabilité conditionnelle d'un évènement A sachant qu'un évènement B est réalisé par :

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

$\mathbf{P}(A|B)$ se lit « probabilité de A sachant que B est réalisé. » On le note également $\mathbf{P}_B(A)$ car il s'agit d'une nouvelle probabilité. On dispose alors de deux mesures de probabilité : $\mathbf{P}(A)$ est la probabilité d'un évènement A sans information supplémentaire, $\mathbf{P}_B(A)$ est la probabilité de A avec l'information supplémentaire que l'évènement B est réalisé.

Comment utiliser des probabilités conditionnelles ?

•**Avec la formule des probabilités totales.** Si on s'intéresse à la probabilité d'un évènement A et que l'on connait les probabilités conditionnelles $\mathbf{P}(A|B)$ et $\mathbf{P}(A|\overline{B})$, alors :

$$\mathbf{P}(A) = \mathbf{P}(A|B)\mathbf{P}(B) + \mathbf{P}(A|\overline{B})\mathbf{P}(\overline{B})$$

Plus généralement, si $B_1, \dots, B_n$ constituent un système complet d'évènements (c'est à dire que $B_1, \dots, B_n$ sont deux à deux disjoints et leur réunion est égale à Ω), alors on aura :

$$\mathbf{P}(A) = \mathbf{P}(A|B_1)\mathbf{P}(B_1) + \dots + \mathbf{P}(A|B_n)\mathbf{P}(B_n)$$

Exemple : on considère deux urnes U_1 et U_2, l'urne U_1 contient une boule blanche et deux boules noires, l'urne U_2 contient une boule noire et deux boules blanches. On choisit l'une des deux urnes aux hasard et on tire une boule, quelle est la probabilité d'obtenir une boule noire ? On note U_i l'évènement « on choisit l'urne i » et N l'évènement « on a tiré une boule noire. » Alors :

$$\mathbf{P}(N) = \mathbf{P}(N|U_1)\mathbf{P}(U_1) + \mathbf{P}(N|U_2)\mathbf{P}(U_2) = \frac{2}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2}$$

On a $\mathbf{P}(U_1) = \mathbf{P}(U_2) = 1/2$ (c'est le choix « au hasard »), $\mathbf{P}(N|U_1) = 2/3$ car il y a 3 boules dans U_1 dont 2 noires et $\mathbf{P}(N|U_2) = 1/3$ car il y a 3 boules dans U_2 dont 1 noire.

•**Avec la formule des probabilités composées.** Par définition d'une probabilité conditionnelle :

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}_A(B)$$

On peut démontrer des résultats analogues pour des intersections d'un nombre quelconque d'évènements :

$$\mathbf{P}(A\cap B\cap C)=\mathbf{P}(A)\mathbf{P}_A(B)\mathbf{P}_{A\cap B}(C)$$
$$\mathbf{P}(A\cap B\cap C\cap D)=\mathbf{P}(A)\mathbf{P}_A(B)\mathbf{P}_{A\cap B}(C)\mathbf{P}_{A\cap B\cap C}(D)$$

etc. Cette méthode est bien adaptée au cas où plusieurs évènements s'enchaînent dans le temps et chaque évènement influe sur ceux qui arrivent ensuite. **Exemple :** on dispose d'une urne contenant 4 boules noires et 2 boules blanches. On effectue des tirages successifs en remettant la boule tirée dans l'urne quand elle est blanche et sans la remettre quand elle est noire. Quelle est la probabilité d'obtenir une boule noire suivie de deux blanches ? Notons B_i (respectivement N_i) l'évènement « la i-ième boule tirée est blanche (respectivement noire) » alors :

$$\mathbf{P}(N_1\cap B_2\cap B_3)=\mathbf{P}(N_1)\mathbf{P}_{N_1}(B_2)\mathbf{P}_{N_1\cap B_2}(B_3)=\frac{4}{6}\times\frac{2}{5}\times\frac{2}{5}$$

$\mathbf{P}_{N_1}(B_2)=2/5$ car en supposant N_1 vrai, il reste 5 boules dans l'urne dont 2 blanches et $\mathbf{P}_{N_1\cap B_2}(B_3)=2/5$ car en supposant N_1 et B_2 vrais, il reste 5 boules dans l'urne dont 2 blanches.

Comment calculer des probabilités conditionnelles ?

•En appliquant la définition. Exemple : on lance un dé équilibré, quelle est la probabilité d'obtenir 6 sachant que le résultat est pair ? On note A l'évènement « on a obtenu un 6 W» (sans information supplémentaire, $\mathbf{P}(A)=1/6$) et B l'évènement « le résultat est pair. » On alors $\mathbf{P}(B)=1/2$, $A\cap B$ est l'évènement « le résultat est pair et égal à 6 » autrement dit $A\cap B=A$ et ainsi :

$$\mathbf{P}(A|B)=\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)}=\frac{1/6}{1/2}=\frac{1}{3}$$

La probabilité de A sachant B réalisé est donc $1/3$.

•En utilisant l'énoncé. Il est en effet fréquent que les informations données dans l'énoncé se traduisent sous forme de probabilités conditionnelles. **Exemple :** on dispose d'un test médical permettant de détecter la présence d'une certaine maladie. Aucun test n'étant parfait, il est possible qu'une personne malade ne soit pas détectée par le test. Il est également possible qu'une personne n'ayant pas contracté la maladie soit quand même testée positivement. Supposons par exemple que le test détecte la maladie pour 99% des personnes qui sont réellement malades et qu'il soit positif pour 0.04% des personnes qui ne sont pas infectées. Si on note T l'évènement « être testé positivement » et M l'évènement « avoir contracté la maladie » ces données numériques se traduisent par les probabilités conditionnelles $\mathbf{P}(T|M)=0.99$ et $\mathbf{P}(T|\overline{M})=0.04\cdot 10^{-2}$.

•En utilisant la formule de Bayes. Cette formule permet de passer d'une probabilité conditionnelle du type $\mathbf{P}(A|B)$ à une probabilité conditionnelle du type $\mathbf{P}(B|A)$. On peut la retrouver très simplement en utilisant la définition des probabilités conditionnelles (et la formule des probabilités totales) :

$$\mathbf{P}(B|A)=\frac{\mathbf{P}(B\cap A)}{\mathbf{P}(A)}=\frac{\mathbf{P}(A|B)\mathbf{P}(B)}{\mathbf{P}(A)}=\frac{\mathbf{P}(A|B)\mathbf{P}(B)}{\mathbf{P}(A|B)\mathbf{P}(B)+\mathbf{P}(A|\overline{B})\mathbf{P}(\overline{B})}$$

Exemple : on reprend les données de l'exemple précédent et on suppose de plus que 0.03% de la population est atteinte de la maladie (maladie rare). On suppose que le résultat du test est positif, quelle est la probabilité que la personne soit malade ? On s'intéresse ici à la probabilité conditionnelle $\mathbf{P}(M|T)$. On a $\mathbf{P}(M)=0.03\cdot 10^{-2}$, $\mathbf{P}(\overline{M})=1-0.03\cdot 10^{-2}$ et avec la formule précédente :

$$\mathbf{P}(M|T)=\frac{\mathbf{P}(T|M)\mathbf{P}(M)}{\mathbf{P}(T|M)\mathbf{P}(M)+\mathbf{P}(T|\overline{M})\mathbf{P}(\overline{M})}=\frac{0.99\times 0.03\cdot 10^{-2}}{0.99\times 0.03\cdot 10^{-2}+0.04\cdot 10^{-2}\times(1-0.03\cdot 10^{-2})}\simeq 0.42$$

Il y a 42% de chances que la personne soit réellement malade.