Calculs de probabilités conditionelles

Mathématiques Générales B
Université de Genève

Sylvain Sardy

20 mars 2008

1. Indépendance

Exemple : On lance deux pièces. Soit A l'évènement 'la première est Pile' et B l'évènement 'la deuxième est Pile'.

Les deux pièces sont équilibrées donc $\mathrm{P(A)} = \frac{1}{2}$ et $\mathrm{P(B)} = \frac{1}{2}$. De même $\mathrm{P(A \bigcap B)}$, la probabilité de deux Piles, est de 1 sur 4. On observe ici que

$$\mathrm{P(A \bigcap B) = P(A)P(B).}$$

Ce n'est pas un hasard : les deux évènements sont **indépendants**.

Définition : Deux évènements A et B sont indépendants si

$$\mathrm{P}(\mathrm{A} \bigcap \mathrm{B}) = \mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B}).$$

Exemple : On tire une carte parmi 52. Soit A l'évènement 'la carte est un As' et B l'évènement 'la carte est un Coeur'.

Clairement $\mathrm{P}(\mathrm{A}) = 4/52 = 1/13$ et $\mathrm{P}(\mathrm{B}) = 13/52 = 1/4$.

La probabilité que la carte soit un As de Coeur $(A \bigcap B)$ est de 1 sur 52.

On voit bien ici aussi que $\mathrm{P}(\mathrm{A} \bigcap \mathrm{B}) = \mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B})$.

Exemple de non-indépendance : On tire 2 cartes parmi 52. Soit A l'évènement 'la première est un Coeur' et B l'évènement 'la deuxième est un Coeur'.

La probabilité que la première carte est un Coeur est $\mathrm{P(A)} = 1/4$. De même pour la deuxième, donc $\mathrm{P(B)} = 1/4$.

Par contre la probabilité que les deux cartes soient des Coeur est $\mathrm{P(A \bigcap B)} = \mathrm{C}_{13,2}/\mathrm{C}_{52,2} = \frac{13 \cdot 12}{52 \cdot 51} < (\frac{1}{4})^2$.

Les deux évènements ne sont donc pas indépendants. On le savait puisque si la première est un Coeur (avec probabilité 13/52), alors la probabilité que la deuxième est un Coeur est plus faible (12/51).

On notera $\mathrm{P(B \mid A)}$ pour 'Probabilité de B sachant A'.

Exemple : On lance 6 dés. Quelle est la probabilité de l'évènement $A =$'On a exactement deux 4' ?

$\Omega = \{1,2,3,4,5,6\}^6$. C'est un ensemble fini.

Un élément de A possible est : $A_k = \boxed{\times}\,\boxed{4}\,\boxed{\times}\,\boxed{\times}\,\boxed{\times}\,\boxed{4}$, où $\times =$ tout sauf un 4.

Le nombre d'éléments de A est $|A| = C_{6,2} = \binom{6}{2}$.

$\mathrm{P}(\times) = 5/6$ et $\mathrm{P}(4) = 1/6$. Par indépendance, chaque A_k se réalise avec probabilité $\mathrm{P(A_k)} = (\frac{5}{6})^4(\frac{1}{6})^2$.

De plus les A_k sont disjoints.

Donc $\mathrm{P(A)} = \mathrm{P}(\bigcup_k \mathrm{A_k}) = \sum_k \mathrm{P(A_k)} = |\mathrm{A}| \cdot (\frac{5}{6})^4(\frac{1}{6})^2 = \binom{6}{2}\left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right)^2$.

Généralisation : la **distribution binomiale** B(n,p).

L'exemple précédent compte la probabilité de $k = 2$ succès ('obtenir un 4') après $n = 6$ expériences indépendantes et, à chaque expérience, la probabilité de succès est $p = 1/6$.

On répète une expérience aléatoire n fois de façon indépendante.
Soit la variable aléatoire $X_i = \begin{cases} 1 & \text{si expérience positive} \\ 0 & \text{sinon} \end{cases}$.
La probabilité d'un succès est $p = \mathrm{P}(\mathrm{X_i} = 1)$.

$$\mathrm{P}(\mathrm{k}\,\text{succès}) = \mathrm{P}(\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{X_i} = \mathrm{k}) = \left(\begin{array}{c} n \\ k \end{array} \right) \mathrm{p^k (1{-}p)^{n-k}} \quad \mathrm{pour} \quad \mathrm{k} \in \{0, 1, 2, \ldots, \mathrm{n}\}.$$

Dans l'exemple précédent : $\left(\begin{array}{c} 6 \\ 2 \end{array} \right) \left(\frac{1}{6}\right)^2 \left(1 - \frac{1}{6}\right)^{6-2}$.

Application : Un étudiant passe un test QCM à 4 possibilités. Il y a en tout 10 questions. Fait étrange, il n'est jamais venu en cours et va donc choisir au hasard. Quelle est la probabilité qu'il ait exactement 3 réponses justes ?

Il s'agit de 10 expériences indépendantes de probabilité de succès $p = 1/4$.

La probabilité d'obtenir exactement 3 réponses justes est donc :

$$\binom{10}{3}(1/4)^3(3/4)^7 = 0.2502823$$

Exemple : Une famille a 3 enfants. La probabilité d'avoir une fille ou un garçon est équitable. Soit $A =$'il y a au plus une fille' et $B =$'la famille a des enfants des deux sexes'. Ces deux évènements sont-ils indépendants ?

$$\mathrm{P(A)} = \mathrm{P}(0) + \mathrm{P}(1) = \mathrm{C}_{3,0}(1/2)^0(1/2)^3 + \mathrm{C}_{3,1}(1/2)^1(1/2)^2 = 1/2.$$

$$\mathrm{P(B)} = \frac{|\{\mathrm{FFG,FGF,GFF,FGG,GFG,GGF}\}|}{2^3} = 6/8 = 3/4.$$

$$\mathrm{P(A \bigcap B)} = \frac{|\{\mathrm{FGG,GFG,GGF}\}|}{2^3} = 3/8.$$

Ces deux évènements sont donc indépendants.

2. Probabilité conditionelle

Motivation

http://math.ucsd.edu/~crypto/Monty/monty.html

Intuition : On répète une expérience n fois est on note :

- le nombre de fois $n(A)$ où l'évènement A se réalise,
- le nombre de fois $n(A \bigcap B)$ où A et B se réalisent ensemble.

La probabilité de B sachant que A se réalise est donc proche de

$$\frac{n(A \bigcap B)}{n(A)} = \frac{n(A \bigcap B)/n}{n(A)/n} \approx \frac{P(B \bigcap A)}{P(A)}.$$

Définition : Soit A et B deux évènements. Si $\mathrm{P(A)} > 0$, on définit

$$\mathrm{P(B \mid A)} = \frac{\mathrm{P(B \bigcap A)}}{\mathrm{P(A)}},$$

et on lit : "Probabilité conditionelle de B sachant A."

Conséquence : Soit deux évènements indépendants A et B. On sait que $\mathrm{P}(\mathrm{A} \bigcap \mathrm{B}) = \mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B})$, donc

$$\mathrm{P}(\mathrm{B} \mid \mathrm{A}) = \frac{\mathrm{P}(\mathrm{B} \bigcap \mathrm{A})}{\mathrm{P}(\mathrm{A})} = \frac{\mathrm{P}(\mathrm{B})\mathrm{P}(\mathrm{A})}{\mathrm{P}(\mathrm{A})} = \mathrm{P}(\mathrm{B}).$$

C'est une conséquence attendue.

De plus si $B = A$, alors $\mathrm{P}(\mathrm{A} \mid \mathrm{A})$ devrait être égale à 1. En effet :

$$\mathrm{P}(\mathrm{A} \mid \mathrm{A}) = \frac{\mathrm{P}(\mathrm{A} \bigcap \mathrm{A})}{\mathrm{P}(\mathrm{A})} = \frac{\mathrm{P}(\mathrm{A})}{\mathrm{P}(\mathrm{A})} = 1.$$

Exemple : Une urne contient 6 boules rouges et 5 boules noires. On tire 2 boules sans remise. Quelle est la probabilité (conditionelle) que la deuxième soit noire sachant que la première est rouge ?

Appelons A=''la première est rouge'' et B=''la deuxième est noire''. La réponse cherchée est

$$P(B \mid A) = \frac{P(B \bigcap A)}{P(A)} = \frac{\frac{(6)(5)}{(11)(10)}}{\frac{(6)(10)}{(11)(10)}} = \frac{1}{2}.$$

Une autre façon de trouver le résultat est que, après que la première est rouge, il reste 5 rouges et 5 noires donc

$$P(B \mid A) = \frac{5}{10} = \frac{1}{2}.$$

Propriétés de multiplication :

1. $\mathrm{P(A \bigcap B) = P(A \mid B)P(B)}$
2. Soit $A_1, \ldots, A_k$ des évènements. Alors

$$\begin{aligned}\mathrm{P(A_1 \bigcap A_2 \bigcap \ldots \bigcap A_k)} &= \mathrm{P(A_k \mid A_{k-1} \bigcap \ldots \bigcap A_1)} \cdot \\ &\quad \mathrm{P(A_{k-1} \mid A_{k-2} \bigcap \ldots \bigcap A_1)} \cdot \\ &\quad \ldots \\ &\quad \mathrm{P(A_3 \mid A_2 \bigcap A_1) \cdot P(A_2 \mid A_1) \cdot P(A_1)}\end{aligned}$$

Exemple : Une urne contient 6 Rouges et 5 Noires. On tire 3 boules sans remise. Quelle est la probabilité qu'elles soient toutes rouges ?

Soit $A_i =$"la $i$ème est rouge". On cherche à calculer :

$$\mathrm{P}(\mathrm{A}_1 \bigcap \mathrm{A}_2 \bigcap \mathrm{A}_3) = \mathrm{P}(\mathrm{A}_3 \mid \mathrm{A}_1 \bigcap \mathrm{A}_2) \cdot \mathrm{P}(\mathrm{A}_2 \mid \mathrm{A}_1) \cdot \mathrm{P}(\mathrm{A}_1) = \frac{4}{9}\frac{5}{10}\frac{6}{11}.$$

Une autre approche est de considérer les arrangements possibles avec trois boules rouges en premier, soit :

$$\mathrm{P}(\mathrm{A}_1 \bigcap \mathrm{A}_2 \bigcap \mathrm{A}_3) = \frac{\mathrm{C}_{6,3}}{\mathrm{C}_{11,3}} = \frac{\frac{6!}{3!3!}}{\frac{11!}{3!8!}}.$$

Formule des probabilités totales.

Soit $\{A_1, \ldots, A_n\}$ une partition de Ω. Alors pour tout évènement A,

$$\mathrm{P}(\mathrm{A}) = \sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{P}(\mathrm{A} \mid \mathrm{A_k})\mathrm{P}(\mathrm{A_k}).$$

Démonstration :

$$\begin{aligned}
\mathrm{P}(\mathrm{A}) &= \mathrm{P}(\mathrm{A} \bigcap (\bigcup_{\mathrm{k}} \mathrm{A_k})) \\
&= \mathrm{P}(\bigcup_{\mathrm{k}} \mathrm{A} \bigcap \mathrm{A_k}) \\
&= \sum_{k} \mathrm{P}(\mathrm{A} \bigcap \mathrm{A_k}) \quad \square
\end{aligned}$$

Exemple : Urne avec n_R Rouge, n_N Noire et n_B Bleu. Soit $n = n_R + n_N + n_B$. Quelle est la probabilité de $A =$"la deuxième tirée sans remise est Rouge" ?

Soit A cet évènement et $A_x =$'la première est de couleur x'.
Notons que $\{A_R, A_N, A_B\}$ est une partition de l'univers, donc :

$$\begin{aligned}
\mathrm{P(A)} &= P(A \mid A_R)\mathrm{P(A_R)} + \mathrm{P(A \mid A_N)P(A_N)} + \mathrm{P(A \mid A_B)P(A_B)} \\
&= \frac{n_R - 1}{n-1}\frac{\mathbf{n_R}}{\mathbf{n}} + \frac{n_R}{n-1}\frac{n_N}{n} + \frac{n_R}{n-1}\frac{n_B}{n} \\
&= \frac{n_R}{n(n-1)}(n_R - 1 + n_N + n_B) \\
&= \frac{\mathbf{n_R}}{\mathbf{n}}.
\end{aligned}$$

Intéressant : C'est la même probabilité que la première boule tirée soit Rouge !

Donc que le tirage soit avec remise ou sans remise, la probabilité d'avoir une Rouge le premier et deuxième tirage est la même.

Problème de génétique. Supposons qu'un gène a 2 allèles : celui des yeux Marrons 'M' et celui des yeux bleus 'b'. 'M' est dominant et 'b' est récessif.

Supposons une population infinie au XVième siècle avec les probabilités suivantes d'avoir M ou b :

Gènes	MM	Mb/bM	bb
Probabilités au temps 0	α_0	β_0	γ_0

Naturellement $\alpha_0 + \beta_0 + \gamma_0 = 1$.

Comment cette population va-t-elle évoluer ?
Quelle est la proportion au XXIième siècle ?

On suppose que les couples se forment au hasard et les allèles sont choisis aléatoirement.

Le premier allèle d'un enfant sera M avec probabilité $p_1 = \alpha_0 + \beta_0/2$.

Le premier allèle d'un enfant sera b avec probabilité $1 - p_1 = \beta_0/2 + \gamma_0$.

Le second allèle est indépendant et de même distribution (probabilité).

La proportion de la première génération est donc

Gènes	MM	Mb/bM	bb
Probabilités au temps 0	α_0	β_0	γ_0
Probabilités au temps 1	α_1	β_1	γ_1

avec $\alpha_1 = p_1^2 \qquad \beta_1 = 2p_1(1 - p_1) \qquad \gamma_1 = (1 - p_1)^2 \quad .$

Un fait remarquable arrive à la deuxième génération :

$$p_2 = \alpha_1 + \beta_1/2 = p_1^2 + 2p_1(1-p_1)/2 =?$$

Par conséquent :

Gènes	MM	Mb/bM	bb
Probabilités au temps 0	α_0	β_0	γ_0
Probabilités au temps 1	α_1	β_1	γ_1
Probabilités au temps 2	α_1	β_1	γ_1
Probabilités au temps ...	α_1	β_1	γ_1
A l'équilibre :	p^2	$2p(1-p)$	$(1-p)^2$

Ce résultat s'appelle le Théorème de Hardy–Weinberg.

3. Formule de Bayes

Soit $\{A_1, \ldots, A_n\}$ une partition de Ω et B un évènement

$$\mathrm{P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_i P(B \mid A_i)P(A_i)}}.$$

Mieux vaut se souvenir de la démonstration :

$$\begin{aligned} \mathrm{P(A_i \mid B)} &= \frac{\mathrm{P(A_i \bigcap B)}}{\mathrm{P(B)}} \\ &= \ldots \end{aligned}$$

$\square$

Exemple : Test pour détecter une maladie.
Sur la boite d'un test médical, il est indiqué que le test est sûr à 95% quand la personne est malade, et que dans 1% des cas le test déclare un 'faux positif'.
De plus, 1 personne sur 100'000 est infectée dans la population.
On vous administre le test qui se révèle positif. Quelle est la probabilité que vous soyez effectivement malade ?

Soit $A =$'être malade' et $B =$'être positif'. On cherche

$$\begin{aligned} \mathrm{P}(\mathrm{A} \mid \mathrm{B}) &= \frac{\mathrm{P}(\mathrm{B} \mid \mathrm{A})\mathrm{P}(\mathrm{A})}{\mathrm{P}(\mathrm{B} \mid \mathrm{A})\mathrm{P}(\mathrm{A}) + \mathrm{P}(\mathrm{B} \mid \mathrm{A}^c)\mathrm{P}(\mathrm{A}^c)} \\ &= \frac{(0.95)(1/100000)}{(0.95)(1/100000) + (0.01)(1 - 1/100000)} \\ &= 0.09\% \end{aligned}$$

Exemple : Problème de Monty-Hall.
Dans un jeu TV, le joueur peut choisir 3 portes. Derrière l'une d'elle se trouve une voiture. Derrière les deux autres se trouve une chèvre.
Hypotèse importante : Monty Hall sait ce qui se trouve derrière chaque porte.

`http://math.ucsd.edu/~crypto/Monty/monty.html`

Soit $V_i =$'La voiture est derrière Porte i'.
Sans perte de généralité on peut supposer que le joueur choisit Porte 1.
Soit $M_{1j} =$'Monty Hall choisit Porte j après que le joueur a choisit Porte 1'.

$$\mathrm{P}(\mathrm{M_{1j}} \mid \mathrm{V_i}) = \begin{cases} 1/2, & i = 1, j = 2, 3 \\ 1, & i = 2, j = 3 \quad \text{ou} \quad \mathrm{i = 3, j = 2} \\ 0, & i = 2, j = 2 \quad \text{ou} \quad \mathrm{i = 3, j = 3} \end{cases} .$$

Sans perte de généralité (en renumérotant les Portes si nécessaire), on suppose que Monty Hall a choisi Porte 2.

On s'intéresse à

$$\begin{aligned}
\mathrm{P}(\mathrm{V}_1 \mid \mathrm{M}_{12}) &= \frac{\mathrm{P}(\mathrm{M}_{12} \mid \mathrm{V}_1)\mathrm{P}(\mathrm{V}_1)}{\mathrm{P}(\mathrm{M}_{12} \mid \mathrm{V}_1)\mathrm{P}(\mathrm{V}_1) + \mathrm{P}(\mathrm{M}_{12} \mid \mathrm{V}_2)\mathrm{P}(\mathrm{V}_2) + \mathrm{P}(\mathrm{M}_{12} \mid \mathrm{V}_3)\mathrm{P}(\mathrm{V}_3)} \\
&= \frac{(1/2)(1/3)}{(1/2)(1/3) + 0 + (1)(1/3)} \\
&= \frac{1/6}{1/2} = 1/3.
\end{aligned}$$

En outre :

$$\begin{aligned} \mathrm{P}(\Omega \mid \mathrm{M}_{12}) &= 1 \\ \text{ssi} \quad \mathrm{P}(\mathrm{V}_1 \mid \mathrm{M}_{12}) + \mathrm{P}(\mathrm{V}_2 \mid \mathrm{M}_{12}) + \mathrm{P}(\mathrm{V}_3 \mid \mathrm{M}_{12}) &= 1 \\ \text{ssi} \quad 1/3 + 0 + \mathrm{P}(\mathrm{V}_3 \mid \mathrm{M}_{12}) &= 1 \end{aligned}$$

Donc $\mathrm{P}(\mathrm{V}_3 \mid \mathrm{M}_{12}) = 2/3$.

Conclusion...