Probabilité conditionnelle ; indépendance de deux événements (on se limitera au cas où l'ensemble d'épreuves est fini). Application à des calculs de probabilités.

Chantal Menini

13 mai 2009

1 Introduction.

Nous commençerons par un exemple.
En 2005-2006 les effectifs (exprimés en milliers) en classes préparatoires se répartissaient de la façon suivante (*Source Insee : Regards sur le parité www.educnet.education.fr/insee/*).

	Filles	Garçons
Scientifique	13.8	33.5
Economique	8.8	7.2
Littéraire	8.5	2.9

On peut alors se poser les questions suivantes : si l'on prend au hasard un élève de classe préparatoire quelle est la probabilité que ce soit une fille ? Quelle est la probabilité que ce soit une fille et qu'elle soit en préparation scientifique ? Sachant que c'est une fille qu'elle est la probabilité qu'elle soit en préparation scientifique ?
Pour répondre à ces questions il faut compléter le tableau en calculant les effectifs totaux suivants

	Filles	Garçons	Total
Scientifique	13.8	33.5	
Economique	8.8	7.2	
Littéraire	8.5	2.9	
Total	31.1	43.6	74.7

Pour le calcul des deux premières probabilités on modélise en prenant pour univers l'ensemble des élèves de classes préparatoires que l'on munit de la probabilité uniforme. La probabilité qu'un élève pris au hasard soit une fille (nous noterons F cet événement) est alors la fréquence des filles soit $P(F) = \frac{31.1}{74.7} \simeq 0.42$ et la probabilité qu'un élève pris au hasard soit une fille et en préparation scientifique (nous noterons $F \cap S$ cet événement) est de même $P(F \cap S) = \frac{13.8}{74.7} \simeq 0.18$.
Pour le troisième calcul la modélisation se fera en prenant pour univers l'ensemble des filles de classe préparatoire muni de la probabilité uniforme et la probalité qu'une fille prise au hasard soit en préparation scientifique est alors $p = \frac{13.8}{31.1} \simeq 0.44$.
On remarquera alors que $p = \frac{P(F \cap S)}{P(F)}$, on appellera cette probabilité *probabilité conditionnée par F* ou encore *probabilité sachant F.*

2 Probabilité conditionnelle

Dans tout ce qui suit nous travaillerons sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$.

Définition 2.1 *Soient A un événement tel que $P(A) > 0$ et B un autre événement. On appelle* ***probabilité de B conditionnée par A*** *ou* ***probabilité de B sachant A*** *le réel, noté $P_A(B)$, défini par*

$$P_A(B) = \frac{P(B \cap A)}{P(A)}.$$

Remarque 2.2 $P_A(A) = 1$

Proposition 2.3 *P_A est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$.*

Preuve. P_A est bien une application définie sur $\mathcal{P}(\Omega)$ et à valeurs dans $\mathbb{R}_+$.
$P_A(\Omega) = \frac{P(\Omega\cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$ et si B et C sont deux événements incompatibles, alors $B\cap A$ et $C\cap A$ sont incompatibles et

$$P_A(B\cup C) = \frac{P((B\cup C)\cap A)}{P(A)} = \frac{P((B\cap A)\cup(C\cap A))}{P(A)} = \frac{P(B\cap A)}{P(A)} + \frac{P(C\cap A)}{P(A)} = P_A(B) + P_A(C).$$

□

Remarque 2.4 *Une conséquence importante est que toutes les propriétés connues des probabilités sont valables pour une probabilité conditionnée selon un événement.*

Proposition 2.5 *Soient A et B deux événements tels que $P(A\cap B) > 0$ alors $(P_A)_B = P_{(A\cap B)}$.*

Preuve. Commençons par remarquer que P_A est bien défini puisque $P(A) \geq P(A\cap B) > 0$ et $P_A(B) > 0$.
Pour tout événement E, $(P_A)_B(E) = \frac{P_A(B\cap E)}{P_A(B)} = \frac{P(A\cap(B\cap E))}{P(A)}\frac{P(A)}{P(A\cap B)} = P_{(A\cap B)}(E)$. □

3 Indépendance.

Définition 3.1 *Soient A et B deux événements d'un espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ avec $P(A) > 0$. Alors B est **indépendant** de A si $P_A(B) = P(B)$.*

Remarque 3.2 *Avec la définition de la probabilité conditionnée par A on voit que $P_A(B) = P(B)$ équivaut à $P(B\cap A) = P(B)P(A)$ ce qui nous conduit à la définition plus générale suivante.*

Définition 3.3 *Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé, alors*

1. *Deux événements A et B sont dits **indépendants** si $P(A\cap B) = P(A)P(B)$.*
2. *N événements $A_1, \ldots, A_N$ sont dits **indépendants dans leur ensemble** si pour tout k compris entre 2 et N, tous indices $1 \leq i_1 < \cdots < i_k \leq N$ on a $P(A_{i_1}\cap\ldots\cap A_{i_k}) = P(A_{i_1})\cdots P(A_{i_k})$.*

Remarque 3.4
- *Indépendance et incompatibilité sont deux choses différentes, en effet si deux événements A et B de probabilité non nulle sont incompatibles alors $P(A\cap B) = 0 \neq P(A)P(B)$.*
- *L'indépendance dans leur ensemble de N événements implique l'indépendance deux à deux (faire $k = 2$ dans la définition) mais la réciproque est fausse. Pour s'en convaincre on peut faire l'exercice suivant.*
  Exercice 1. *On jette deux dés de façon indépendante et équiprobable. Soient A l'événement "le chiffre du premier dé est impair", B l'événement "le chiffre du deuxième dé est pair" et C l'événement "les deux dés ont la même parité". Montrer que A, B et C sont indépendants deux à deux mais ne sont pas indépendants dans leur ensemble.*

Proposition 3.5 *Soient A et B deux événements indépendants alors A et $\overline{B}$ sont indépendants.*

Cette propriété se généralise à un nombre fini d'événements.
Preuve. $P(A\cap\overline{B}) = P(A) - P(A\cap B) = P(A) - P(A)P(B) = P(A)(1-P(B)) = P(A)P(\overline{B})$. □

4 Deux résultats de décomposition.

4.1 Probabilités conditionnelles composées.

Proposition 4.1 *Soient $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé, A et B deux événements avec $P(A) > 0$ alors $P(A\cap B) = P(A)P_A(B)$.*

Preuve. Ceci résulte directement de la définition de la probabilité conditionnée par A. □

Proposition 4.2 *Soient $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé, $(A_i)_{1\leq i\leq n}$ une suite finie d'événements tels que $P\left(\bigcap_{1\leq i\leq n-1} A_i\right) > 0$; on a*

$$P\left(\bigcap_{1\leq i\leq n} A_i\right) = P(A_1)P_{A_1}(A_2)P_{(A_1\cap A_2)}(A_3)\cdots P_{(A_1\cap\ldots\cap A_{n-1})}(A_n).$$

Preuve. Commençons par remarquer que toutes les probabilités conditionnées introduites sont bien définies puisque pour tout entier k compris entre 1 et $n-1$, $P\left(\bigcap_{1\leq i\leq k} A_i\right) \geq P\left(\bigcap_{1\leq i\leq n-1} A_i\right) > 0$.
La preuve se fait par récurrence sur n. Soit (H_n) l'hypothèse de récurrence : "$(A_i)_{1\leq i\leq n}$ une suite finie d'événements tels que $P\left(\bigcap_{1\leq i\leq n-1} A_i\right) > 0$ alors $P\left(\bigcap_{1\leq i\leq n} A_i\right) = P(A_1)P_{A_1}(A_2)P_{(A_1\cap A_2)}(A_3)\cdots P_{(A_1\cap\ldots\cap A_{n-1})}(A_n)$".
(H_2) est vérifiée (c'est la prosition 4.1). Montrons que (H_n) implique (H_{n+1}) pour $n\geq 2$.
$P\left(\bigcap_{1\leq i\leq n} A_i\right) > 0$ et avec la proposition 4.1

$$P\left(\bigcap_{1\leq i\leq n+1} A_i\right) = P(A_1\cap\ldots\cap A_n)P_{(A_1\cap\ldots\cap A_n)}(A_{n+1}).$$

Puis en utilisant (H_n) (cf. $P(A_1\cap\ldots\cap A_{n-1}) \geq P(A_1\cap\ldots\cap A_n) > 0$) on a (H_{n+1}).
(H_2) est vérifiée, (H_n) implique (H_{n+1}) pour tout $n\geq 2$ donc (H_n) est vérifié pour tout $n\geq 2$. □

4.2 Formule des probabilités totales.

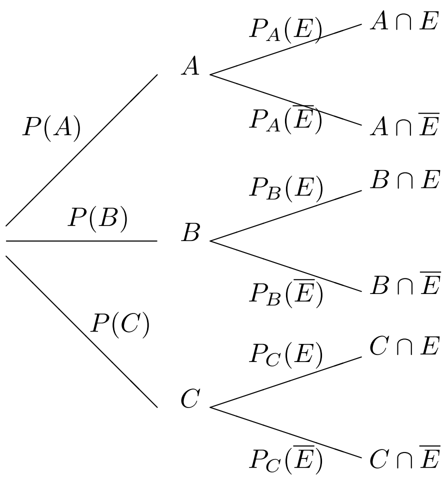
Définition 4.3 *Une famille finie d'événements $(A_i)_{1\leq i\leq n}$ deux à deux incompatibles et tels que $\bigcup_{1\leq i\leq n} A_i = \Omega$ est appelée **système complet d'événements**.*

Théorème 4.4 *Soient $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé et $(A_i)_{1\leq i\leq n}$ un système complet d'événements ayant tous une probabilité strictement positive, alors pour tout événement A*

$$P(A) = \sum_{i=1}^{n} P(A_i)P_{A_i}(A).$$

Preuve. $P(A) = P(A\cap(\bigcup_{1\leq i\leq n} A_i)) = \sum_{i=1}^{n} P(A\cap A_i)$ et le résultat final avec la proposition 4.1. □

4.3 Arbre de probabilité.



Les règles de calculs dans un arbre de probabilités sont les suivantes :
- Pour calculer la probabilité d'un événement figurant au bout d'une branche, on fait le produit des probabilités figurant sur les branches conduisant à cet événement (on parlera de la probabilité du chemin). Cette règle n'est autre que la propriété des probabilités composées. Par exemple ici $P(A\cap E) = P(A)P_A(E)$.
- La probabilité d'un événement est la somme des probabilités des chemins conduisant à cet événement. Cette règle est la formule des probabilités totales. Par exemple ici $P(E) = P(A)P_A(E) + P(B)P_B(E) + P(C)P_C(E)$.
- La somme des probabilités des branches ayant même origine vaut 1. Ceci provient du fait que sur les branches figurent des probabilités et que l'on a un système complet d'événements. Par exemple ici $P_A(E) + P_A(\overline{E}) = 1$.

Exercice. 2 Une expérience est conduite pour étudier la mémoire des rats. Un rat est mis devant trois couloirs. Au bout de l'un d'eux se trouve la nourriture qu'il aime, au bout des deux autres, il reçoit une décharge électrique. Cette expérience élémentaire est répétée au plus 50 fois et s'arrête dès que le rat a trouvé le bon couloir. Sous chacune des hypothèses suivantes :

(H_1) le rat n'a aucun souvenir des expériences antérieures,
(H_2) le rat se souvient de l'expérience immédiatement précédente,
(H_3) le rat se souvient des deux expériences précédentes,
avec qu'elle probabilité la première tentative réussie est-elle la première ? la deuxième ? la troisième ? la k-ième ? Avec qu'elle probabilité aura t-il échoué à chacune des 50 tentatives ?

5 Probabilité des causes.

Très souvent lorsque l'on connait $P_A(B)$, l'événement A est considéré comme la *cause* et l'événement B comme la *conséquence* ; par exemple si on considère un test de dépistage d'une maladie A sera l'événement "la personne est malade" et B "le test est positif". Mais souvent on a besoin de connaitre la probabilité de la cause sachant la conséquence ; si on reprend l'exemple du test de dépistage, sachant que le test est positif on veut connaitre la probabilité que la personne soit malade. Nous utiliserons pour cela la formule de Bayes.

Lemme 5.1 *Soient $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé, A_1 et A_2 deux événements de probabilité non nulle alors*

$$P_{A_1}(A_2) = P_{A_2}(A_1)\frac{P(A_2)}{P(A_1)}.$$

Preuve. D'après la proposition 4.1 $P(A_1)P_{A_1}(A_2) = P(A_1 \cap A_2) = P(A_2)P_{A_2}(A_1)$. □

Théorème 5.2 ***(Formule de Bayes)*** *Soient $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé et $(A_i)_{1\leq i\leq n}$ un système complet d'événements tels que pour tout i $P(A_i) > 0$. Alors pour tout événement A de probabilité non nulle et pour tout i*

$$P_A(A_i) = \frac{P(A_i)P_{A_i}(A)}{\sum_{j=1}^{n} P(A_j)P_{A_j}(A)}.$$

Preuve. Avec le lemme ci-dessus $P_A(A_i) = \frac{P(A_i)P_{A_i}(A)}{P(A)}$ et on termine avec la formule des probabilités totales. □
Exercice 3. On considère un test servant à dépister une maladie. Soient les événements M ="l'individu est malade" et P_0 ="le test est positif". On sait expérimentalement que $P_M(\overline{P_0}) = 10^{-3}$ et que $P_{\overline{M}}(P_0) = 2 \times 10^{-3}$ appliqué à une maladie qui touche un individu sur 10000. Calculer la probabilité que l'individu soit malade lorsque le test est positif.

6 Modélisation d'un sytème évolutif.

Un exercice type que l'on trouve dans tout chapitre sur les probabiltés conditionnelles est le suivant
Exercice 4. On dispose de deux urnes U_1 et U_2, dans l'urne U_1 il y a 7 boules noires et 3 blanches et dans l'urne U_2 il y a 9 boules noires et 11 blanches, les boules sont indiscernables au toucher. On choisit une urne au hasard puis on tire une boule, qu'elle est la probabilité de tirer une boule noire ?
Exercice que l'on résolvera de la façon suivante. On note U_i l'événement "choix de l'urne U_i" et N (resp. B) l'événement "choix de la boule noire (resp. blanche)".
Le choix des urnes est équiprobable donc $P(U_1) = P(U_2) = 1/2$. Dans chaque urne les boules sont indiscernables au toucher donc il y a encore équiprobabilité, d'où $P_{U_1}(N) = \frac{7}{7+3}$ et $P_{U_2}(N) = \frac{9}{9+11}$. Avec la formule des probabilités totales $P(N) = \frac{1}{2}(\frac{7}{10} + \frac{9}{20}) = 0.575$.
On constate que l'information donnée par l'enoncé induit la connaissance de P_{U_i} et non celle du modèle probabiliste $(\Omega, \mathcal{P}(\Omega), P)$ pour lequel P_{U_i} est la probabilité conditionnée par l'événement U_i. On doit donc se poser la question de l'existence d'un tel modèle.

Proposition 6.1 *Soient Ω_i, $1 \leq i \leq n$ des ensembles finis ; on suppose connu*
- *la probabilité sur Ω_1 notée p_1,*
- *pour tout $2 \leq k \leq n$ et tout $(\omega_1, \ldots, \omega_{k-1}) \in \Omega_1 \times \cdots \times \Omega_{k-1}$, la probabilité sur Ω_k notée $p_k^{(\omega_1,\ldots,\omega_{k-1})}$.*

Alors il existe un espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ tel que

$$\begin{aligned} P(\{\omega_1\} \times \Omega_2 \times \cdots \Omega_n) &= p_1(\{\omega_1\}) \\ P_{(\{\omega_1\}\times\cdots\times\{\omega_{k-1}\}\times\Omega_k\times\cdots\times\Omega_n)}(\{\omega_1\} \times \cdots \times \{\omega_k\} \times \Omega_{k+1} \times \cdots \times \Omega_n) &= p_k^{(\omega_1,\ldots,\omega_{k-1})}(\{\omega_k\}) \end{aligned} \qquad (1)$$

On peut prendre $\Omega = \Omega_1 \times \cdots \times \Omega_n$ et P défini par $P(\{\omega_1\} \times \cdots \times \{\omega_n\}) = p_1(\{\omega_1\})p_2^{(\omega_1)}(\{\omega_2\}) \cdots p_n^{(\omega_1,\ldots,\omega_{n-1})}(\{\omega_n\})$.

Remarque 6.2 *Lorsque pour tout k $p_k^{(\omega_1,\ldots,\omega_{k-1})}$ ne dépend pas de $(\omega_1,\ldots,\omega_{k-1})$ mais uniquement de k, alors, la probabilité P obtenue est la probabilité produit et les événements $\Omega_1\times\cdots\times\Omega_{k-1}\times\{\omega_k\}\times\Omega_{k+1}\cdots\times\Omega_n$ et $\Omega_1\times\cdots\times\Omega_{j-1}\times\{\omega_j\}\times\Omega_{j+1}\cdots\times\Omega_n$ sont indépendants pour tout $k\neq j$.*

Preuve.
P est donné sur les événements élémentaires de Ω et est à valeurs positives. Rappelons que l'application définie sur $\mathcal{P}(\Omega)$ par $P(E)=\sum_{\omega\in E}P(\{\omega\})$ sera bien une probabilité si et seulement si $\sum_{\omega\in\Omega}P(\{\omega\})=1$ (cf. leçon 04). En utilisant que pour tout k, $p_k^{(\omega_1,\ldots,\omega_{k-1})}$ est une probabilité nous avons que

$$\sum_{\omega\in\Omega}P(\{\omega\})=\sum_{\omega_1\in\Omega_1}p_1(\{\omega_1\})\left(\sum_{\omega_2\in\Omega_2}p_2^{(\omega_1)}(\{\omega_2\})\ldots\left(\sum_{\omega_n\in\Omega_n}p_n^{(\omega_1,\ldots,\omega_{n-1})}(\{\omega_n\})\right)\ldots\right)=1.$$

En outre la relation (1) se vérifie aisément. □

Si l'on revient à notre exemple avec les urnes, $\Omega_1=\{1,2\}$ et $p_1(\{1\})=p_1(\{2\})=1/2$ (modélisation du choix de l'urne) ; $\Omega_2=\{n,b\}$ et $p_2^{(1)}(\{n\})=\frac{7}{10}$, $p_2^{(2)}(\{n\})=\frac{9}{20}$. Alors $\Omega=\{(1,n),(2,n),(1,b),(2,b)\}$ et $P(\{(1,n)\})=\frac{1}{2}\frac{7}{10}$,etc...En utilisant que $U_i=\{(i,n),(i,b)\}$ pour $i=1$ ou 2 et que $N=\{(1,n),(2,n)\}$ on a bien que $P(U_i)=1/2$ et $P_{U_1}(N)=\frac{7}{10}$ etc...

Remarque 6.3 *On notera que dans les exercices 2 et 3 aussi l'enoncé nous donne des indications directes sur les probabilités conditionnés par la ou les étapes précédentes.*

7 Quelques classiques.

7.1 Ruine du joueur.

Un joueur a a euros en poche et il décide de jouer à un jeu pour lequel il mise à chaque partie k euros, s'il perd la partie il perd sa mise, s'il gagne la partie il remporte le double de sa mise. Les parties sont indépendantes et il a la probabilité p de gagner une partie. Il s'est fixé comme objectif de gagner $2a$ euros, il arrête de jouer et repart avec cette somme dès qu'il a atteint cet objectif ou bien il repart ruiné dès qu'il n'a plus d'argent disponible. Le but est de calculer la probabilité $P(n)$ que le joueur atteigne son objectif de gagner $2a$ euros lorsqu'il a n euros en poche.
Avec la formule des probabilités totales (en admettant que la probabilité qu'il atteigne son objectif lorsqu'il a n euros en poche sachant qu'il gagne la partie suivante est égal à la probabilité qu'il atteigne son objectif lorsqu'il a $n+k$ euros en poche)

$$P(n)=pP(n+k)+(1-p)P(n-k).$$

En prenant pour simplifier a un multiple de k, nous obtenons que la suite de terme général $u_l=P(lk)$ satisfait la relation de récurrence

$$u_{l+2}-\frac{1}{p}u_{l+1}+\frac{1-p}{p}u_l=0$$

avec pour conditions “aux bords” $u_0=P(0)=0$ et $u_{\frac{2a}{k}}=P(2a)=1$.
Nous avons une relation de récurrence d'ordre 2 et pour déterminer le terme général de la suite il faut chercher les racines de l'équation du second degré

$$x^2-\frac{1}{p}x+\frac{1-p}{p}=0$$

qui sont 1 et $\frac{1-p}{p}$.
Elles sont distinctes pour $p\neq 1/2$ (pour tous les jeux de casino p est strictement inférieur à 1/2) et alors $u_l=\alpha+\beta\left(\frac{1-p}{p}\right)^l$, α et β étant déterminés par les conditions $u_0=0$ et $u_{\frac{2a}{k}}=1$. Ce qui donne

$$P(a)=u_{a/k}=\frac{1-(\frac{1-p}{p})^{a/k}}{1-(\frac{1-p}{p})^{2a/k}}=\frac{1}{1+(\frac{1-p}{p})^{a/k}},$$

probabilité qui est maximale lorsque l'on mise le tout pour le tout c'est à dire $k=a$ et vaut alors p (ce qui n'est guère surprenant).
Si on refait les calculs dans le cas d'un jeu équitable ($p=1/2$), l'équation $x^2-\frac{1}{p}x+\frac{1-p}{p}=0$ admet alors 1 pour racine double et $u_l=\alpha+l\beta$, les conditions “aux bords” nous permettent d'en déduire que $u_l=l\frac{k}{2a}$ puis que $P(a)=u_{a/k}=1/2$ quelque soit le montant k de la mise que l'on a décidé de mettre à chaque partie.

7.2 Difficulté d'interprétation des résultats.

Un test médical a pour but de comparer deux méthodes de traitements de calculs (rénaux), la méthode par chirurgie notée C et la méthode par ultra-sons notée U. 700 patients ont été traités, 350 par chirurgie et 350 par ultra-sons. 273 traités par la méthode C ont guéri et 289 par la méthode U. On en déduit donc, en prenant pour univers l'ensemble des patients traités muni de la probabilité uniforme et en notant R l'événement une personne prise au hasard parmi les patients traités a été guérie, que $P_C(R) = \frac{273}{350} \simeq 0.78$ et que $P_U(R) = \frac{289}{350} \simeq 0.83$. Nous sommes donc tenté d'en conclure que la méthode par ultra-sons est plus performante.
Une analyse plus fine tenant compte de la taille des calculs traités donne les valeurs suivantes :

	Méthode C	Méthode U
Gros calculs C_g	191 réussites/71 échecs	60 réussites/26 échecs
Petits calculs C_p	82 réussites/6 échecs	229 réussites/35 échecs

Nous avons alors les probabilités conditionnelles $P_{C\cap C_g}(R) = \frac{191}{262} \simeq 0.73$ et $P_{U\cap C_g}(R) = \frac{60}{86} \simeq 0.7$, la méthode par ultra-sons est moins performante sur les gros calculs.
Ainsi que les probabilités conditionnelles $P_{C\cap C_p}(R) = \frac{82}{88} \simeq 0.93$ et $P_{U\cap C_p}(R) = \frac{229}{264} \simeq 0.87$, la méthode par ultra-sons est aussi moins performante sur les petits calculs...

8 Commentaires.

- Compte-tenu du titre il n'est pas indispensable (et même peut-être à éviter) de parler de l'indépendance de N événements dans leur ensemble puisqu'il est indiqué *indépendance de deux événements*. Il est bon tout de même d'avoir un minimum de connaissance dessus.
- La partie modélisation d'un sytème évolutif est techniquement complexe, il est bon de l'avoir comprise au moins dans le cas de l'exemple des urnes puisque vous l'aurez constaté la plupart des énoncés nous donnent des probabilités conditionnées sans se préocuper du modèle probabiliste sous-jacent.

9 Bibiographie.

Issac R., *Une initiation aux probabilités*, Vuibert.
Ouvrard J-Y., *Probabilités 1 - Capes Agrégation*, Cassini.
Ramis J.P., *Mathématiques tout-en-un L2*, Dunod.