Eléments du Calcul des Probabilités
Chapitres 1 et 2

Louis Wehenkel

Département d'Electricité, Electronique et Informatique - Université de Liège
B24/II.93 - L.Wehenkel@ulg.ac.be

MATH0062-1 : 2BacIng - 7/2/2017

Find slides: *http://montefiore.ulg.ac.be/∼lwh/Probas/*

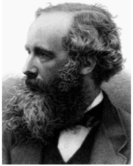
Dixit ...

"La théorie des probabilités n'est rien d'autre que le bon sens réduit sous forme de calcul."

Pierre Simon Laplace, 1749 - 1827

"The true logic of this world lies in the calculus of probabilities."

James Clerk Maxwell, 1831 - 1879



Chapitre 1. Introduction

D'abord, comprendre et apprendre à résoudre des problèmes en présence d'incertitudes et/ou de phénomènes aléatoires

Ensuite, modéliser, analyser, concevoir, et optimiser des systèmes techniques complexes

Chapitre 1. Introduction

D'abord, comprendre et apprendre à résoudre des problèmes en présence d'incertitudes et/ou de phénomènes aléatoires

Ensuite, modéliser, analyser, concevoir, et optimiser des systèmes techniques complexes

1.1 Motivation

1.2 Probabilités vs statistique

1.3 Trois problèmes typiques de raisonnement sous incertitude

1.4 Organisation du cours (Théorie/Répétitions/Travaux ; Evaluation)

NB.

- Calcul des probabilités ⇒ discipline mathématique
  - Préférer la rigueur à l'intuition
  - Préférer la précision à l'approximation
- Prérequis : analyse, géométrie, algèbre, analyse numérique, informatique
- Débouchés : plein d'applications et de méthodes de résolution de problèmes

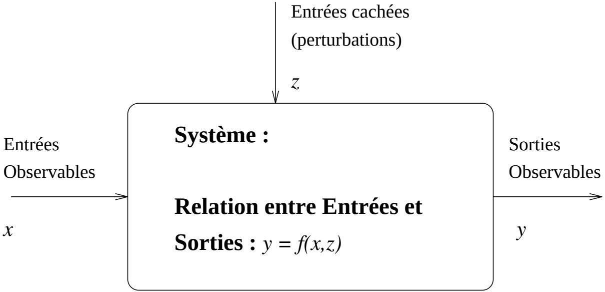
1.1 Motivation

- Du point de vue applicatif
  1. Nécessité croissante de traiter des systèmes très complexes
  2. Difficulté intrinsèque de prédire certains facteurs dimensionnants et certains phénomènes physiques
  3. Impact très important des technologies de l'information sur la manière d'aborder ces problèmes
- Du point de vue pédagogique
  1. Bachelier:
     - Calcul des probabilités : bases logiques et mathématiques du raisonnement en présence d'incertitudes
     - Statistique : méthodes pour tirer des conclusions fiables à partir de données d'observation
     - Processus aléatoires : modélisation opérationnelle des phénomènes spatio-temporels
  2. Master : applications dans les différentes disciplines

Exemples de domaines concernés

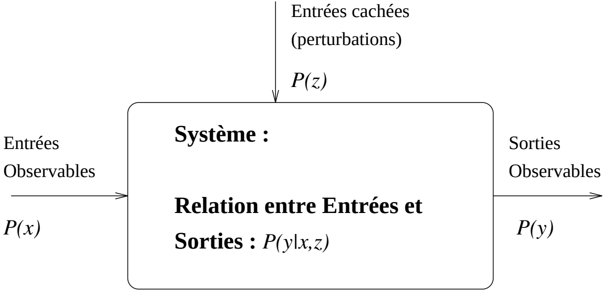
- La *gestion des risques*
- Les *arbitrages économiques*
- La *théorie des systèmes* et le *traitement du signal*
- L'*informatique*
- L'ingénierie des *procédés chimiques*
- La *production mécanique*
- L'*aéronautique* et le domaine des *transports*
- La *géologie* et la *construction*
- Les *télécommunications*
- L'*ingénierie biomédicale*
- etc. etc.

Représentation d'un système (vision déterministe)



Disons que le système est une voiture : les entrées observables x seraient alors les actions prises par le conducteur (position du volant, de l'accélérateur, du frein, etc.), la sortie y serait la trajectoire du véhicule, et les entrées "cachées" z (i.e. les perturbations non observables) seraient par exemple les forces exercées par la route sur les pneus, et les mouvements des passagers.
La relation déterministe entrées-sorties $y = f(x, z)$ permet de calculer les sorties si on connait les entrées observables et cachées.

Représentation d'un système (vision stochastique)



La loi *$P(y|x,z)$ tient compte de la connaissance imparfaite du système* (usure des pneus, température du moteur, etc.), la loi $P(z)$ exploite nos connaissances a priori (conditions hivernales, nombre de passagers) quant aux entrées cachées, et la loi $P(x)$ permet de modéliser les erreurs de mesure sur les entrées observables.
Les méthodes probabilistes permettent alors de calculer une loi de probabilité $P(y)$ sur l'ensemble des trajectoires possibles de la voiture étant données ces informations.

1.2 Probabilités vs statistique

- Probabilités :
  - *le modèle est complètement spécifié*
  - le but essentiel est d'*exploiter le modèle* pour prendre des décisions
  - nécessite rigueur et cohérence (OK, pour un ordinateur...)
  - *Raisonnement déductif*
- Statistique :
  - *le modèle est inconnu*, mais on dispose d'*observations*
  - le but essentiel est de *compléter le modèle* à l'aide des observations
  - nécessite intuition et sens physique (NOK, pour un ordinateur...)
  - *Raisonnement inductif*

1.2 Probabilités vs statistique

- Probabilités :
  - *le modèle est complètement spécifié*
  - le but essentiel est d'*exploiter le modèle* pour prendre des décisions
  - nécessite rigueur et cohérence (OK, pour un ordinateur...)
  - *Raisonnement déductif*
- Statistique :
  - *le modèle est inconnu*, mais on dispose d'*observations*
  - le but essentiel est de *compléter le modèle* à l'aide des observations
  - nécessite intuition et sens physique (NOK, pour un ordinateur...)
  - *Raisonnement inductif*

NB.
- Dans de nombreuses applications, on peut très bien utiliser le calcul de probabilités sans faire appel à la statistique.
- Il est presque impossible de faire de la statistique sans faire appel au calcul des probabilités.

Exemple de problème de statistique

Lors d'un sondage d'opinions, on interroge un échantillon de 500 personnes habitant des zones rurales ainsi que 500 personnes habitant des zones urbaines, sur leurs intentions de vote, au second tour des élections présidentielles en France.
La table suivante reprend le résultat

	Rural	Urbain	*Total*
Candidat 1	234	245	479
Candidat 2	266	255	521
Total	500	500	1000

Peut-on conclure que les intentions de vote des deux sous-populations dont sont issus les deux échantillons ont des préférences électorales différentes ?
Est-il probable que le candidat 2 va remporter les élections ?

Exemple de problème de raisonnement probabiliste

- Supposons qu'1% de la population souffre d'une certaine pathologie.
- Un test de dépistage de cette maladie a une précision de 90%, ce qui veut dire ici que le test est positif chez 90% des personnes souffrant de la maladie et chez 10% des personnes saines.
- Une personne λ se soumet au test, et celui-ci s'avère positif.
- Quelle est la probabilité pour que cette personne souffre de la maladie ?

NB: il s'agit d'un exemple du "conditionnement".

1.3 Trois problèmes de raisonnement sous incertitude

- Problème de prédiction :

  On cherche à établir les conséquences probables d'un certain nombre de choix, sachant que les conséquences peuvent aussi être influencées par des facteurs aléatoires exogènes.

- Problème de diagnostic :

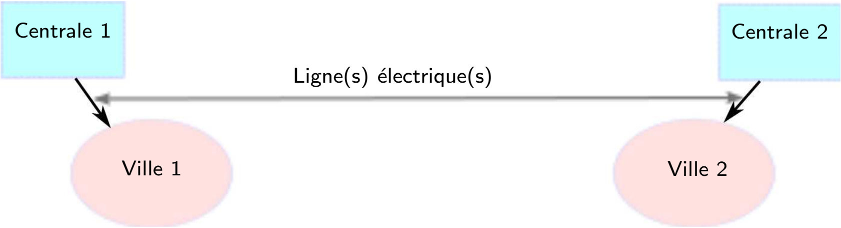
  On veut inférer les causes probables d'un phénomène observé à partir de l'observation de ses conséquences.

- Problème de décision séquentielle :

  On cherche à établir une stratégie permettant d'adapter son comportement à des informations qui seront collectées ultérieurement.

Exemple de problème de prédiction

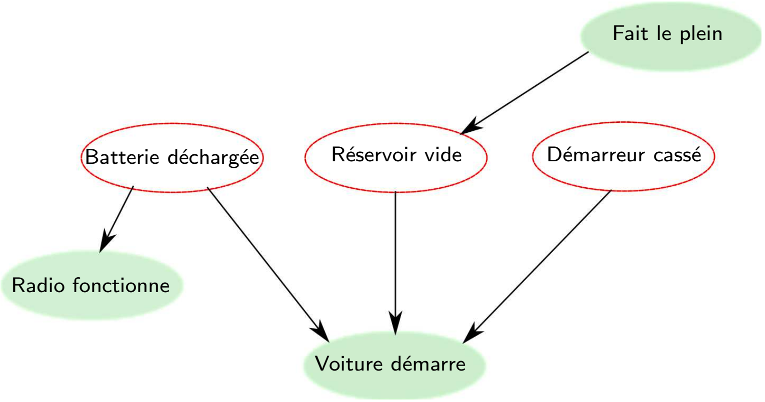
Il faut décider s'il est intéressant d'investir dans la construction d'une ligne électrique pour relier deux villes (ou deux pays, ou deux îles, ou deux continents...).



Pour ce faire, il faut prédire la probabilité de défaillance, et la quantité moyenne d'énergie non desservie, dans deux hypothèses : sans la ligne, et avec la (ou les) ligne(s) en place, puis arbitrer en fonction du coût de construction d'une ligne.

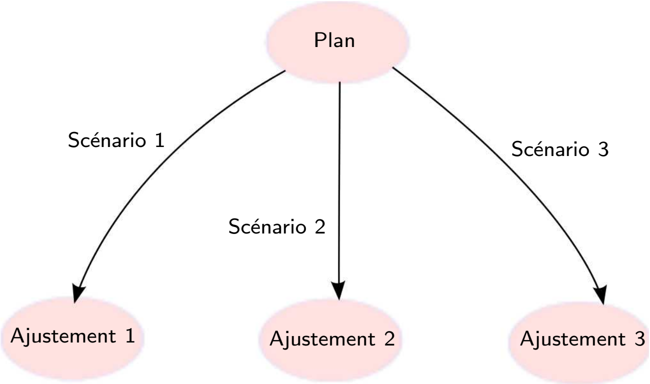
Exemple de problème de diagnostic

Ayant l'information que la voiture ne démarre pas, et sachant qu'on a fait récemment le plein et que la radio fonctionne encore, on cherche l'explication la plus probable de la panne (noeuds entourés en rouge).



Exemple de problème de décision séquentielle

Malgré l'incertitude sur les conditions du marché de l'électricité du lendemain, il faut planifier la veille quelles centrales électriques seront en service le lendemain, sachant qu'il est possible d'ajuster le lendemain le niveau de production de ces centrales. On souhaite maximiser le bénéfice économique.



1.4 Organisation du cours

- Composantes didactiques (voir agenda page suivante)
  - 6 leçons ex cathedra
  - 6 séances de répétitions
  - 2 travaux personnels à réaliser sur ordinateur (avec MatLab)
  - **séance spéciale de présentation des TPs et projets ce mercredi 8/2 de 12h40 à 13h20 à l'amphi 01 (B37)**
  - 2 séances d'information pour les projets
- Supports (et pages web) :
  - Transparents leçons et syllabus : http://montefiore.ulg.ac.be/∼lwh/Probas/
  - Répétitions et travaux pratiques: http://www.montefiore.ulg.ac.be/∼lduchesne/proba.html
- Evaluation
  - Travaux pratiques personnels : pondération ($\simeq$ 20%)
  - Examen écrit partie théorie à livre fermé ($\leq$ 40%)
  - Examen écrit partie exercices (avec formulaire fourni) ($\geq$ 40%)

Agenda des leçons et répétitions

7/2	**Leçon 1:** Introduction et Modèle probabiliste	(8h15; Amphi 202, B7b)
8/2	**Séance info TPs et projets**	(12h40; Amphi 01, B37)
14/2	**Répétition 1:** Modèle probabiliste	(8h15; R52-R53-R54; Europe)
21/2	**Leçon 2:** Variables aléatoires I	(8h15; Amphi 202, B7b)
28/2	Congé de Carnaval	
7/3	**Répétition 2:** Variables aléatoires I	(8h15; R52-R53-R54; Europe)
14/3	**Leçon 3:** Variables aléatoires II	(8h15; Amphi 202, B7b)
21/3	**Répétition 3:** Variables aléatoires II	(8h15; R52-R53-R54; Europe)
28/3	**Leçon 4:** Ensembles de variables aléatoires I	(8h15; Amphi 202, B7b)
4/4	**Répétition 4:** (Récup.) Ensembles de var. aléat. I	(8h15; Salles 01-02-S39; B37)
11/4	Congé de Pâques	
18/4	**Leçon 5:** Ensembles de var. aléat. II	(8h15; Amphi 202, B7b)
25/4	**Répétition 5:** Ensembles de variables aléatoires II	(8h15; R52-R53-R54; Europe)
2/5	**Leçon 6:** Vecteurs/processus aléatoires gaussiens	(8h15; Amphi 202, B7b)
9/5	**Répétition 6:** Vecteurs/processus aléatoires gaussiens	(8h15; Amphi 202, B7b)
16/5	**Récapitulatif et Session de Questions & Réponses**	(8h15; Amphi 202, B7b)

Chapitre 2. Le modèle probabiliste

D'abord, comprendre le modèle probabiliste...

Chapitre 2. Le modèle probabiliste

D'abord, comprendre le modèle probabiliste...

2.1 Notion de probabilité - Interprétations

2.2 Eléments de base du calcul des probabilités

2.3 Espaces de probabilités produits

2.4 Le problème du Monty-Hall

NB.

- Calcul des probabilités $\Rightarrow$ discipline mathématique
  - Préférer la rigueur à l'intuition
  - Préférer la précision à l'approximation
- Prérequis : analyse, géométrie, algèbre, analyse numérique, informatique
- Débouchés : plein d'applications et de méthodes de résolution de problèmes

2.1 Notion de probabilité - Interprétations

- Notion d'expérience aléatoire
- Notion mathématique d'espace de probabilité
- Interprétations de la notion de probabilité
- Ensembles universels non dénombrables

Notion d'expérience aléatoire

Une *expérience* est qualifiée d'aléatoire si on ne peut pas prévoir par avance son résultat, et si, répétée dans des conditions apparemment identiques, elle pourrait donner lieu à des *résultats* différents.

Pour étudier une telle expérience on s'intéresse tout d'abord à l'*univers* de tous les *résultats* (ou objets) possibles : on note

- *L'univers* (= ensemble des résultats possibles) par Ω
- *Un résultat* (= un élément quelconque de Ω) par ω

Exemple 1: Lancers de pièce

- Expérience 1.1 : Lancer simple

$$\text{Univers} : \Omega = \{P, F\}$$

Un résultat : $\omega = F$

- **Expérience 1.2 : Triple lancer**

$$\Omega = \{PPP, PPF, PFP, PFF, FPP, FPF, FFP, FFF\}.$$

Un résultat : $\omega = FPF$

- Expérience 1.3 : Lancer jusqu'au double pile

$$\Omega = \{PP, FPP, PFPP, FFPP, \ldots\},$$

Un résultat : $\omega = FFFFFPFPFFFPP$

Autres exemples

- Lancer de dés
- Distribution de cartes
- Roue de la fortune
- Diagnostic médical
- Trafic internet
- etc. etc.

NB: Des exemples de problèmes d'ingénieurs viendront plus tard

Homework : lire section 2.1.1 du syllabus...

Notion d'événement

Dans la terminologie du calcul des probabilités, un **événement** désigne une **assertion logique vérifiable** relative au résultat d'une expérience. Un événement définit donc un **sous-ensemble** de Ω.

A tout événement on peut donc faire correspondre un sous-ensemble de Ω. En particulier, à tout $\omega \in \Omega$ on peut associer un **événement élémentaire** correspondant au singleton $\{\omega\}$.

Expérience 1.2 : Triple lancer de pièces

"P est observé au plus une fois": $A = \{PFF, FPF, FFP, FFF\} \subset \Omega$
"F est observé au plus une fois": $B = \{PPP, PPF, PFP, FPP\} \subset \Omega$
"P est observé trois fois": $C = \{PPP\} \subset \Omega$ (événement élémentaire)
"F est observé au plus 3 fois": Ω (événement certain)
"F et P sont observés au moins deux fois": $\emptyset$ (événement impossible)

Algèbre des événements

On souhaite pouvoir combiner des événements pour en définir d'autres :

- $A^c = \Omega \setminus A$: complément de l'événement (≡ négation logique)
- $A \cup B$: union d'événements (≡ disjonction logique)
- $A \cap B$: intersection d'événements (≡ conjonction logique)
- $A \setminus B = A \cap B^c$
- *On désigne par $\mathcal{E}$ l'ensemble de tous les événements intéressants*

Expérience 1.2 : Triple lancer de pièces

"P est observé au moins 2 fois": $A^c = \{PPP, PPF, PFP, FPP\} \subset \Omega$
"F ou P sont observés au plus 1 fois": $B \cup A = \Omega$
"F et P sont observés au plus 1 fois": $B \cap A = \emptyset$
"P et F sont observés au moins 2 fois": $A^c \cap B^c = \emptyset$

Notion de probabilité $P(A)$ d'un événement A

Pour définir complètement le modèle probabiliste de l'expérience aléatoire (après avoir précisé son univers Ω et sa collection d'événements $\mathcal{E}$), il faut encore préciser sa mesure de probabilité

- On dit que $A \in \mathcal{E}$ se réalise, si on observe $\omega \in A$ lorsqu'on effectue l'expérience
- $\forall A \in \mathcal{E}$, la *probabilité $P(A) \in [0, 1]$* représente le degré de certitude qu'on peut associer *a priori* à la réalisation de A.
- L'ensemble des valeurs $P(A)$, $\forall A \in \mathcal{E}$ est appelé *mesure de probabilité*.
- *La mesure de probabilité traduit notre connaissance de l'expérience* **avant** *de réaliser celle-ci.*
- Si $P(A) \geq P(B)$, on dit que l'événement A est *a priori* plus probable que l'événement B.

Propriétés souhaitables d'une mesure de probabilité

Mais, quelles sont les propriétés qu'une mesure de probabilité devrait satisfaire pour qu'elle soit compatible avec notre intuition ?

- Monotonicité : si $A \subset B$ alors $P(A) \leq P(B)$. Donc
  - $P(\Omega) \geq P(A), \forall A \in \mathcal{E}$ et $P(\emptyset) \leq P(A), \forall A \in \mathcal{E}$
  - $P(A \cap B) \leq \min\{P(A), P(B)\}$ et $P(A \cup B) \geq \max\{P(A), P(B)\}$
- Cohérence :
  - $P(A) = P(B) \Leftrightarrow P(A^c) = P(B^c)$
  - $P(A) > P(B) \Leftrightarrow P(A^c) < P(B^c)$
- ...

NB: La propriété d'*additivité* ($A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$), implique la monotonicité et la cohérence.

Exemple d'assignation de mesure de probabilité

Expérience 1.2 : Triple lancer de pièces

NB: ici Ω contient 8 éléments et l'algèbre $\mathcal{E}$ contient tous les $2^8 = 256$ sous-ensembles de Ω (l'ensemble de toutes les parties de Ω est noté 2^Ω).

- Imposons que la mesure P soit *uniforme* : i.e. $P(A)$ ne dépend que du nombre d'éléments de A (i.e. de sa cardinalité $|A|$)
- Imposons aussi, par convention, que $P(\Omega) = 1$ (*événement certain*).
- Puisque $P(\Omega) = 1$, la propriété d'additivité impose $P(\emptyset) = 0$.
- Puisque $P(\Omega) = 1$, la propriété d'additivité impose aussi que $P(\{\omega\}) = 1/8$ pour tout événement élémentaire.
- En conséquence, la probabilité $P(A) = |A|/8, \forall A \in 2^\Omega$.
- *NB: La monotonicité et la cohérence sont assurées par ce choix !*

Exemple d'assignation de mesure de probabilité (suite)

Expérience 1.2 : Triple lancer de pièces *(P uniforme)*

- **"P est observé au plus une fois"**: $A = \{PFF, FPF, FFP, FFF\} \subset \Omega$
  $P(A) = 1/2$
- **"F est observé au plus une fois"**: $B = \{PPP, PPF, PFP, FPP\} \subset \Omega$
  $P(B) = 1/2$
- **"P est observé trois fois"**: $C = \{PPP\} \subset \Omega$ (événement élémentaire)
  $P(C) = 1/8$
- **"P est observé au premier lancer"**: $D = \{PPP, PPF, PFP, PFF\} \subset \Omega$
  $P(D) = 1/2$
- **"F est observé au plus 3 fois"**: Ω (événement certain)
  $P(\Omega) = 1$
- **"F et P sont observés au moins deux fois"**: $\emptyset$ (événement impossible)
  $P(\emptyset) = 0$
- etc. etc.

Formalisation : σ-algèbre $\mathcal{E}_\Omega$ d'événements

Définition de la notion de σ-algèbre sur Ω.

Une σ-algèbre $\mathcal{E}_\Omega$ d'événements, ou *tribu*, définie sur un univers Ω est une partie de 2^Ω (i.e. un ensemble de sous-ensembles de Ω) qui vérifie les propriétés suivantes :

T1. $\Omega \in \mathcal{E}_\Omega$;

T2. $A \in \mathcal{E}_\Omega \Rightarrow A^c \in \mathcal{E}_\Omega$;

T3. $\forall A_1, A_2, \ldots \in \mathcal{E}_\Omega$ (en nombre fini ou dénombrable) : $\bigcup_i A_i \in \mathcal{E}_\Omega$.

NB: nous verrons plus loin les implications en ce qui concerne les univers non finis (dénombrables ou non dénombrables)

Formalisation : mesure de probabilité P_Ω

Axiomes de Kolmogorov.

On appelle *mesure* de probabilité sur $(\Omega, \mathcal{E}_\Omega)$ (ou loi de probabilité) une fonction $P_\Omega(\cdot)$ définie sur $\mathcal{E}_\Omega$ telle que :

K1. $P_\Omega(A) \in [0,1], \forall A \in \mathcal{E}_\Omega$;

K2. $P_\Omega(\Omega) = 1$;

K3. $\forall A_1, A_2, \ldots \in \mathcal{E}_\Omega$ incompatibles deux-à-deux (et en nombre fini ou dénombrable) : $P_\Omega(\bigcup_i A_i) = \sum_i P_\Omega(A_i)$.

Dans la suite, s'il n'y a pas de risque de confusion, nous utilisons P à la place de P_Ω et $\mathcal{E}$ à la place de $\mathcal{E}_\Omega$ pour alléger les notations.

Conséquences immédiates

(A rapprocher de nos desiderata en ce qui concerne les propriétés souhaitables d'une mesure de probabilité.)

1. $P(\emptyset) = 0$.
2. $P(A^c) = 1 - P(A)$.
3. $A \subset B \Rightarrow P(A) \leq P(B)$.
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
5. $P(\bigcup_i A_i) \leq \sum_i P(A_i)$.
6. $P(A) = 1 \Rightarrow P(A \cup B) = 1, \forall B \in \mathcal{E}$.
7. $P(A) = 1 \Rightarrow P(A \cap B) = P(B), \forall B \in \mathcal{E}$.
8. $P(A) = 0 \Rightarrow P(A \cap B) = 0, \forall B \in \mathcal{E}$.
9. $P(A) = 0 \Rightarrow P(A \cup B) = P(B), \forall B \in \mathcal{E}$.

Synthèse : espace de probabilité $(\Omega, \mathcal{E}_\Omega, P_\Omega)$

L'espace $(\Omega, \mathcal{E}_\Omega, P_\Omega)$ constitue notre modèle de l'expérience aléatoire. Il est construit en définissant dans l'ordre :

1. *L'univers* Ω : un ensemble (abstrait) de résultats possibles de l'expérience (ceux que nous souhaitons ou pouvons distinguer dans le contexte applicatif considéré)
2. *La σ-algèbre d'événements* $\mathcal{E}_\Omega$: les sous-ensembles de Ω (auxquels nous souhaitons ou pouvons attribuer des probabilités)
3. *La mesure de probabilité* P_Ω : elle doit satisfaire les axiomes de Kolmogorov

Le calcul de probabilités (capacités et limitations)

- dit comment **exploiter** un modèle probabiliste de façon cohérente
- est **agnostique** en ce qui concerne l'interprétation physique du modèle
- est donc, en principe, **incompétent** pour justifier la validité du modèle

Propriétés de base : 1. Théorème des probabilités totales

Définition de la notion de système complet d'événements.

$A_1, \ldots, A_n \in \mathcal{E}$ forment un *système complet d'événements* (on dit qu'ils forment une partition de Ω) si

- $\forall i \neq j \; : A_i \cap A_j = \emptyset$ (ils sont incompatibles deux à deux)
- et si $\bigcup_{i=1}^{n} A_i = \Omega$ (ils couvrent Ω).

NB: On supposera la plupart du temps que tous les A_i sont non-vides.

Théorème des probabilités totales (version 1)

Soit $B_1, \ldots, B_n$ un système complet d'événements, alors on a :
$\forall A \in \mathcal{E} \; : P(A) = \sum_{i=1}^{n} P(A \cap B_i)$.

Propriétés de base : 2. Formule de Poincaré

Formule de Poincaré

Il s'agit de la généralisation de la formule $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. On a

$$\begin{aligned}
P(\bigcup_{i=1}^{n} A_i) \quad = \quad & \sum_{\{i_1 : 1 \leq i_1 \leq n\}} P(A_{i_1}) \\
& - \sum_{\{(i_1,i_2) : 1 \leq i_1 < i_2 \leq n\}} P(A_{i_1} \cap A_{i_2}) \\
& + \sum_{\{(i_1,i_2,i_3) : 1 \leq i_1 < i_2 < i_3 \leq n\}} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) \\
& - \ldots \\
& + (-1)^{n-2} \sum_{\{(i_1,\ldots,i_{n-1}) : 1 \leq i_1 < \cdots < i_{n-1} \leq n\}} P(A_{i_1} \cap \cdots \cap A_{i_{n-1}}) \\
& + (-1)^{n-1} P(\bigcap_{i=1}^{n} A_i).
\end{aligned}$$

Différentes interprétations de la notion de probabilité

L'interprétation du calcul des probabilités, et même sa formalisation mathématique, restent encore aujourd'hui sujet à débat scientifique (entre mathématiciens et philosophes).

Il y a différents points de vue : le point de vue objectiviste (notion mesurable de la probabilité), le point de vue classique (arguments de symétrie, d'équiprobabilité), et le point de vue subjectiviste (notion liée à l'agent qui raisonne).

Ces débats ne remettent cependant nullement en question l'intérêt pratique du calcul des probabilités. Nous y reviendrons à la fin de ce cours.

Homework : à ce stade, lire la section 2.1.3 du syllabus, et l'appendice C!

Univers Ω fini, dénombrable, et non-dénombrable

De la nécessité d'être mathématiquement précautionneux !!!

- Univers fini : ne pose pas de problèmes techniques
  C'est notre "bac à sable" du point de vue mathématique, que nous utiliserons dans la mesure du possible pour introduire les nouvelles notions de manière aussi simple que possible.
- Univers dénombrable : questions de convergence apparaissent
  Nous utiliserons ce cas comme entraînement entre le "bac à sable" et la "jungle"...
- Univers infini non dénombrable : problèmes d'une nature nouvelle apparaissent (paradoxe de Tarski-Banach, etc.)
  Nous donnerons les consignes de sécurité pour aller dans la "jungle", car elle contient de nombreux trésors pour résoudre des problèmes pratiques.

Homework : à ce stade, lire la section 2.1.4 du syllabus, et l'appendice B!

Conditionnement

- Un des objectifs principaux de ce cours est d'apprendre à conditionner : utiliser de manière effective la notion de probabilité conditionnelle et la notion d'espérance conditionnelle.
- Dans ce chapitre nous allons introduire la notion de probabilité conditionnelle pour la première fois.
- De cette notion découlent aussi les notions très importantes d'indépendance et d'indépendance conditionnelle.

Conditionnement

Scénario :

- Deux agents a_1 et a_2, tous deux de bonne foi, se sont mis d'accord sur un modèle $(\Omega, \mathcal{E}, P)$ en ce qui concerne une expérience aléatoire.
- L'expérience est réalisée par l'agent a_1, qui a pu en observer le résultat, et qui décide de révéler une information partielle à l'agent a_2, en lui indiquant que $\omega \in B$ (B étant un certain événement).
- L'agent a_2 se demande comment mettre à jour son appréciation de la probabilité pour qu'un événement quelconque A soit aussi réalisé sachant que B l'est.
- Nous noterons par $P(A|B)$ cette probabilité mise à jour, pour la distinguer de la probabilité $P(A)$ qu'il aurait assignée en l'absence de l'information qui lui a été communiquée :
  - Doit-il réviser son jugement : est-ce que $P(A|B)$ est différente de $P(A)$?
  - Si oui, comment doit-il réviser son jugement sur base des informations dont il dispose (le modèle $(\Omega, \mathcal{E}, P)$ et le fait que B soit réalisé) ?

Conditionnement : l'essence du calcul des probabilités

- Si $P(B) = 0$ notre agent a_2 doit considérer que l'agent a_1 ment, puisque cela est impossible selon le modèle probabiliste dont il dispose.
- Si $P(B) = 1$, il peut considérer que ce que lui dit l'agent est inintéressant, puisqu'il ne sait pas alors distinguer cette information du simple fait que $\omega \in \Omega$, ce qu'il savait déjà à l'avance.
- Par contre, si $P(B) \in]0, 1[$ il est obligé de réviser son jugement
  - En effet, dans ce cas il doit maintenant estimer que $P(B|B) = 1$ et $P(B^c|B) = 0$, ce qui n'était pas le cas avant qu'il ne dispose de l'information fournie par a_1.
  - De même pour tout événement C tel que $B \subset C$, il doit estimer que $P(C|B) = 1$ et pour tout événement C' tel que $C' \subset B^c$ il doit maintenant estimer que $P(C'|B) = 0$.
- Après mure réflexion, l'agent a_1 révise ses croyances en attribuant à chaque événement la probabilité conditionnelle

$$P(A|B) = P(A \cap B)/P(B).$$

Que pensez-vous de cette façon de remettre à jour les probabilités ?

Exemple de conditionnement

Expérience 1.2 : Triple lancer de pièces *(P uniforme)*

- Rappel : $\Omega = \{PPP, PPF, PFP, PFF, FPP, FPF, FFP, FFF\}$ et tous les résultats possibles sont équiprobables, et de probabilité $P(\omega) = 1/8$.
- **"P est observé au premier lancer"**: $D = \{PPP, PPF, PFP, PFF\} \subset \Omega$
  $P(D) = 1/2$
- **"P est observé au plus une fois"**: $A = \{PFF, FPF, FFP, FFF\} \subset \Omega$
  $P(A) = 1/2$
- **"P est observé au premier lancer et au plus une fois"**: $A \cap D = \{PFF\}$
  $P(A \cap D) = 1/8$
- Probabilité que **"P est observé au plus une fois"** sachant que **"P est observé au premier lancer"** :

$$P(A|D) = \frac{P(A \cap D)}{P(D)} = \frac{1/8}{1/2} = 1/4.$$

L'agent doit considérer que l'information qui lui est révélée a pour effet de diminuer la probabilité de l'événement A.

Conditionnement : l'essence du calcul des probabilités

Adoptons la définition : $P(A|B) = P(A \cap B)/P(B)$ et regardons un peu les conséquences de cette adoption.

- $P(A \cap B) = P(B)P(A|B)$: la probabilité que A et B se réalisent est égale au produit de la probabilité que B se réalise par la probabilité conditionnelle que A soit réalisé sachant que B l'est.
- Comme $P(A \cap B) = P(B \cap A)$ on a $P(B)P(A|B) = P(A)P(B|A)$.
- Quid si on lui avait dit que B^c est réalisé ? $P(A \cap B^c) = P(B^c)P(A|B^c)$ et $P(A) = P(A \cap B) + P(A \cap B^c)$. Donc
$$P(A) = P(B)P(A|B) + P(B^c)P(A|B^c)$$
- On déduit que
$$P(B|A) = \frac{P(B)P(A|B)}{P(B)P(A|B) + P(B^c)P(A|B^c)}$$
- Quid si on lui dit d'abord que B_1 est réalisé puis que B_2 est réalisé, vs on lui dit tout de suite que $B = B_1 \cap B_2$ est réalisé.
$$P(A|B_1 \cap B_2) = P(A \cap B_2|B_1)/P(B_2|B_1) = P(A \cap B_1|B_2)/P(B_1|B_2)$$

Probabilité conditionnelle : définition

Probabilité conditionnelle de A sachant B (avec $P(B) > 0$)

$$P(A|B) \triangleq \frac{P(A \cap B)}{P(B)}. \tag{1}$$

- Cette formule définit une mesure de probabilité (qui satisfait les axiomes de Kolmogorov) sur $(\Omega, \mathcal{E})$.
- Le conditionnement est une opération symétrique : conditionner d'abord sur B puis sur C est équivalent à conditionner d'abord sur C puis sur B.

Notion d'événements indépendants (définition 1)

Evénements indépendants (première définition)

On dit que A est indépendant de B si $P(A|B) = P(A)$, c'est-à-dire si le fait de savoir que B est réalisé ne change en rien la probabilité de A. Pour dire que A est indépendant de B on utilisera la notation

$$A \perp B. \tag{2}$$

Notons que si $P(B) \in]0,1[$, on a A indépendant de B si, et seulement si, $P(A|B) = P(A|B^c)$.

(*Suggestion : calculer alors* $P(A)$ *par le théorème des probabilités totales.*)

Notion d'événements indépendants (propriétés 1)

Propriétés "positives" :

- $\emptyset$ est indépendant de tout autre événement.
- Un événement de probabilité nulle est indépendant de tout autre événement :
  $P(A) = 0 \Rightarrow P(A \cap B) = 0 \Rightarrow P(A|B) = 0$.
- Tout événement est indépendant de Ω.
- Tout événement est indépendant de tout événement certain :
  $P(A) = 1 \Rightarrow P(A \cap B) = P(B)$ et donc $P(B|A) = P(B)$.
- "A indépendant de B" $\Leftrightarrow P(A \cap B) = P(A)P(B)$
  (conséquence directe de la définition, lorsque $P(B) > 0$; si $P(B) = 0$, la condition $P(A \cap B) = P(A)P(B)$ est encore vérifiée).
- "A indépendant de B" $\Rightarrow$ "B indépendant de A".
- "A indépendant de B" $\Rightarrow$ "A^c indépendant de B".
- "A indépendant de B" $\Rightarrow$ "A indépendant de B^c".
- "A indépendant de B" $\Rightarrow$ "A^c indépendant de B^c".

Notion d'événements indépendants (définition 2)

On peut donc utiliser en lieu et place de la définition de l'indépendance la définition suivante.

Evénements indépendants (seconde définition)

Deux événements A et B sont indépendants si et seulement si

$$P(A \cap B) = P(A)P(B).$$

Il est à noter que cette définition couvre aussi le cas où $P(A)$ et/ou $P(B)$ sont nulles, la propriété étant trivialement vérifiée dans ces cas (car $0 \leq P(A \cap B) \leq \min\{P(A), P(B)\}$).

On voit bien que la relation $\perp$ est symétrique.

Exemple d'événements indépendants

Expérience 1.2 : Triple lancer de pièces *(P uniforme)*

- Rappel : $\Omega = \{PPP, PPF, PFP, PFF, FPP, FPF, FFP, FFF\}$ et tous les résultats possibles sont équiprobables, et de probabilité $P(\omega) = 1/8$.
- **"P est observé au premier lancer"**: $A = \{PPP, PPF, PFP, PFF\}$, $P(A) = 1/2$
- **"P est observé au troisième lancer"**: $B = \{PPP, PFP, FPP, FFP\}$, $P(B) = 1/2$
- **"P est observé au second lancer"**: $C = \{PPP, PPF, FPP, FPF\}$, $P(C) = 1/2$
- **"On observe le même résultat au premier et au troisième lancer"**: $D = \{PPP, PFP, FPF, FFF\}$, et donc $P(D) = 1/2$
- $A \cap B = \{PPP, PFP\}$. Ainsi $P(A \cap B) = 1/4 = P(A)P(B)$ et donc $A \perp B$.
- De même, on montre que $A \perp C$ et $B \perp C$.
- $A \cap D = \{PPP, PFP\}$ et donc $P(A \cap D) = 1/4$ et donc $A \perp D$.
- De même, on montre que $B \perp D$ et $C \perp D$.

Notion d'événements indépendants (propriétés 2)

Propriétés "négatives"

- Un événement quelconque non trivial n'est jamais indépendant de lui-même!
- A indépendant de B et B indépendant de $C \not\Rightarrow A$ indépendant de C. *(Suggestion : à titre de contre-exemple, prendre A et B tous deux de probabilité non nulle et indépendants, et puis considérer la cas où $C = A$).*
- A dépendant de B et B dépendant de $C \not\Rightarrow A$ dépendant de C. *(Suggestion : prendre A et C indépendants, et $B = A \cap C$ en supposant que $P(B) > 0$.)*
- A indépendant de $B \not\Rightarrow A$ indépendant de B conditionnellement à C. *(Suggestion : prendre comme exemple le double "pile ou face" avec une pièce équilibrée, comme événements A "face au premier lancer", B "face au second lancer", C "même issue aux deux lancers").*

Indépendance mutuelle vs indépendance 2-à-2

Indépendance mutuelle de *n* événements.

On dira que les événements $A_1, A_2, \ldots, A_n$ sont *mutuellement* indépendants si pour toute partie I de l'ensemble des indices allant de 1 à n on a :

$$P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i). \tag{3}$$

Il est important de noter que l'indépendance mutuelle est une condition plus forte que l'indépendance deux à deux, qui elle est définie comme suit.

Indépendance deux à deux de *n* événements.

On dira que les événements $A_1, A_2, \ldots, A_n$ sont indépendants *deux à deux* si $\forall i \neq j$ on a :

$$P\left(A_i \cap A_j\right) = P(A_i)P(A_j). \tag{4}$$

Exemple: indépendance mutuelle vs indépendance 2-à-2

Expérience 1.2 : Triple lancer de pièces *(P uniforme)*

- Rappel : $\Omega = \{PPP, PPF, PFP, PFF, FPP, FPF, FFP, FFF\}$ et tous les résultats possibles sont équiprobables, et de probabilité $P(\omega) = 1/8$.
- **"P est observé au premier lancer"**: $A = \{PPP, PPF, PFP, PFF\}$, $P(A) = 1/2$
- **"P est observé au troisième lancer"**: $B = \{PPP, PFP, FPP, FFP\}$, $P(B) = 1/2$
- **"P est observé au second lancer"**: $C = \{PPP, PPF, FPP, FPF\}$, $P(C) = 1/2$
- **"On observe le même résultat au premier et au troisième lancer"**: $D = \{PPP, PFP, FPF, FFF\}$, et donc $P(D) = 1/2$
- A, B, C, D sont indépendants deux-à-deux (voir transparent 47)
- A, B, C sont aussi mutuellement indépendants : en effet, en plus de leur indépendance 2-à-2 on a $P(A \cap B \cap C) = 1/8 = P(A)P(B)P(C)$
- Mais les événements A, B et D ne sont pas mutuellement indépendants : en effet on a que $P(A \cap B \cap D) = 1/4 \neq P(A)P(B)P(D) = 1/8$.

Formules de Bayes

Passage de $P(A|B)$ à $P(B|A)$:

Première formule de Bayes ($P(A)$ et $P(B)$ non nulles)

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}. \tag{5}$$

Conséquence immédiate de la définition de la probabilité conditionnelle.

On peut aussi reformuler le théorème des probabilités totales en exploitant la définition de la probabilité conditionnelle :

Théorème des probabilités totales (version 2)

Si $B_1, B_2, \ldots, B_n$ est un système complet d'événements non triviaux ($P(B_i)$ non nulles) alors le théorème des probabilités totales peut s'écrire sous la forme suivante

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i). \tag{6}$$

Formules de Bayes (suite)

Dès lors la formule de Bayes peut aussi s'écrire sous la forme suivante.

Deuxième formule de Bayes

Si $B_1, B_2, \ldots, B_n$ est un système complet d'événements non triviaux alors la première formule de Bayes peut s'écrire sous la forme suivante

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{k=1}^{n} P(A|B_k)P(B_k)}, \tag{7}$$

qui s'appelle aussi théorème sur la "probabilité des causes".
En effet, il permet de calculer les probabilités de chaque cause possible d'un événement (la conséquence) sachant que celui-ci s'est réalisé, à partir de la probabilité de celui-ci conditionnellement à n'importe quelle cause, et de la probabilité a priori de chaque cause.
Ce théorème n'est valable que sous l'hypothèse qu'une seule cause à la fois peut se réaliser (caractère mutuellement exclusif des causes B_i).

Problème à résoudre

Vous pouvez maintenant essayer de résoudre le problème suivant :

- Supposons qu'1% de la population souffre d'une certaine pathologie.
- Un test de dépistage de cette maladie a une précision de 90%, ce qui veut dire ici que le test est positif chez 90% des personnes souffrant de la maladie et chez 10% des personnes saines.
- Une personne λ se soumet au test, et celui-ci s'avère positif.
- Quelle est la probabilité pour que cette personne souffre de la maladie ?

NB: il s'agit d'un exemple du "conditionnement".

Je vous suggère de rendre les choses concrètes en supposant que la population est composée de 1 000 000 individus.

2.3 Espaces de probabilités produits

- La notion de "produit" de deux espaces de probabilité $(\Omega_1, \mathcal{E}_1, P_1)$ et $(\Omega_2, \mathcal{E}_2, P_2)$ permet de construire un modèle de probabilité $(\Omega = \Omega_1 \times \Omega_2, \mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2, P = P_1 P_2)$ relatif à la combinaison de deux expériences aléatoires indépendantes, à partir des modèles de ces expériences.
- Cette notion peut-être généralisé au produit d'un nombre fini quelconque d'espaces de probabilités.
- Nous reviendrons sur cela plus tard dans ce cours.

Homework : à ce stade, lire la section 2.3 du syllabus !

2.4 Le problème du Monty-Hall

In the September 9, 1990 issue of Parade magazine, the columnist Marilyn vos Savant responded to this letter:

> *Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?*
>
> *Craig. F. Whitaker*
> *Columbia, MD*

The letter roughly describes a situation faced by contestants on the 1970's game show Let's Make a Deal, hosted by Monty Hall and Carol Merrill. Marilyn replied that the contestant should indeed switch. But she soon received a torrent of letters - many from mathematicians - telling her that she was wrong. The problem generated thousands of hours of heated debate.

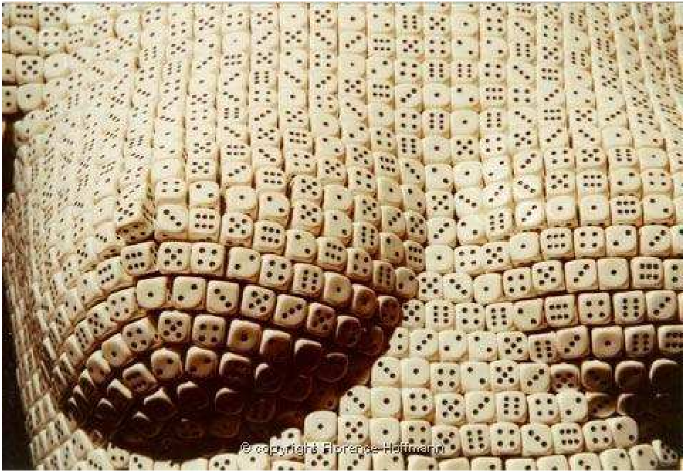
Ce problème sera traité lors de la séance de demain ! Ensuite, lire attentivement la section 2.4 du syllabus.

Conseils de lecture du syllabus

Signification des notations • et ◦ utilisées dans les notes :

- • : ce symbole désigne des sections fort importantes, mais dont la matière est difficile. Il faut les lire une première fois quand la matière est vue au cours, puis relire plus tard à tête reposée.
- ◦ : ces sections comportent des notions utiles pour les enseignements ultérieurs, mais ne sont pas fondamentales pour comprendre la suite de ce cours; à lire quand même !

Exemple de “processus aléatoire spatial” (cf Chapitre 5)



Sculpture “Le hasard fait bien les choses” de Florence Hoffmann (Luxembourg)