Notes de cours en calcul des probabilités (JJ Bellanger)
I : Espaces Probabilisés

I : ESPACES PROBABILISES

I.1-Expérience aléatoire

Intuitivement une expérience aléatoire est une expérience dont on ne peut pas prévoir le résultat exact : en la répétant plusieurs fois dans des conditions strictement identiques (pour ce qui est des paramètres maîtrisés par l'expérimentateur) les résultats obtenus pourrons différer. Le résultat d'un essai est a priori non prévisible. La caractérisation d'une expérience aléatoire peut évidemment se faire en décrivant par le menu le protocole suivi ('matériel' utilisé, réglages,…) . Un autre point de vue, dit phénoménologique (et qui ne s'intéresse en quelque sorte qu'à 'l'extérieur des choses ') consiste à ne considérer que l'ensemble des résultats possibles et à essayer de quantifier le 'pourcentage de chances' pour que le résultat obtenu présente telle où telle caractéristique comme par exemple le résultat est pair dans le lancement d'un dé, la casaque du jockey vainqueur est rouge, etc… Ceci se formalise en introduisant

- l'ensemble Ω des résultats possibles,
- des sous-ensembles de Ω sur lesquels une même propriété est vérifiée (pair, rouge,..) et qu'on appellera événements (c'est le fait qu'on ne sache pas à l'avance si oui ou non le résultat $\omega \in \Omega$ obtenu vérifiera cette propriété qui constitue un événement à l'issue de l'expérience),
- des valeurs affectées à chaque événement, dîtes valeurs de probabilité, comprises entre 0 et 1 et qui s'interprètent comme un pourcentage de chances (ramené entre 0 et 1) pour que l'événement se réalise.

La formalisation mathématique intégre des notions nécessaires du point de vue de l'utilisateur et permet

1) d'en assurer la cohérence logique,
2) d'introduire un calcul d'une part sur les événements et d'autre part sur les probabilités d'événements.

Le premier type de calcul utilise des opérations ensemblistes qui portent sur les événements (qui sont des sous-ensembles de Ω) et le deuxième utilise des opérations arithmétiques puisque les quantités manipulées sont alors des nombres réels (compris entre 0 et 1). C'est pour disposer d'un cadre non ambigu que sont introduites les structures mathématiques 'espace probabilisable' et 'espace probabilisé', définies dans ce qui suit.

I.2-Espaces probabilisables

1.Définition

On appelle <u>espace probabilisable</u> une paire (Ω, τ) dans laquelle Ω est un ensemble quelconque[1] et τ une collection de sous-ensembles de Ω appelée <u>tribu</u> (on dit <u>tribu d'événements</u>) qui répond par définition aux contraintes (axiomes) suivantes :

1. $\Omega \in \tau$
2. si $E \in \tau$ alors $\overline{E} \in \tau$ où $\overline{E}$ désigne le complément de E dans Ω
3. si $E \in \tau$ et $F \in \tau$ alors $E \cup F \in \tau$
4. pour toute famille d'événements $E_i, i \in I$ où I est un ensemble dénombrable d'indices $\bigcup_I E_i \in \tau$.

[1] quelconque du point de vue mathématique mais pas de celui de l'utilisateur qui l'interprétera comme l'ensembles des résultats possibles d'une expérience aléatoire

Corollaires :

1. $\Phi \in \tau$
2. toute intersection finie ou dénombrable d'événements d'une tribu est un événement de cette même tribu. Autrement dit l'opération d'intersection dénombrable est fermée dans une tribu

Remarque : les axiomes 2. et 3. et 4. expriment la *fermeture* des opérations ensemblistes de complémentation et d'union respectivement *non dénombrable* (pour 3.) *et dénombrable* (pour 4.). Si on ne conserve que les trois premiers axiomes la collection d'événements est appelée *algèbre d'événements*. Toute tribu est évidemment une algèbre alors que la réciproque est fausse.

2.Langage ensembliste et langage des événements

Le résultat d'une expérience étant un élément ω de Ω on dira que l'événement (sous ensemble) E est réalisé si $\omega \in \Omega$ et que E n'est pas réalisé si c'est l'événement complémentaire $\overline{E}$ qui contient ω. On a une correspondance immédiate entre le langage des sous ensembles et celui des probabilités. Celle ci est précisée dans le tableau ci-dessous :

Langage ensembliste	*Langage probabiliste*
Sous ensemble vide Φ	Evénement impossible Φ (jamais réalisé)
Ensemble fondamental Ω	Evénement certain Ω (toujours réalisé)
Sous ensembles E et F disjoints	Evénements E et F incompatibles
Union de E et F	E est réalisé ou F est réalisé (ou non exclusif)
Intersection de E et F	E est réalisé et F est réalisé
E$\subset$F	E implique F

3.Rappel (*de quelques propriétés des opérations ensemblistes*)

- $E \cap (\cup_i F_i) = \cup_i (E \cap F_i)$, $E \cup (\cap_i F_i) = \cap_i (E \cup F_i)$
- formules de De Morgan : $\overline{\cup_i E_i} = \cap_i \overline{E}_i$ et $\overline{\cap_i E_i} = \cup_i \overline{E}_i$
- Si $(E_i, i \in I)$ est une partition de Ω (cad si les E_i sont disjoints 2 à 2 et leur union sur *I* est égale à Ω) alors pour tout *F* de Ω on a : $F = \cup_{i \in I}(E_i \cap F)$

- Soit une application quelconque $f : \Omega_1 \to \Omega_2$ définie sur un ensemble quelconque vers un autre ensemble également quelconque. En notant

$f^{-1}(E) = \{\omega \in \Omega_1 : f(\omega) \in E \subset \Omega_2\}$ on a :

$f^{-1}(E \cup F) = f^{-1}(E) \cup f^{-1}(F)$, $f^{-1}(E \cap F) = f^{-1}(E) \cap f^{-1}(F)$, $f^{-1}(\overline{E}) = \overline{f^{-1}(E)}$

4.Choix d'une tribu sur un ensemble donné Ω.

Sur un même ensemble Ω (à condition qu'il contienne au moins 2 éléments) il est toujours possible de sélectionner plus d'une tribu.

Par exemple :

- la tribu minimale (la plus petite) est $\tau_m = \{\Phi, \Omega\}$,
- à chaque événement *E* on peut associer la tribu $\tau_E = \{\Phi, \Omega, E, \overline{E}\}$,
- à toute partition $D = (E_i, i \in I)$ de Ω on peut associer la tribu τ_D constituée par tous les ensembles de la forme $\cup_{i \in I} G_i$ où, pour tout *i*, G_i prend soit la valeur E_i soit la valeur Φ. Formellement on a donc $\tau = \{(\cup_{i \in J} E_i) : J \in \text{Pt}(I)\}$ où Pt(*I*) désigne l'ensemble des parties de *I*,

- $\tau_M = \mathrm{Pt}(\Omega)$, l'ensemble des parties, se trouve être toujours, par construction, une tribu sur Ω. C'est la tribu la plus grosse (la plus grande) sur Ω puisqu'elle contient tous ses sous-ensembles.

5.Propriété

Soit $(\tau_\alpha, \alpha \in \mathrm{A})$ une famille (finie, dénombrable ou même non dénombrable) de tribus sur un même ensemble Ω. Alors il est très simple de vérifier que la collection des sous-ensembles qui sont, chacun, dans chaque tribu, c'est à dire la collection $\cap_{\alpha \in A} \tau_\alpha = \{E : \forall \alpha, E \in \tau_\alpha\}$ *obtenue par intersection des tribus est elle même une tribu*.
Cette propriété permet d'introduire la notion de *tribu engendrée* par une famille d'événements (sous-ensembles).

6.Définition

On appelle tribu engendrée sur Ω par la collection $C = (E_t, t \in T)$ où T peut être *dénombrable ou non* l'intersection (qui a un sens d'après la propriété ci-dessus) de toutes les tribus sur Ω qui contiennent, chacune, cette collection. On notera $\tau(C)$ cette tribu.

On a donné plus haut l'exemple des tribus τ_E, τ_D qui sont engendrées respectivement, au sens de la définition donnée ci-dessus, pour la première par l'événement E et pour la seconde par la partition C.

7.Quelques exemples classiques d'espaces probabilisables

- *Cas des espaces discrets.*

On dit que Ω est discret si Ω est dénombrable : $\Omega = \{\omega_i, i \in I \subset N\}$. Dans ce cas on prend fréquemment, bien que cela ne soit bien entendu pas obligatoire, $\tau = Pt(\Omega)$.

- *Cas où $\Omega = R$ (droite réelle).*

On n'utilise jamais l'ensemble des parties pour des raisons techniques (difficulté de construction d'une mesure de probabilité, notion abordée plus loin dans ce chapitre, sur cette tribu). La tribu de loin la plus utilisée dans ce cas est la tribu appelée *tribu des boréliens* de R ou encore tribu de Borel sur R (du nom du mathématicien français Emile Borel). Elle peut être définie comme étant la tribu engendrée par la famille des intervalles fermés $[a,b]$ de R. Il est cependant vérifiable que cette tribu coïncide, *entre autres*, avec chacune des tribus suivantes :
  - tribu engendrée par les $]a,b[$,
  - tribu engendrée par les $[a,b]$,
  - tribu engendrée par les $[a,b[$,
  - tribu engendrée par les $]-\infty,a[$,
  - tribu engendrée par les ouverts de la topologie usuelle sur R.

- *Produit d'espaces probabilisables*

Soient (Ω_1, τ_1) et (Ω_2, τ_2) 2 espaces probabilisables. On introduit alors l'espace probabilisable (Ω_3, τ_3) défini par :
  - $\Omega_3 = \Omega_1 \times \Omega_2$
  - $\tau_3 = \tau(\{E_1 \times E_2, E_1 \in \tau_1, E_2 \in \tau_2\})$

La tribu τ_3 est appelée produit (ou produit tensoriel) des tribus τ_1 et τ_2 et est notée $\tau_3 = \tau_1 \otimes \tau_2$. L'espace (Ω_3, τ_3) est appelé produit des espaces (Ω_1, τ_1) et (Ω_2, τ_2).
Ceci peut s'étendre au produit d'un nombre quelconque d'espaces probabilisables. C'est ainsi que peuvent être introduits les espaces probabilisables (R^N, B^N) en faisant le produit de *N* espaces *(R,B)*. La tribu B^N est appelée tribu des boréliens sur R^N.

- *Tribus définies à partir d'une application* $f : \Omega_1 \to \Omega_2$
  - Si τ_2 est une tribu sur Ω_2 la collection $\{f^{-1}(E), E \in \Omega_2\}$ de sous ensembles de Ω_1 est une tribu sur Ω_1 (exercice : le vérifier). Cette tribu est appelée tribu *engendrée* sur Ω_1 par f.
  - Si τ_1 est une tribu sur Ω_1 la collection $\{E \in \Omega_2 / f^{-1}(E) \in \tau_1\}$ de sous ensembles de Ω_2 est une tribu sur Ω_2 (exercice : le vérifier). Cette tribu est appelée tribu *induite* sur Ω_2 par f.

I.2-Espaces probabilisés

1.Définition d'une mesure de probabilité
Etant donné un espace probabilisable (Ω, τ) on appelle mesure de probabilité sur cet espace toute application de la forme :
$P : E \in \tau \to P(E) \in [0,1]$
qui associe à tout événement *E* une valeur de probabilité comprise entre 0 et 1 et qui est telle que :
i) normalisation : $P(\Omega) = 1$
ii) additivité dénombrable : pour toute famille dénombrable $E_i, i \in I$ d'événements dans τ, 2 à 2 incompatibles, on a : $P(\bigcup_{i \in I} E_i) = \sum_{i \in I} P(E_i)$

2.Propriétés d'une mesure de probabilité
- $P(\varnothing) = 0$,
- $P(\overline{E}) = 1 - P(E)$,
- $\forall E, F$ dans τ $P(E \cap F) \le P(E)$ et $E \subset F \Rightarrow P(E) \le P(F)$
- pour toute famille dénombrable $E_i, i \in I$ on a : $P(\bigcup_{i \in I} E_i) \le \sum_{i \in I} P(E_i)$,
- $P(A \bigcup B \bigcup C) = P(A) + P(B) + P(C) - P(A \bigcap B) - P(B \bigcap C) - P(A \bigcap C) + P(A \bigcap B \bigcap C)$
  qui se généralise pour $P(\bigcup_{n=1..N} E_n)$ par la formule de Poincaré (voir exercice),
- formule des probabilités totales : si $\Omega = \sum_{i \in Idén} E_i$ ($\sum$ désigne ici une union d'événements <u>disjoints</u>), tous les E_i étant dans τ, alors pour tout *F* dans τ on a $P(F) = \sum_{i \in I} P(F \cap E_i)$,
- continuité :
  - pour toute suite non croissante d'événements $E_n \supset E_{n+1}, n \ge 1$ on a $P(\bigcap_{n \ge 1} E_n) = \lim_{n \to \infty} P(E_n)$,

- pour toute suite non décroissante d'événements $E_{n+1} \supset E_n, n \geq 1$ on a $P(\bigcup_{n\geq 1} E_n) = \lim_{n\to\infty} P(E_n)$,
- En posant respectivement $\lim_{n\to\infty} E_n = \bigcap_{n\geq 1} E_n$ dans le cas non croissant et $\lim_{n\to\infty} E_n = \bigcup_{n\geq 1} E_n$ dans l'autre cas ces 2 propriétés se résument par

$P(\lim_{n\to\infty} P(E_n)) = \lim_{n\to\infty} P(E_n)$

Vocabulaire : tout triplet (Ω, τ, P) où (Ω, τ) est un espace probabilisable et P une mesure sur cet espace sera appelé **espace probabilisé.**

I.3-Probabilités conditionnelles

1.Définition

Soit l'espace probabilisé (Ω, τ, P). Pour les événements E et F dans τ avec $P(F) > 0$ on appelle probabilité conditionnelle de E si F le nombre $P(E/F) = \dfrac{P(E \cap F)}{P(F)}$ qui est compris entre 0 et 1 (par construction).

Si $P(F) > 0$ et $P(E) > 0$ on peut écrire : $P(E \cap F) = P(E/F)P(F) = P(F/E)P(E)$

2.Propriété

L'application $P_F : E \in \tau \to P_F(E) = P(E/F)$ est une nouvelle mesure de probabilité sur (Ω, τ) appelée mesure conditionnée par E.

Remarque : si $E \subset F$ on a obligatoirement $P(E/F) \geq P(E)$ et si $E \cap F = \varnothing$ on a obligatoirement $P(E/F) = 0$

3.Formule de Bayes

Soit une partition $\Omega = \sum_{i \in I \text{ dén}} E_i$ avec les probabilités $P(E_i) = \pi_i > 0 \, \forall i$ connues. Pour tout événement E pour lequel les probabilités conditionnelles $P(F/E_i)$ sont également connues on pourra écrire :

$$P(E_i / F) = \frac{P(F/E_i)}{\sum_{i \in I} \pi_k P(F/E_k)}, i \in I$$

Cette formule peut être utilisée en pratique lorsque qu'un expérimentateur se demande, à l'issue de l'expérience aléatoire, laquelle des 'hypothèses' E_i a été réalisée (sans qu'il puisse l'observer directement), alors que par contre il est averti de la réalisation d'un événement F. En l'absence de cette dernière information la probabilité, dite probabilité a priori, pour que E_i soit réalisé est π_i. Lorsqu'on est averti de la réalisation de F, la probabilité pour que E_i soit réalisé conditionnellement à F, appelée probabilité a posteriori, peut être calculée avec la formule de Bayes.

I.4-Exemples d'espaces probabilisés

Notes de cours en calcul des probabilités (JJ Bellanger)

I : Espaces Probabilisés

1.Les espaces discrets

L'espace probabilisé (Ω,τ,P) est dit discret si $\Omega=\bigcup_{i\in I}\omega_i$ où I est soit fini soit infini dénombrable et les ω_i sont des singletons. La tribu choisie est alors généralement l'ensemble des parties $Pt(\Omega)$. Pour spécifier une mesure P sur (Ω,τ) il suffit de se donner les valeurs de probabilité pour tous les singletons de Ω : $p_i=P(\{\omega_i\}),\omega_i\in\Omega$. On a :

- $\forall i\ p_i\geq 0$ et $\sum_i p_i=1$
- $\forall E\in\tau\ P(E)=\sum_{i:\omega_i\in E}p_i$

Dans le cas particulier où Ω est fini, de cardinal N et où $p_1=..=p_N=\dfrac{1}{N}$ on a :

$\forall E\in\tau\ P(E)=\dfrac{card(E)}{card(\Omega)}=\dfrac{\text{'nombre de cas favorables'}}{\text{'nombre de cas possibles'}}$ dans le vocabulaire traditionnel.

2.Les espaces infinis non dénombrables

Nous nous contentons ici de considérer le cas où $(\Omega,\tau)=(R^N,B^N)$. La tribu B^N choisie ordinairement est celle engendrée par tous les pavés N-dimensionnels de R^N, cad tous les ensembles de la forme $[a_1,b_1]\times..\times[a_N,b_N]$. Elle est appelée tribu des boréliens sur R^N comme dans le cas uni-dimensionnel.

Un résultat élémentaire de l'expérience aléatoire étant ici une collection de N nombres réels on se trouve, pour spécifier P, dans la même démarche que celle consistant à spécifier la loi d'une variable aléatoire N-dimensionnelle. On se reportera pour cela en **II.2** : loi de probabilité d'une variable aléatoire.

I.4-Indépendance entre événements

1.Définition

Soit un espace probabilisé (Ω,τ,P). Par définition les N événements $E_1,..,E_N$ pris dans τ (et qui correspondent donc à N sous-ensembles de Ω) sont dits indépendants dans l'ensemble pour la mesure de probabilité P si pour tout sous-ensemble de K indices distincts $i_1,..,i_K$ pris dans $\{1,..,N\}$, avec $2\leq K\leq N$ on a $P(\bigcap_{k=1..K}E_k)=\prod_{k=1..K}P(E_k)$.

Remarques :

- On a donc en particulier
  $P(\bigcap_{k=1..N}E_k)=\prod_{k=1..N}P(E_k)$ et $\forall(i,j),1\leq i,j\leq N),i\neq j,P(E_i\cap E_j)=P(E_i)P(E_j)$
  <u>mais ces seules conditions ne sont pas en général équivalentes à l'indépendance dans l'ensemble (elles ne suffisent pas).</u>
- Des événements de (Ω,τ) indépendants dans l'ensemble pour une mesure P ne seront pas, en général, indépendants dans l'ensemble pour une autre mesure Q sur le même espace.

2.Définition

N tribus $\tau_1,..,\tau_N$ sur (Ω,τ,P) sont dites indépendantes dans l'ensemble si $\forall E_1,..,E_N$ pris respectivement dans $\tau_1,..,\tau_N$, $E_1,..,E_N$ est une famille indépendante dans l'ensemble pour P.