Université Paul Verlaine – Metz
Préparation concours 2 (L2, BE) Année universitaire 2010-2011
A. Pasquale

Quelques notions de probablité

Une expérience dont le résultat ne peut pas être prévu avec certitude, est dite *expérience aléatoire*. En général, on suppose que l'ensemble des résultats possibles de l'expérience est connus.
Donné une expérience aléatoire, soit Ω l'ensemble des résultats possibles de l'expérience. Il est dit l'*ensemble fondamentale* ou *univers*. Un élément de Ω est dit une *éventualité*.
Un *événement* est une partie A de Ω. L'*ensemble des événements* est une famille $\mathcal{T}$ de parties de Ω vérifiant :
(1) $\Omega \in \mathcal{T}$;
(2) si $A \in \mathcal{T}$, alors $\Omega \setminus A \in \mathcal{T}$
(3) Si $\{A_n\}_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{T}$, alors $\cup_{n\in\mathbb{N}} A_n \in \mathcal{T}$
Le couple $(\Omega, \mathcal{T})$ est appelé *espace probabilisable*.

Propriétés : Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Alors :
(1) $\emptyset \in \mathcal{T}$
(2) Si $\{A_n\}_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{T}$, alors $\cap_{n\in\mathbb{N}} A_n \in \mathcal{T}$
(3) Si $A, B \in \mathcal{T}$, alors $A \cap B$, $A \cup B$ et $A \setminus B$ appartiennent à $\mathcal{T}$.

Soit $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω. Alors $\mathcal{P}(\Omega)$ satisfait les propriétés de l'ensemble des événements $\mathcal{T}$.
Si Ω est un ensemble fini ou dénombrable, alors on choisit toujours $\mathcal{T} = \mathcal{P}(\Omega)$. Ainsi un espace probabilisable avec ensemble fondamentale fini ou dénombrable est $(\Omega, \mathcal{P}(\Omega))$.

Deux événements $A, B \in \mathcal{T}$ sont dits *incompatibles* si $A \cap B = \emptyset$.

Un *système complet d'événements* d'un espace probabilisable est une famille finie ou dénombrable $\{A_n\}_{n\in I}$, où $I = \mathbb{N}$ ou $I = \{1, 2, \ldots, n\}$, telle que :
(1) pour tout $n \in I$ on a $A_n \in \mathcal{T}$
(2) pour tous n, m avec $n \neq m$ on a $A_n \cap A_m = \emptyset$
(3) $\Omega = \cup_{n\in\mathbb{N}} A_n$.

On appelle *probabilité* sur l'espace probabilisable $(\Omega, \mathcal{T})$ toute application $P : \mathcal{T} \to [0, +\infty[$ vérifiant :
(1) $P(\Omega) = 1$
(2) Pour toute suite $\{A_n\}_{n\in\mathbb{N}}$ d'éléments de $\mathcal{T}$ tels que $A_n \cap A_m = \emptyset$ pour tous $n \neq m$, on a

$$P\big(\cup_{n\in\mathbb{N}} A_n\big) = \sum_{n=0}^{+\infty} P(A_n)\,.$$

La valeur $P(A)$ est dite la *probabilité de l'événement* A.
Un *espace probabilisé* est un triplet $(\Omega, \mathcal{T}, P)$ où $(\Omega, \mathcal{T})$ est un espace probabilisable et P est une probabilité sur $(\Omega, \mathcal{T})$

Propriétés :
(1) Pour tout $A \in \mathcal{T}$ on a $P(\Omega \setminus A) = 1 - P(A)$ (car $1 = P(\Omega) = P(A) \cup P(\Omega \setminus A)$)
En particulier, $0 \leq P(A) \leq 1$, et $P(\emptyset) = 0$.
(2) Si $A, B \in \mathcal{T}$ et $A \subset B$, alors $P(A) \subset P(B)$.
(3) Si $A, B \in \mathcal{T}$, alors $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
(4) Si $A, B \in \mathcal{T}$, alors $P(A \cup B) \leq P(A) + P(B)$.
Puisque $0 \leq P(A) \leq 1$, on a que P est une application à valeurs dans $[0, 1]$.
$P(A \cup B)$ est la *probabilité totale* de A et B.
$P(A \cap B)$ est la *probabilité composée* de A et B.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé.
Tout événement $A \in \mathcal{T}$ tel que $P(A) = 0$ est dit *négligeable* ou *quasi-impossible*.
Tout événement $A \in \mathcal{T}$ tel que $A \neq \Omega$ et $P(A) = 1$ est dit *presque-certain* ou *quasi-certain*.
Une propriété sur un ensemble de probabilité 1 est dite *vérifiée presque sûrement*.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et soit $A \in \mathcal{T}$ un événement fixé tel que $P(A) \neq 0$. L'application $P_A : \mathcal{T} \to [0, +\infty[$ définie par $P_A(B) = \frac{P(B\cap A)}{P(A)}$ est une probabilité sur $(\Omega, \mathcal{T}, P)$ dite *probabilité conditionnelle relative à* A. Pour $B \in \mathcal{T}$, on pose $P(B|A) := P_A(B)$. On appelle $P(B|A)$ la *probabilité de B sachant A*.

Propriétés :

(1) Pour $A, B \in \mathcal{T}$ on a : $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$.
(2) Pour $A, B \in \mathcal{T}$ tels que $P(A)P(B) \neq 0$ on a : $P(A|B) = \frac{P(A)P(B|A)}{P(B)}$.

Deux événements $A, B \in \mathcal{T}$ sont dits *indépendants* si $P(A \cap B) = P(A)P(B)$.
Propriétés :

(1) $A, B \in \mathcal{T}$ sont indépendants $\Leftrightarrow P(B|A) = P(B)$
(2) Si $A, B \in \mathcal{T}$ sont incompatibles et non négligeables, alors A et B ne sont pas indépendants
(3) Si $A, B \in \mathcal{T}$ sont indépendants, alors les événements A et $\Omega \setminus B$, les événements $\Omega \setminus A$ et B, et les événements $\Omega \setminus A$ et $\Omega \setminus B$ sont également indépendants.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une *variable aléatoire réelle* sur Ω est une application $X : \Omega \to \mathbb{R}$ telle que pour tout $x \in \mathbb{R}$ on a $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{T}$.
Si Ω est fini ou dénombrable, alors $\mathcal{T} = \mathcal{P}(\Omega)$, d'où $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{T} = \mathcal{P}(\Omega)$ pour tout $x \in \mathbb{R}$. Ainsi, dans ce cas, une variable aléatoire réelle est simplement une application $X : \Omega \to \mathbb{R}$.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et soit $X : \Omega \to \mathbb{R}$ une variable aléatoire réelle. On note :

$$P([X \leq x]) = P\big(\{\omega \in \Omega : X(\omega) \leq x\}\big)\,.$$
$$P([X = x]) = P\big(\{\omega \in \Omega : X(\omega) = x\}\big)\,.$$
$$P([a < X < b]) = P\big(\{\omega \in \Omega : a < X(\omega) < b\}\big)\,.$$

etc.

La *fonction de répartition* d'une variable aléatoire réelle X est l'application $F_X : \mathbb{R} \to \mathbb{R}$ définie par $F_X(x) = P([X \leq x])$.
Propriétés : F_X est une fonction croissante et satisfait : $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$

Variables discrètes : une variable aléatoire réelle X est dite *discrète* si elle ne peut prendre qu'un nombre fini ou dénombrable de valeurs réelles $x_1, x_2, \dots$. L'ensemble de ces valeurs est dit l'*ensemble de définition de* X, noté $X(\Omega)$. La *loi de probabilité* de la variable discrète X est la fonction $P_X : X(\Omega) \to \mathbb{R}$ définie par $P_X(x) = P([X = x])$. On a $\sum_{x \in X(\Omega)} P_X(x) = 1$.
Une variable aléatoire discrète X est dite une *variable aléatoire discrète finie* si elle ne peut prendre qu'un nombre fini de valeurs ; sinon elle est dite *variable aléatoire discrète infinie.*
Soit X une variable aléatoire discrète finie telle que $X(\Omega) = \{x_1, \dots, x_n\}$ avec $x_1 < x_2 < \cdots < x_n$. La fonction de répartition de X est alors la fonction $F_X : \mathbb{R} \to \mathbb{R}$ en escalier croissante donnée par :

$$F_(x) = \begin{cases} 0 & \text{si } x < x_1 \\ P([X = x_1]) + \cdots + P([X = x_k]) & \text{si } x \in [x_k, x_{k+1}[\\ 1 & \text{si } x \geq x_n \end{cases}$$

Soit X une variable aléatoire discrète finie X telle que $X(\Omega) = \{x_1, \dots, x_n\}$.
Soit $k \in \mathbb{N}$, $k \geq 1$. Le *moment d'ordre* k de X est le réel

$$m_k(X) = \sum_{i=1}^{n} x_i^k P([X = x_i])\,.$$

Soit X est une variable aléatoire discrète infinie X telle que $X(\Omega) = \{x_1, x_2 \dots\}$. On dit que X *admet un moment d'ordre* k si la série $\sum_{i=1}^{+\infty} x_i^k P([X = x_i])$ converge absolument (c'est-à-dire si $\sum_{i=1}^{+\infty} |x_i^k P([X = x_i])|$ converge). Dans ce cas, on définit le *moment d'ordre* k de X par

$$m_k(X) = \sum_{i=1}^{+\infty} x_i^k P([X = x_i])\,.$$

Soit X une variable aléatoire discrète finie ou bien une une variable aléatoire discrète infinie qui admet moment d'ordre 1. L'*espérance de* X, notée $E(X)$, est $E(X) = m_1(X)$.
Soit X une variable aléatoire discrète finie ou bien une une variable aléatoire discrète infinie qui admet moment d'ordre 1 et telle que $X - E(X)$ admet moment d'ordre 1. La *variance de* X, notée $V(X)$, est alors définie par $V(X) = m_2(X - E(X))$.
Théorème de Koenig-Huygens : une variable alétoire discrète admet une variance si et seulement si X admet un moment d'ordre 2. Dans ce cas, on a : $V(X) = E(X^2) - \big(E(x)\big)^2$.