Chapitre 2
VARIABLES ALEATOIRES REELLES

2.1 NOTIONS GENERALES

2.1.1 Variable aléatoire réelle (v.a.r.)

Dans une expérience aléatoire modélisée par l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ on s'intéresse souvent davantage à une quantité dépendant du résultat ω qu'à ω lui-même. Exemples :

– dans une course de chevaux, le gain (ou la perte...) plutôt que les numéros des chevaux
– dans un problème de trafic, le nombre de connexions dans un intervalle de temps donné plutôt que leurs instants précis .

Lorsqu'on creuse cette idée, on est conduit à la

Définition

Une **variable aléatoire réelle** (en abrégé v.a.r.) est une application $X : \Omega \to \mathbb{R}$ telle que l'image réciproque par X de tout intervalle soit un évènement('question pertinente').

Mais les intervalles de $\mathbb{R}$ engendrant la tribu borélienne $\mathcal{B}(\mathbb{R})$, il est équivalent d'écrire

$$\boxed{\forall B \in \mathcal{B}(\mathbb{R}) \quad X^{-1}(B) \in \mathcal{A}}$$

On retombe sur une notion vue en Calcul Intégral : X est une application *mesurable* de $(\Omega, \mathcal{A}, \mathbb{P})$ vers $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On sait que ces applications ont d'excellentes propriétés de stabilité par les opérations usuelles sur les applications : opérations algébriques (somme,produit,...), sup, inf, et lim.

Exemples

1. variable indicatrice $X = \mathbb{I}_A$ d'un évènement A

$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

2. variable aléatoire binômiale
3. variable aléatoire uniforme sur $]0, 1[$

Ce dernier exemple est essentiel dans les questions de simulation de variables aléatoires (voir Travaux Pratiques) : les logiciels de calcul génèrent des nombres (pseudo)-aléatoires «au hasard »entre 0 et 1, censés simuler le comportement d'une telle v.a.r.

2.1.2 Loi de probabilité d'une v.a.r.

Comme le montrent les exemples précédents, si on ne s'intéresse qu'aux évènements liés à la variable aléatoire X, alors l'espace probabilisé initial $(\Omega, \mathcal{A}, \mathbb{P})$ peut être remplacé par un autre mieux adapté $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$, dans lequel $\mathbb{P}_X$ est définie par :

$$\boxed{\forall B \in \mathcal{B}(\mathbb{R}) \quad \mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}([X \in B])}$$

$\mathbb{P}_X$ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, c'est la **loi** ou **distribution de probabilité** de X.
Interprétation de cette formule : on peut tirer au sort un élément $\omega \in \Omega$ selon la règle donnée par $\mathbb{P}$, puis prendre son image par X. Il revient au même de tirer directement au sort un $x \in \mathbb{R}$ selon la règle de tirage au sort donnée par $\mathbb{P}_X$.
Il existe des lois de probabilité compliquées. Dans la suite, on se limitera à deux types simples : les lois **discrètes** et les lois **à densité**.

2.1.3 v.a.r. discrètes

Définition

La v.a.r. X est **discrète** si $X(\Omega)$ est un sous-ensemble fini ou dénombrable de $\mathbb{R}$. P_X est alors une mesure de probabilité discrète, combinaison de masses de Dirac.
Si $X(\Omega) = \{x_1, x_2, \ldots\}$, $\mathbb{P}_X$ est determinée par les nombres $p_k = \mathbb{P}_X(\{x_k\}) = \mathbb{P}([X = x_k])$ et

$$\boxed{\forall B \in \mathcal{B}(\mathbb{R}) \quad \mathbb{P}_X(B) = \mathbb{P}([X \in B]) = \sum_{x_k \in B} p_k}$$

On doit évidemment avoir la relation

$$\boxed{1 = \sum_{x_k \in \mathbb{R}} p_k = \sum_{x_k \in X(\Omega)} p_k}$$

Principales lois discrètes
- loi de Dirac δ_{x_0}
- loi de Bernoulli $\mathcal{B}(1, p)$
- loi binômiale $\mathcal{B}(n, p)$
- loi géométrique $\mathcal{G}(p)$
- loi de Poisson $\mathcal{P}(\lambda)$

(voir tables et figures après le chapitre 3)

2.1.4 v.a.r. à densité

Définition

La v.a.r. X est **à densité** s'il existe une fonction $f_X : \mathbb{R} \to \bar{\mathbb{R}}_+$ borélienne telle que

$$\boxed{\forall B \in \mathcal{B}(\mathbb{R}) \quad \mathbb{P}_X(B) = \mathbb{P}([X \in B]) = \int_B f_X(x)dx}$$

La loi de probabilité $\mathbb{P}_X$ est une mesure de densité f_X par rapport à la mesure de Lebesgue sur $\mathbb{R}$.

On a bien sûr

$$\boxed{1 = \int_{\mathbb{R}} f_X(x)dx}$$

Principales lois à densité

- densité uniforme $\mathcal{U}([a,b])$
- densité exponentielle $\mathcal{E}(\lambda)$
- densité gamma $\gamma(a,\lambda)$
- densité gaussienne $\mathcal{N}(m,\sigma^2)$
- densité de Cauchy $\mathcal{C}(c)$

(voir tables et figures après le chapitre 3)

2.1.5 Fonction de répartition (F.R.)

Il existe une autre façon de caractériser la loi de probabilité d'une v.a.r., indépendamment du type de cette loi. Elle repose sur le fait que les intervalles de la forme $]-\infty, x]$ engendrent la tribu $\mathcal{B}(\mathbb{R})$.

Définition

L'application $F_X : \mathbb{R} \to [0,1]$ telle que

$$\boxed{\forall x \in \mathbb{R} \quad F_X(x) = \mathbb{P}([X \leq x])}$$

s'appelle la **fonction de répartition** de X.

- F_X est croissante, continue à droite, et possède une limite à gauche (finie) en chaque point.
- $\lim\limits_{x \to -\infty} F_X(x) = 0$
- $\lim\limits_{x \to +\infty} F_X(x) = 1$
- $\boxed{\mathbb{P}([a < X \leq b]) = F_X(b) - F_X(a)}$ pour $a < b$
- $[F_X]' = \mathbb{P}_X$ (au sens des distributions)

• si X est discrète, F_X est une fonction en escalier, et inversement.

• si X est à densité, F_X est continue, dérivable presque partout avec $F'_X = f_X$ p.p. et

$$\boxed{\forall x \in \mathbb{R} \quad F(x) = \int_{-\infty}^{x} f_X(t)dt}$$

2.2 ESPERANCE MATHEMATIQUE

Soit X une v.a.r. sur $(\Omega, \mathcal{A}, \mathbb{P})$. On cherche à donner un sens à la notion de «valeur moyenne »de X.

Définition

On appelle **espérance mathématique** de X le nombre réel, s'il existe

$$\boxed{\mathbb{E}(X) = \int_{\Omega} X(\omega) d\mathbb{P}(\omega)}$$

NB : cette intégrale n'existe que si d'abord $\int_{\Omega} |X(\omega)| d\mathbb{P}(\omega) < +\infty$

Théorème de transfert : si $\mathbb{E}(X)$ existe

$$\boxed{\mathbb{E}(X) = \int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}} x d\mathbb{P}_X(x)}$$

• si X est discrète, $\boxed{\mathbb{E}(X) = \sum_k x_k p_k}$

• si X est à densité, $\boxed{\mathbb{E}(X) = \int_{\mathbb{R}} x f_X(x) dx}$

▶ Exemples

On a (cf cours de Calcul Intégral) les

Propriétés de l'espérance mathématique

- $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
- $\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$
- si $X \geq 0$ alors $\mathbb{E}(X) \geq 0$
- si $X \geq Y$ alors $\mathbb{E}(X) \geq \mathbb{E}(Y)$
- $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$
- $\mathbb{E}(|X|) = 0 \Leftrightarrow X = 0 \quad p.s.$

Très souvent on aura à considérer le cas suivant : une v.a.r. $Y = \varphi(X), \varphi : \mathbb{R} \to \mathbb{R}$, donnée comme fonction d'une autre v.a.r. X. Pour calculer $\mathbb{E}(Y)$ on peut utiliser le

Théorème de transfert généralisé :

$$\boxed{\mathbb{E}(Y) = \mathbb{E}(\varphi(X)) = \int_{\Omega} \varphi(X(\omega))d\mathbb{P}(\omega) = \int_{\mathbb{R}} \varphi(x)d\mathbb{P}_X(x)}$$

sous réserve d'existence.

2.3 VARIANCE

Soit X une v.a.r. sur $(\Omega, \mathcal{A}, \mathbb{P})$, ayant une espérance $\mathbb{E}(X)$. On cherche à donner un sens à la notion de $\ll$variation moyenne $\gg$de X autour de $\mathbb{E}(X)$.

Définition

On appelle **variance** de X le nombre réel positif, s'il existe

$$\boxed{\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \int_\Omega (X(\omega) - \mathbb{E}(X))^2 d\mathbb{P}(\omega)}$$

Propriétés de la variance

- $\boxed{\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2}$
- $\mathbb{V}(X) \geq 0$
- $\mathbb{V}(X) = 0 \Leftrightarrow X = \mathbb{E}(X) \quad p.s.$
- $\mathbb{V}(\alpha X + \beta) = \alpha^2 \mathbb{V}(X)$
- $\boxed{\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X,Y)}$

La **covariance** de X et Y étant :
$\boxed{Cov(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))}$ ou encore $\boxed{Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}$

NB :**en général** $\mathbb{V}(X+Y) \neq \mathbb{V}(X) + \mathbb{V}(Y)$

2.4 V.A.R. INDEPENDANTES

Définition

Deux v.a.r. X_1 et X_2 définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sont $\mathbb{P}$-**indépendantes** si tout évènement relatif à X_1 est $\mathbb{P}$-indépendant de tout évènement relatif à X_2.Formule :$\forall A_1, A_2 \in \mathcal{B}(\mathbb{R})$

$$\boxed{\mathbb{P}([X_1 \in A_1] \cap [X_2 \in A_2]) = \mathbb{P}([X_1 \in A_1])\mathbb{P}([X_2 \in A_2])}$$

On peut prouver que cette propriété possède une formulation équivalente en termes d'espérances :X_1 et X_2 sont $\mathbb{P}$-**indépendantes** si et seulement si

$$\boxed{\mathbb{E}(h_1(X_1)h_2(X_2)) = \mathbb{E}(h_1(X_1))\mathbb{E}(h_2(X_2))}$$

pour **toutes** les fonctions $h_1, h_2 : \mathbb{R} \to \mathbb{R}$ telles que les espérances existent.

• Application : si X et Y sont indépendantes et ont une variance $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

• Généralisation à une famille de v.a.r. $(X_i)_{i \in I}$:

Les X_i sont **mutuellement** $\mathbb{P}$-**indépendantes** si pour tout sous-ensemble FINI J de I

$$\boxed{\mathbb{P}(\bigcap_{i \in J} [X_i \in A_i]) = \prod_{i \in J} \mathbb{P}([X_i \in A_i])}$$

ATTENTION !

- indépendance deux à deux n'implique pas indépendance mutuelle.
- ça ne dépend pas que des X_i, mais aussi de $\mathbb{P}$

2.5 CALCULS DE LOIS DE PROBABILITE

- **méthodes directes**
- **méthode de la fonction de répartition**
  On cherche à déterminer la fonction F_Y définie sur $\mathbb{R}$ par : $F_Y(y) = \mathbb{P}[Y \leq y]$
- **utilisation de l'espérance**
  (ou méthode de la fonction muette)
  • Elle est basée sur le résultat suivant de Calcul Intégral : égalité de deux mesures bornées

$$\boxed{\mu = \nu \iff \int h d\mu = \int h d\nu}$$

pour toute fonction-test $h \in \mathcal{C}_b(\mathbb{R})$

• En pratique, si $Y = \varphi(X)$ on cherche à calculer $\mathbb{E}(h(Y)) = \int h(y) d\mathbb{P}_Y(y)$ en utilisant le théorème de transfert.
(Dans ce calcul la fonction h continue bornée quelconque ne sert que de «révélateur»pour $\mathbb{P}_Y$).

2.6 SIMULATION DE VARIABLES ALEATOIRES REELLES

- **Simulation de la loi uniforme sur** $]0,1[$
  Nombreuses méthodes.
  Une méthode classique : générateur congruentiel
  $$x_{n+1} = ax_n + b \qquad modulo\ m$$
  La suite $u_n = \frac{x_n}{m}$, pour des valeurs bien choisies de a, b et m, est pseudo-aléatoire.
  Exemple : $a = 7^5 = 16807$, $b = 0$,
  $m = 2^{31} - 1 = 2147483647$ (Gentle) donne une suite de période maximale m.
  Sous SCILAB, la commande *rand* utilise
  $a = 843314861$, $b = 453816693$, $m = 2^{31}$. Le terme initial x_0 (le *germe*) est fixé par défaut à 0, le premier appel de *rand* renvoie toujours
  $u_1 = 453816693/2^{31} \approx 0.2113249$

- **Théorème fondamental de la simulation**
  Toute v.a.r. X peut être simulée sous la forme
  $X = \phi(U)$ avec U de loi uniforme sur $]0,1[$ et ϕ borélienne continue p.p.

- **Méthode de la fonction de répartition**
  Pour toute fonction de répartition F, et pour toute v.a.r. U de loi uniforme sur $]0,1[$, la v.a.r.
  $X = F^{-1}(U)$ a pour fonction de répartition F.
  si F n'est pas bijective, $F^{-1}(u) = \inf\{t \in \mathbb{R}, F(t) \geq u\}$