INSA

Signaux aléatoires

Notes de cours
Version 2.0

D. Arzelier

Avertissement : Ce document est constitué de notes de cours et ne prétend donc ni à l'exhaustivité ni à l'originalité. Ces notes doivent en effet beaucoup aux emprunts faits aux ouvrages référencés en bibliographie.

Notations

- $\mathbb{R}$: corps des nombres réels.

- A' : matrice transposée de la matrice A.

- $A > \mathbf{0}$: A matrice définie positive.

- $A \geq \mathbf{0}$: A matrice semi-définie positive.

- $\| \bullet \|$: norme Euclidienne pour un vecteur et induite par la norme Euclidienne pour une matrice.

- $P[\bullet]$: probabilité simple.

- $P[\bullet/\bullet]$: probabilité conditionnelle.

- v.a. : variable aléatoire.

- V.A. : vecteur aléatoire.

- $E[\bullet]$: espérance mathématique.

- $E[x/y]$: espérance conditionnelle de x sachant y.

- $p_x(\alpha)$: densité de probabilité de la v.a. x.

- $p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n)$: densité de probabilité conjointe.

- $p_{x/y}(\alpha/\beta)$: densité de probabilité conditionnelle.

- $Z \sim \mathcal{N}(m_Z, Q_Z)$: vecteur gaussien de moyenne m_Z et de matrice de covariance Q_Z.

- $R(t,\tau)$: matrice d'autocorrélation.

- $P(t,\tau)$: matrice d'autocovariance.

- δ_{kl} : symbole de Kronecker.
- $\delta(t)$: impulsion de Dirac.
- $\mathbf{1}_n$: matrice identité de dimension n.
- $\mathbf{0}_{n\times m}$: matrice nulle de dimensions $n \times m$.

Table des matières

Introduction générale

La nécessité d'introduire la notion de signal aléatoire ou de processus stochastique provient de la constatation facile à établir que le cadre défini par les signaux déterministes est insuffisant pour décrire et étudier l'ensemble des problèmes rencontré en Automatique, en théorie de l'estimation et en traitement du signal. En effet, partant du fait qu'un signal "naturel" est plus ou moins imprévisible et à ce titre transporte de l'information, la limitation du modèle déterministe des signaux apparaît clairement sachant qu'il ne comprend aucun élément d'incertitude et ne transporte donc aucune information. Si l'on ajoute qu'en tout état de cause aucun modèle mathématique ne peut prétendre représenter exactement la réalité, qu'il est également nécessaire de prévoir et de modéliser les perturbations non prévisibles de manière déterministe et que tout système de mesure fournit des données incomplètes et bruitées, il semble alors naturel de souhaiter disposer d'un modèle complémentaire reflétant l'aspect aléatoire entrant dans toute représentation et modélisation des signaux et systèmes.

Dans le cadre de ce cours, nous nous limiterons volontairement à l'étude de la modélisation des signaux aléatoires et à leur transmission par des systèmes linéaires. Afin de mieux circonscrire mathématiquement l'objet en question, nous nous plaçons volontairement dans un cadre probabiliste. Nous utiliserons donc de manière privilégiée le cadre mathématique de la théorie axiomatique des probabilités développée par Kolmogorov. Il sera donc supposé que nous disposons de suffisamment de connaissance a priori sur le signal pour permettre une description appropriée de ses propriétés moyennes. La majeure partie du cours sera donc consacrée à la définition des caractéristiques statistiques essentielles ainsi qu'aux caractéristiques temporelles. Les propriétés de ces caractéristiques et les rapports qu'elles entretiennent permettront de définir les notions essentielles d'ergodicité et de stationnarité. Un chapître sera également consacré à l'étude de signaux aléatoires remarquables. Finalement, la transmission des signaux aléatoires par les systèmes linéaires sera étudiée en détails. Dans tout ce qui précède, seul le terme de signal aléatoire a été employé. Le terme de processus stochastique est également fréquent dans la littérature. Un signal aléatoire est l'image d'une grandeur associée à un phénomène physique aléatoire généralement appelé processus stochastique. Dans la suite, ces deux termes seront confondus et utilisés indifféremment.

Chapitre 1

Introduction à l'étude des processus stochastiques

1.1 Introduction et définition

La notion de processus stochastique est une notion fondamentale qui va servir d'une part à la définition de modèles pour les perturbations, les phénomènes de bruit mais également à la définition des équations aux différences et équations différentielles stochastiques. Les solutions de ces équations stochastiques définiront également des processus stochastiques possédant des propriétés particulières.

Définition 1.1.1 *: processus stochastique*
Un processus stochastique *(scalaire ou vectoriel) est défini comme une famille de variables ou vecteurs aléatoires indexée par un ensemble de paramètres $t \in T$, que l'on va considérer dans la suite du cours comme étant le temps. La notation est $\{x_t(\omega) \mid t \in T\}$. T peut être un ensemble discret ou continu.*

C'est donc une fonction de deux paramètres : le temps et ω paramètre aléatoire (résultat d'une expérience).

- Pour chaque t, $x_t(\bullet)$ est **une variable aléatoire** égale à l'état du processus considéré à l'instant t.

- Pour ω fixé, $x_\bullet(\omega)$ est **une réalisation du processus** qui est une fonction du temps.

- Pour t et ω fixés, $x_t(\omega)$ est **un nombre**.

Si la variable (vecteur) aléatoire prend ses valeurs dans un espace discret, on dira que c'est **une chaîne**, ou **un processus stochastique à état discret**, qui peut être défini en temps discret ou en temps continu. Dans le cas contraire, le processus sera défini comme **un processus stochastique à état continu** en temps discret ou continu. Dans le cadre de ce cours, nous nous intéresserons principalement aux processus à état continu qui, en temps discret sont appelés **séquences aléatoires** et dans le cas continu plus simplement **processus stochastiques** si aucune confusion n'est possible. La notation est alors simplifiée en $x(t)$ pour un processus en temps continu et x_t pour une séquence.

<u>Exemple :</u>

Soit un satellite en orbite périodique autour de la terre pour lequel on souhaite connaître les paramètres orbitaux afin d'analyser et de corriger d'éventuelles phases de dérive. On dispose pour cela de plusieurs instruments de mesure fonctionnant en parallèle. La restitution de l'altitude $x(t)$ du satellite par l'ensemble de ses capteurs définit un signal aléatoire analogique en temps continu.

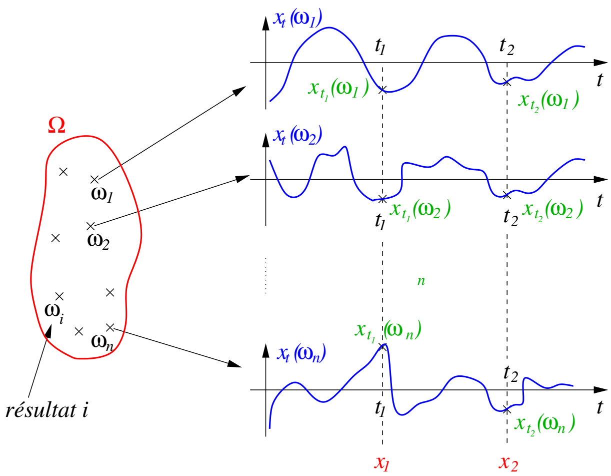


FIG. 1.1 – *restitution d'un paramètre orbital d'un satellite*

1.2 Caractérisations probabilistes des processus stochastiques

1.2.1 Loi de distribution

Soit $\{x_t(\omega) \mid t \in T\}$ un processus stochastique. T peut être continu ou discret. Pour n'importe quelle partition $\{t_1, \cdots, t_n\}$ de T, qui peut être obtenue par discrétisation dans le cas continu, la loi de distribution conjointe des variables (vecteurs) aléatoires $x_{t_1}, \cdots, x_{t_n}$ est appelée **loi de distribution finie** du processus. Le processus stochastique peut être caractérisé par la définition de la loi de distribution finie pour tous les ensembles finis $\{t_i\} \in T$, donnée de manière équivalente par :

- La fonction de répartition conjointe :

$$F_{x(t_1),\cdots,x(t_n)}(\alpha_1, \cdots, \alpha_n)$$

- La fonction densité de probabilité :

$$p_{x(t_1),\cdots,x(t_n)}(\alpha_1, \cdots, \alpha_n)$$

pour toutes les partitions $\{t_i\} \in T$.
Spécifier l'une des deux fonctions précédentes $\forall\ \{t_i\} \in T$, consiste à définir **la loi de probabilité** du processus stochastique.

Exemple :
Un processus stochastique $x(t)$ peut être défini par :

$$x(t) = a + bt$$

où $a,\ b$ sont des variables aléatoires de distribution connue. On suppose de plus que $a,\ b$ sont deux variables aléatoires conjointement normalement distribuées. Puisque l'on peut écrire :

$$\begin{bmatrix} x(t_1) \\ \vdots \\ x(t_n) \end{bmatrix} = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

On peut dire que le vecteur aléatoire $[x(t_1), \cdots, x(t_n)]'$ est un vecteur gaussien dont la moyenne et la covariance peuvent être données à partir de la moyenne et de la covariance de $[a \ \ b]'$.

1.2.2 Processus du second ordre

Soit un processus stochastique $\{x_t(\omega) \mid t \in T\}$. Bien que non suffisantes dans le cas le plus général, la spécification des densités de probabilité du **premier ordre** et du **deuxième ordre** est souvent intéressante :

- Densité du premier ordre :

$$\boxed{p_{x(t)}(\alpha) = p(\alpha, t)}$$

- Densité du deuxième ordre :

$$\boxed{p_{x(t),x(\tau)}(\alpha, \beta) = p(\alpha, \beta, t, \tau)}$$

Ces fonctions ne suffisent, en général, pas à la caractérisation complète des processus stochastiques mais il existe certaines classes de processsus (gaussiens - markoviens) pour lesquelles la loi de distribution est entièrement déterminée par la connaissance des fonctions densité de probabilité, au second ordre.

Définition 1.2.1 *: processus du second ordre*
Un processus stochastique caractérisé entièrement par ses lois de distributions au premier et deuxième ordre est appelé **processus du second ordre.**

1.2.3 Caractéristiques statistiques

A partir de ces fonctions densité de probabilité, peuvent être définis les moments du processus stochastique.

Définition 1.2.2 *: moyenne*
Soit un processus stochastique $\{x_t(\omega) \mid t \in T\}$ *noté* $x(t)$*, alors* **la moyenne de** $x(t)$ *est définie par :*

$$\boxed{m_x(t) = E[x(t)] = \int_{-\infty}^{+\infty} \alpha p(\alpha, t) d\alpha}$$

C'est une fonction du paramètre t.

Définition 1.2.3 *: matrice d'autocorrélation et d'autocovariance*
On peut définir **la matrice d'autocorrélation**

$$\boxed{R(t, \tau) = E[x(t)x(\tau)']}$$

dont les éléments sont :

$$r_{ij}(t,\tau) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \alpha_i \beta_j p[\alpha,t,\beta,\tau] d\alpha d\beta$$

et la matrice d'autocovariance :

$$\boxed{P(t,\tau) = E[(x(t)-m_x(t))(x(\tau)-m_x(\tau))']}$$

On vérifie aisément les propriétés suivantes :

Propriété 1.2.1 *:*

- $P(t,\tau) = P(\tau,t) \quad \forall\ t,\tau$
- $R(t,\tau) = R(\tau,t) \quad \forall\ t,\tau$
- $R(t,t) \geq 0 \quad \forall\ t$
- $Q(t) = P(t,t) \geq 0 \quad \forall\ t$
- $P(t,\tau) = R(t,\tau) - m_x(t)m_x(\tau)' \quad \forall\ t,\tau$

Le préfixe "auto", souvent omis, n'est utilisé que lorsqu'il peut y avoir confusion avec l'intercorrélation et l'intercovariance, définies par rapport à deux processus stochastiques distincts $x(t)$ et $y(t)$ par :

$$\Gamma_{xy}(t,\tau) = E[x(t)y(\tau)']$$

pour l'intercorrélation et :

$$C_{xy}(t,\tau) = E[(x(t)-m_x(t))(y(\tau)-m_y(\tau))']$$

pour l'intercovariance.
La moyenne et la matrice de corrélation sont généralement plus faciles à manipuler que la loi de probabilité et n'en donne pas moins de nombreuses informations sur le processus.

Exemple :

Soit le processus stochastique :

$$x(t) = a + bt$$

de l'exemple précédent. Les moments de ce processus sont donnés par :

$$m_x(t) = E[a] + tE[b] = m_a + tm_b$$

$$r(t,\tau) = E[a^2] + (t+\tau)E[ab] + t\tau E[b^2]$$

$$p(t,\tau) = var(a) + cov(a,b)(t+\tau) + var(b)t\tau$$

On peut noter que ces moments ne dépendent que des moments de a et b.

1.3 Caractéristiques importantes

La stationnarité fixe la propriété d'invariance par rapport au temps t des lois de probabilité qui caractérisent le processus stochastique. Elle est en pratique très importante et peut être définie de différentes façons.

1.3.1 Stationnarité au sens strict

Définition 1.3.1 *: stationnarité au sens strict*

$x(t)$ est dit **stationnaire au sens strict**, *sur T intervalle de définition, si pour toute partition $\{t_1, t_2, \cdots, t_n\}$ de T :*

$$\forall\ n,\ \forall\ \theta,\quad p_{x(t_1),\cdots,x(t_n)}(\alpha_1,\cdots,\alpha_n) = p_{x(t_1+\theta),\cdots,x(t_n+\theta)}(\alpha_1,\cdots,\alpha_n)$$

1.3.2 Stationnarité au sens large

La stationnarité au sens large est une notion plus pratique qui est définie à l'ordre deux et généralement à partir des deux premiers moments du processus.

Définition 1.3.2 *: stationnarité au sens large*

$x(t)$ est un processus stochastique **stationnaire au sens large** *si*

$$\begin{cases} m_x(t) = m_x \quad \text{(indépendant du temps)} \\ \\ P(t,\tau) = P(t-\tau) \end{cases}$$

Dans la condition précédente, la covariance peut être de façon équivalente remplacée par la corrélation.

Exemple :
Le processus

$$x(t) = acos(2\pi t) + bsin(2\pi t)$$

où a, b sont des variables aléatoires non corrélées de moyenne nulle et de variance unité est stationnaire au sens large :

$$m(t) = 0$$

$$r(t+\tau, t) = cos(2\pi\tau)$$

Nota : il est à noter qu'un processus stationnaire au sens strict est stationnaire au sens large mais la réciproque n'est pas nécessairement vraie.

1.3.3 Ergodicité

On a défini précédemment des moyennes statistiques à des instants donnés. Il est également intéressant de définir des moyennes temporelles.

Définition 1.3.3 *: moyenne temporelle*

On définit la **moyenne temporelle** *d'un échantillon du processus par la relation :*

$$< x(t) >= \lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} x(t)dt$$

lorsque cette limite existe.

Nota : d'autres moyennes temporelles peuvent être définies. Par exemple :

$$< x^2(t) >= \lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} x^2(t)dt$$

La limite ainsi définie peut ne pas exister pour certains échantillons ou dépendre de l'échantillon mais elle ne dépend pas du temps. Toutefois, il existe des cas où la moyenne temporelle et la moyenne statistique sont égales.

Définition 1.3.4 *: ergodicité*

Un processus stochastique $x(t)$ est **ergodique** *si toutes les moyennes temporelles existent et ont même valeur quelque soit l'échantillon.*

$$\forall\ f, \quad \lim_{t_1\to\infty} [\frac{1}{t_1 - t_0} \int_{t_0}^{t_1} f[x(t)]dt] = E[f(x(t))] \tag{1.1}$$

Cette notion d'ergodicité est très importante du fait que pratiquement, pour évaluer les moyennes statistiques, l'on ne dispose généralement que d'un échantillon sur lequel on fait une moyenne temporelle. Ce procédé n'a de valeur que si le processus stochastique est stationnaire et ergodique.

1.3.4 La densité spectrale de puissance

En théorie du signal, les spectres sont associés à la transformée de Fourier. Dans le cas de signaux déterministes, c'est une façon de représenter les signaux comme une superposition d'exponentielles. Pour les signaux aléatoires, il s'agit principalement d'étudier la transformée des moyennes.

Définition 1.3.5 *: puissance*

On définit **la puissance moyenne** *d'un processus stochastique stationnaire au sens large par :*

- *Cas continu* $x(t)$ *:*

$$\boxed{p = E[x'(t)x(t)]}$$

- *Cas discret* x_N *:*

$$\boxed{p = \lim_{N \to +\infty} \frac{1}{2N+1} \sum_{-N}^{N} x_N^2}$$

Propriété 1.3.1 *:*

$$p = Trace[R_x(0)]$$

Soit un processus stochastique $x(t)$ stationnaire au sens large et possédant une matrice de corrélation $R_x(t-\tau) = R_x(\tau)$.

Définition 1.3.6 *: densité spectrale de puissance*

La densité spectrale de puissance $\Psi_x(\omega)$ *du processus stochastique* $x(t)$ *est définie comme la transformation de Fourier, si elle existe, de la matrice de corrélation* $R_x(\tau)$ *:*

$$\boxed{\Psi_x(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_x(\tau) d\tau}$$

Réciproquement,

$$\boxed{R_x(\tau) = \tfrac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} \Psi_x(\omega) d\omega}$$

Les deux relations précédentes sont connues sous le nom de formules de Wiener-Khinchin. $\Psi_x(\omega)$ est une matrice complexe possédant les propriétés :

Propriété 1.3.2 *:*

- $\Psi_x(-\omega) = \Psi_x(\omega)' \;\; \forall \; \omega$
- $\Psi_x^*(\omega) = \Psi_x(\omega) \;\; \forall \; \omega$

- $\Psi_x(\omega) \geq 0 \quad \forall\ \omega$

- $\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} \Psi_x(\omega) d\omega = R_x(0)$

Exemple :

Soit le processus stochastique $v(t)$ stationnaire au sens large de matrice de corrélation :

$$R_v(t-\tau) = \sigma^2 exp\left[-\frac{|t-\tau|}{\theta}\right] \qquad \theta > 0$$

Il est facile de montrer par le calcul que :

$$\Psi_v(\omega) = \frac{2\sigma^2\theta}{1+\omega^2\theta^2}$$

Nous donnons quelques exemples de densité spectrale de puissance de pro-

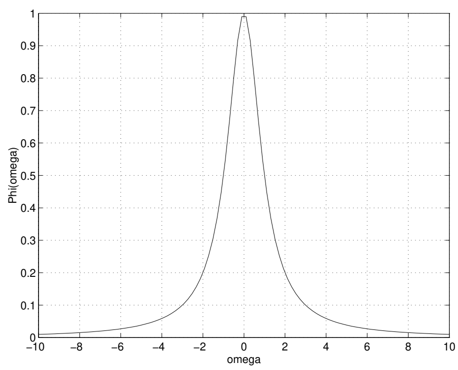


FIG. 1.2 – *Spectre de $v(t)$*

cessus stochatiques définis par leur matrice de corrélation.

$R_x(\tau)$	$\Psi_x(\omega)$
$\delta(\tau)$	1
$e^{j\beta\tau}$	$2\pi(\delta(\omega-\beta)$
$e^{-\alpha\|\tau\|}$	$\dfrac{2\alpha}{\alpha^2+\omega^2}$
$e^{-\alpha\|\tau\|}cos(\beta\tau)$	$\dfrac{\alpha}{\alpha^2+(\omega-\beta)^2}+\dfrac{\alpha}{\alpha^2+(\omega+\beta)^2}$
1	$2\pi\delta(\omega)$
$cos(\beta\tau)$	$\pi\delta(\omega+\beta)+\pi\delta(\omega-\beta)$

L'introduction de la densité spectrale de puissance peut se justifier à travers le résultat suivant.

Théorème 1.3.1 *:*
Soit le processus stochastique $x(t)$, alors pour toute matrice symétrique $W(t)$:

$$E[x(t)'W(t)x(t)] = Trace[W(t)R_x(t,t)]$$

Si, de plus, $x(t)$ est stationnaire au sens large de moyenne nulle et W est constante alors :

$$E[x(t)'Wx(t)] = Trace[WR_x(0)] = \frac{1}{2\pi}Trace[\int_{-\infty}^{\infty} W\Psi_x(\omega)d\omega]$$

Ce résultat permet de donner une interprétation au terme densité spectrale de puissance. En effet, la "puissance" totale $E[x(t)'Wx(t)]$ d'un processus stochastique de moyenne nulle et stationnaire au sens large est obtenue par intégration de $Trace[W\Psi_x(\omega)]$ sur toutes les fréquences. Ainsi,

$Trace[W\Psi_x(\omega)]$ peut être considérée comme une mesure de la densité de puissance à la pulsation ω. La matrice W est une matrice de pondération qui détermine la contribution des divers composantes de $x(t)$ à la puissance. La densité spectrale de puissance représente la répartition harmonique de la puissance moyenne p de $x(t)$.

1.3.5 Indépendance et décorrélation

Il est possible d'étendre les notions d'indépendance et de décorrélation aux processus stochastiques.

Définition 1.3.7 *: indépendance*
Soient $v_1(t)$, $v_2(t)$ deux processus stochastiques. Alors $v_1(t)$, $v_2(t)$ sont dits **indépendants** *si $\{v_1(t_1), v_1(t_2), \cdots, v_1(t_l)\}$ et $\{v_2(t'_1), v_2(t'_2), \cdots, v_2(t'_m)\}$ sont des ensembles indépendants de vecteurs aléatoires pour tous les $t_1,\ t_2, \cdots, t_l, t'_1, \cdots, t'_m$ et pour tout m et l.*

De même,

Définition 1.3.8 *: décorrélation*
$v_1(t)$, $v_2(t)$ sont dits **décorrélés** *si $v_1(t_1)$ et $v_2(t_2)$ sont des vecteurs aléatoires décorrélés pour tous les t_1, t_2.*

Soient deux processus stochastiques $x(t)$ et $y(t)$.

Définition 1.3.9 *: matrice d'intercorrélation et d'intercovariance*
On définit

- **la matrice d'intercorrélation** *par :*

$$\boxed{R_{xy}(t_1, t_2) = E[x(t_1)y'(t_2)]}$$

- **la matrice d'intercovariance** *par :*

$$\boxed{P_{xy}(t_1, t_2) = R_{xy}(t_1, t_2) - m_x(t_1)m'_y(t_2)}$$

Une notion importante liée à ces concepts est l'orthogonalité.

Définition 1.3.10 *: orthogonalité*
$x(t)$, $y(t)$ sont deux processus stochastiques **orthogonaux** *si*

$$R_{xy}(t_1, t_2) = 0 \qquad \forall\ t_1,\ t_2$$

De plus,

Définition 1.3.11 *: décorrélation*
$x(t)$, $y(t)$ sont deux processus stochastiques **décorrélés** *si*

$$P_{xy}(t_1, t_2) = 0 \qquad \forall\ t_1,\ t_2$$

1.3.6 Densité spectrale croisée

Définition 1.3.12 *:*

La densité spectrale croisée *ou* **interspectre de puissance** *a pour définition :*

$$\Psi_{xy}(\omega) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j\omega\tau} d\tau$$

aussi bien que :

$$\Psi_{yx}(\omega) = \int_{-\infty}^{\infty} R_{yx}(\tau) e^{-j\omega\tau} d\tau$$

Réciproquement, on aura :

$$R_{xy}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_{xy}(\tau) e^{j\omega\tau} d\omega$$

$$R_{xy}(\omega) = \int_{-\infty}^{\infty} \Psi_{xy}(\tau) e^{j\omega\tau} d\omega$$

Du fait des propriétés de l'intercorrélation, on a :

$$\Psi_{xy}(\omega) = \Psi_{yx}(-\omega) = \Psi^{*}_{xy}(\omega)$$

Chapitre 2

Processus stochastiques remarquables

2.1 Processus gaussien

Les processus gaussiens sont des processus très importants du fait qu'on les rencontre très souvent en pratique et du fait que de nombreux processus physiques sont approximativement gaussiens. Ceci est la conséquence du théorème de la limite centrale. Sans entrer dans les détails, le théorème de la limite centrale établit que si l'on considère la somme :

$$x = x_1 + \cdots + x_n$$

où x_i sont des variables aléatoires indépendantes, alors sous certaines conditions, la distribution de probabilité de x approche la distribution gaussienne quand $n \to \infty$.

Définition 2.1.1 *: processus gaussien*
Le processus stochastique $x(t)$ est **gaussien** *si pour toute partition $\{t_1, t_2, \cdots, t_n\}$ de T, le vecteur des variables aléatoires $[x(t_1), \cdots, x(t_n)]$ est gaussien.*

La loi de distribution d'un vecteur gaussien est entièrement déterminée par la connaissance des moments du premier ordre (moyenne et covariance) qui, eux-mêmes, ne nécessitent pour leur calcul que les fonctions densité de probabilité du premier et du second ordre $p_x(\alpha, t)$, $p_x(\alpha, t, \beta, \tau)$. Il s'agit donc bien d'un processus du second ordre. Noter que la fonction du premier ordre peut être calculée à partir de celle du second ordre par la fonction marginale :

$$p_x(\alpha, t) = \int_{-\infty}^{+\infty} p_x(\alpha, t, \beta, \tau) d\beta$$

Propriété 2.1.1 *:*
Un processus gaussien stationnaire au sens large l'est au sens strict.

Exemple :
Soit le processus défini par :

$$x(t) = a + bt$$

où $a,\ b$ sont des variables aléatoires gaussiennes non corrélées de moyenne nulle et de variance unité. Alors, réécrivant

$$\begin{bmatrix} x(t_1) \\ \vdots \\ x(t_n) \end{bmatrix} = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

on peut trouver que :

$$E[x(t_1), \cdots, x(t_n)]' = 0$$

$$P = \begin{bmatrix} 1 + t_1^2 & \cdots & 1 + t_1 t_n \\ \vdots & \vdots & \vdots \\ 1 + t_n t_1 & \cdots & 1 + t_n^2 \end{bmatrix}$$

Le processus est de plus gaussien.

De par le théorème de la limite centrale dont nous avons parlé en début de section, un processus gaussien est souvent un bon modèle pour le bruit dans les circuits physiques tels que les transmetteurs et récepteurs radios, les radars etc.

2.2 Processus markovien

Cette classe de processus stochastiques est importante du fait qu'elle généralise à l'univers stochastique une propriété fondamentale des équations différentielles ordinaires. En effet, la solution $x(t_2) = g(t_2, x(t_1), t_1)$ de l'équation différentielle $\dot{x}(t) = f(x(t))$ est une fonction de $x(t_1)$ et ne dépend pas de $x(\tau),\ \tau < t_1$.

Définition 2.2.1 *: processus markovien*
Le processus stochastique $x(t)$ est **markovien** *si $\forall\ \{t_1, t_2, \cdots, t_n\}$ de T :*

$$P[x(t_n) < \alpha_n / x(t_{n-1}), \cdots, x(t_1)] = P[x(t_n) < \alpha_n / x(t_{n-1})]$$

soit en termes de fonction densité de probabilité :

$$p_{x(t_n)/x(t_{n-1}),\cdots,x(t_1)}(\alpha_n/\alpha_{n-1}, \cdots, \alpha_1) = p_{x(t_n)/x(t_{n-1})}(\alpha_n/\alpha_{n-1})$$

Toute l'information sur le passé est concentrée dans le dernier état "observé". Un processus de Markov est donc un processus stochastique dont le passé n'a pas d'influence sur le futur si le présent est connu.

Définition 2.2.2 *: densité de probabilité de transition*

$p_{x(t_n)/x(t_{n-1})}(\alpha_n/\alpha_{n-1})$ *est appelée* **fonction densité de probabilité de transition**.

Propriété 2.2.1 *:*

Le processus markovien est un processus du second ordre.

En effet, la fonction densité de probabilité conjointe de $(x(t_1), \cdots, x(t_n))$ s'exprime ainsi :

$$\begin{gathered} p_{x(t_n),x(t_{n-1}),\cdots,x(t_1)}(\alpha_n, \alpha_{n-1}, \cdots, \alpha_1) = p_{x(t_n)/x(t_{n-1})}(\alpha_n/\alpha_{n-1}) \bullet \\ p_{x(t_{n-1})/x(t_{n-2})}(\alpha_{n-1}/\alpha_{n-2}) \cdots p_{x(t_2)/x(t_1)}(\alpha_2/\alpha_1) p_{x(t_1)}(\alpha_1) \end{gathered}$$

qui se détermine donc à partir de la connaissance de $p_{x(t)}(\alpha) = p_x(\alpha, t)$, la densité de probabilité simple et de $p_{x(t)/x(t_0)}(\alpha/x_0)$, la densité de probabilité de transition. On rappelle que :

- $p_x(\alpha, t) = \dfrac{\partial F(\alpha, t)}{\partial \alpha} = \dfrac{\partial P[x(t) \leq \alpha]}{\partial \alpha}$
- $p_{x(t)/x(t_0)}(\alpha/x_0) = p_{x(t)/x(t_0)}(\alpha/x(t_0) = x_0)$
- $p_x(\alpha, t) = \displaystyle\int_{-\infty}^{+\infty} p_x(x_0, t_0) p_{x(t)/x(t_0)}(\alpha/x_0) dx_0$
- $p_{x(t)/x(t_0)} \xrightarrow{t \to t_0} \delta(\alpha - x_0)$

Exemple :

On considère l'équation aux différences :

$$x_{n+1} = x_n + w_n \quad n = 0, 1, \cdots$$

où les variables w_i sont des variables aléatoires indépendantes gaussiennes et la condition initiale x_0 est gaussienne et indépendante de w_i, $\forall\ i$. La séquence $\{x_n\}_n$ est clairement markovienne puisque pour x_n donnée, x_{n+1} dépend uniquement de w_n qui est indépendant de $x_{n-1}, \cdots, x_0$. De plus, elle est également gaussienne puisque combinaison linéaire de variables gaussiennes.

Comme nous le verrons dans la suite de ce cours, les processus de Markov ont une grande importance dans l'étude des équations aux différences stochastiques (systèmes dynamiques en temps discret) et dans l'étude des équations différentielles stochastiques (systèmes dynamiques en temps continu).

2.3 La marche aléatoire

La marche aléatoire est le nom donné à un processus stochastique en temps discret et à état discret que l'on peut décrire de la façon suivante.

On lance une pièce non truquée toutes les T secondes et suivant que l'on obtient face ou pile, on effectue un pas à droite ou à gauche de longueur s. Le processus débute pour $t = 0$ et la position du marcheur aléatoire à l'instant t suit une fonction en escalier où les discontinuités sont aux points $t = nT$.

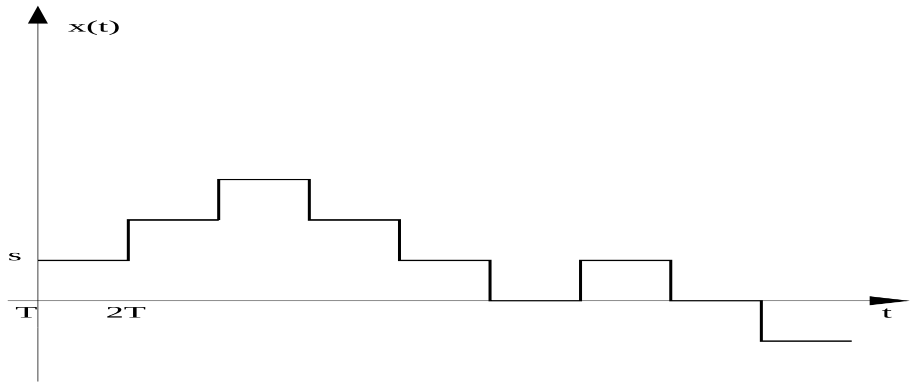


FIG. 2.1 – *Exemple de marche aléatoire*

On suppose maintenant que l'on a effectué n tirages de la pièce, pour lesquels on a observé k fois pile et donc $n - k$ fois face. La position du marcheur aléatoire à l'instant $t = nT$ est :

$$x(nT) = ks - (n-k)s = ms \quad m = 2k - n$$

De plus, la probabilité d'avoir k fois pile dans n tirages est alors :

$$P[x(nT) = ms] = \frac{\mathcal{C}_n^k}{2^n} \qquad k = \frac{m+n}{2}$$

En remarquant que :

$$x(nT) = x_1 + x_2 + \cdots + x_n$$

où x_i est la taille du i^{eme} pas et les x_i sont indépendantes. En prenant la valeur $\pm s$ avec $E[x_i] = 0$, $E[x_i^2] = s^2$, on obtient :

$$\boxed{E[x(nT)] = 0 \qquad E[x^2(nT)] = ns^2}$$

Quand le nombre de tirages augmente, il est possible de montrer, en utilisant le théorème de DeMoivre-Laplace que :

$$\mathcal{C}_n^k\, p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-(k-np)^2/2npq}$$

pour k dans un $\sqrt{npq}$ voisinage de np.

Ici, $p = q = 0.5$ et $m = 2k - n$, ce qui conduit à :

$$P[x(nT) = ms] \simeq \frac{1}{\sqrt{n\pi/2}} e^{-m^2/2n}$$

pour m de l'ordre de $\sqrt{n}$.

2.4 Bruit blanc

Le bruit blanc est un processus stochastique utilisé afin de modéliser les bruits intervenant dans toute modélisation de systèmes dynamiques.

Définition 2.4.1 *: séquence blanche*
Une séquence aléatoire blanche $\{x_n \;,\; n = 1, 2, \cdots\}$ *est une séquence de markov telle que :*

$$p_{x_k/x_l}(\alpha_k/\alpha_l) = p_{x_k}(\alpha_k) \qquad (k > l)$$

Cela signifie que les x_k sont mutuellement indépendants. Connaître la réalisation de x_l n'aide pas à connaître celle de x_k. Une séquence blanche est totalement imprévisible.
Si les x_k sont normalement distribués alors c'est **une séquence blanche gaussienne** définie par sa moyenne et sa covariance :

$$E[x_n] = m_n$$

$$P(n,k) = V_n \delta_{nk} \quad \delta_{nk} = \begin{cases} 1 & n = k \\ 0 & n \neq k \end{cases} \;,\; V_n \geq 0 \text{ Intensité}$$

En général, le bruit du à la superposition d'un grand nombre de faibles effets aléatoires indépendants est gaussien (ceci provenant de l'application du théorème de la limite centrale).

Puisque une séquence blanche gaussienne est un bon modèle pour le bruit affectant un système dynamique décrit par un modèle en temps discret, on souhaite disposer d'un équivalent pour les bruits affectant les modèles dynamiques en temps continu. Dans ce cas, le bruit blanc peut être défini *formellement* par extension du cas discret.

Définition 2.4.2 *: bruit blanc*
$x(t)$ est **un bruit blanc** *s'il vérifie :*

$$p_{x(t)/x(\tau)}(\alpha/\beta) = p_{x(t)}(\alpha) \quad t > \tau$$

Cette définition formelle quoique possédant les propriétés indispensables caractérisant un bruit, n'est pas d'une utilisation très facile.
Nous définissons donc formellement un bruit blanc gaussien par sa moyenne et sa covariance :

Définition 2.4.3 *: bruit blanc gaussien*
Un bruit blanc gaussien *est un processus stochastique gaussien défini par :*

$$E[x(t)] = m_x(t)$$

$$P(t,\tau) = E[(x(t) - m_x(t))(x(\tau) - m_x(\tau))'] = V(t)\delta(t-\tau)$$

où :

- $V(t) \geq 0$ *intensité du bruit blanc.*
- $\delta(\bullet)$ *impulsion de Dirac.*

Dans le cas où l'intensité du bruit blanc est constante, le processus est stationnaire au sens large et il est possible d'introduire sa densité spectrale de puissance qui est donnée par :

$$\Psi_x(\omega) = V$$

Cela montre qu'un bruit blanc gaussien stationnaire a une densité de puissance identique à toutes les fréquences. Cela justifie la dénomination de bruit blanc par analogie avec la lumière blanche. Toutefois, si l'on calcule la puissance totale d'un bruit blanc en utilisant le résultat du théorème 6, nous obtenons une valeur infinie qui montre que ce type de processus n'existe pas dans le monde physique. De plus, du point de vue strictement mathématique, le bruit blanc n'est pas rigoureusement défini. En fait, on peut montrer qu'un bruit blanc gaussien est le processus "dérivé" d'un processus de Wiener bien que le processus de Wiener ne soit pas dérivable au sens classique. Dans le cadre de ce cours, nous nous contenterons de sa définition formelle du fait qu'un traitement rigoureux irait au delà des objectifs de ce cours.

2.5 Bruit blanc à bande limitée

Définition 2.5.1 *:*

Un signal aléatoire est dit être un **bruit blanc à bande limitée** *s'il satisfait la condition :*

$$\Psi_x(\omega) = \begin{cases} \dfrac{\eta}{2} & |\omega| \leq \omega_B \\ \\ 0 & |\omega| \geq \omega_B \end{cases}$$

D'où l'on déduit que :

$$R_x(\tau) = \frac{1}{2\pi} \int_{-\omega_B}^{\omega_B} \frac{\eta}{2} e^{j\omega\tau} d\omega = \frac{\eta\omega_B}{2\pi} \frac{\sin \omega_B \tau}{\omega_B \tau}$$

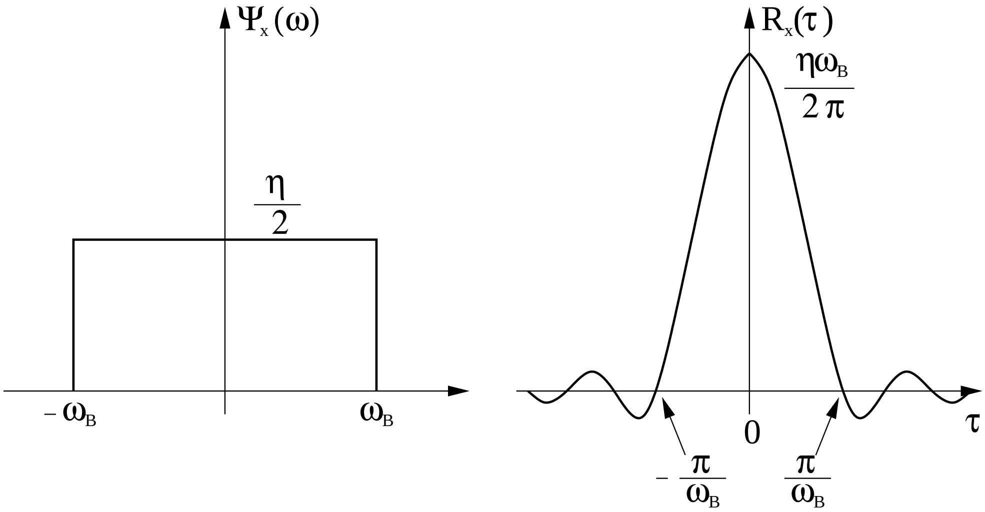


FIG. 2.2 – *bruit blanc à bande limitée*

2.6 Processus aléatoire à bande étroite

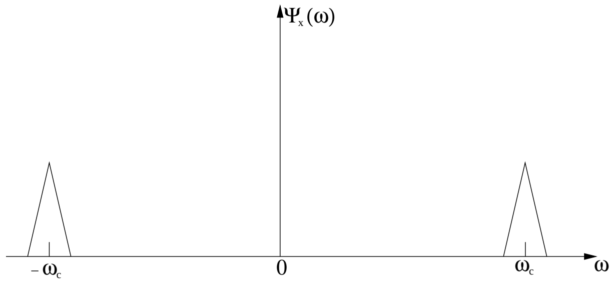
Un autre type de processus entrant dans la définition de processus stochastiques plus complexes sont les processus à bande étroite.

Définition 2.6.1 *: processus aléatoire à bande étroite*

Soit $x(t)$ un processus SSL de moyenne nulle et de densité spectrale de puissance $\Psi_x(\omega)$ non nulle dans une bande de largeur $2W$ dont la valeur est faible par rapport à la fréquence centrale ω_c de cette bande. Le processus est dit à bande étroite.

$$x(t) = V(t)\cos(\omega_p t + \phi(t))$$

où $V(t)$ est un processus aléatoire appelé **enveloppe** *et $\phi(t)$ est un processus aléatoire appelé* **phase**.



Propriété 2.6.1 *: décomposition cartésienne*

- *Tout processus à bande étroite peut être factorisé comme :*

$$x(t) = X_c(t)\cos\omega_p t - X_s(t)\sin\omega_p t$$

  où $X_c = V(t)\cos\phi(t)$ est la composante en phase et $X_s(t) = V(t)\sin\phi(t)$ est la composante en quadrature.

$$V(t) = \sqrt{X_c(t)^2 + X_s(t)^2} \quad \phi(t) = \mathrm{atan}\frac{X_s(t)}{X_c(t)}$$

- *X_s et X_c ont une densité spectrale identique :*

$$\Psi_{X_c} = \Psi_{X_s} = \Psi_x(\omega - \omega_p) + \Psi_x(\omega + \omega_p) \quad |\omega| \leq W$$

- *X_c et X_s ont même moyenne et même écart type que $x(t)$ et sont décorrélés.*

2.7 Processus de Wiener

Le processus suivant a une importance théorique toute particulière dans l'étude générale des processus stochastiques. Avant de le définir rigoureusement, il est nécessaire d'introduire une définition intermédiaire.

Définition 2.7.1 *: incréments stationnairement indépendants*
Un processus stochastique $\{x_t(\omega)\ ,\ t \in T\}$ *noté* $x(t)$ *a des* **incréments stationnairement indépendants** *si pour tout les ensembles finis* $\{t_i :\ t_i < t_{i+1}\} \in T$ *:*

- $x(t_2)-x(t_1), \cdots, x(t_n)-x(t_{n-1})$ *sont des vecteurs aléatoires indépendants.*
- $x(t+h) - x(\tau+h)$ *a la même distribution que* $x(t) - x(\tau)\ \ \forall\ t > \tau \in T,\ h > 0.$

Définition 2.7.2 *: processus de Wiener*
Un processus de Wiener *est un processus stochastique continu* $\{x_t,\ t \geq 0\}$ *si :*

- *il est à incréments stationnairement indépendants*
- *il est normalement distribué*
- *sa moyenne est nulle*
- $P[\{x_0 = 0\}] = 1$

Il est remarquable de noter que pour la spécification de la loi de probabilité d'un processus à incréments indépendants, il suffit de spécifier la distribution de $x(t)$ et la distribution de $x(t) - x(\tau),\ \ \forall\ t > \tau \in T$. Du fait des propriétés vues dans le chapitre 1 alors $x(t)-x(\tau)$ est aussi normalement distribué pour tout $t,\ \tau \geq 0$. Ainsi, afin de spécifier la loi de probabilité d'un processus de Wiener, il reste à spécifier la distribution de $x(t)-x(\tau)$ pour $t > \tau \geq 0$ qui est gaussienne et est donc spécifiée par sa moyenne et sa matrice de covariance. On a clairement :

$$E[x(t) - x(\tau)] = 0$$

De plus, pour $t \leq \tau$:

$$\begin{aligned} R(t,\tau) = E[x(t)x(\tau)'] &= E[((x(t) - x(t_0))(x(t) - x(\tau) + x(\tau) - x(t_0))'] \\ &= E[(x(t) - x(t_0))(x(t) - x(t_0))'] = E[x(t)x(t)'] \\ &= Q(t) \end{aligned}$$

De même, pour $t \geq \tau$:

$$R(t,\tau) = Q(\tau)$$

donc :

$$\boxed{R(t,\tau) = Q(min(t,\tau))}$$

On peut montrer également que la matrice $Q(t) = E[x(t)x(t)']$ est une fonction non décroissante monotone de t, soit :

$$Q(\tau) \geq Q(t) \qquad \forall\ \tau \geq t$$

Si l'on suppose de plus que $Q(t)$ est absolument continue alors il est possible d'écrire :

$$Q(t) = E[x(t)x(t)'] = \int_0^t V(s)ds$$

où $V(s) = V(s)' \geq 0$.

Cela permet de réécrire alors :

$$R(t,\tau) = \int_0^{min(t,\tau)} V(s)ds \qquad \forall\ t,\ \tau > 0$$

$$P(t,\tau) = E[((x(t) - x(\tau))(x(t) - x(\tau))'] = Q(t) - Q(\tau) = \int_\tau^t V(s)ds$$

Il est à noter qu'un processus de Wiener est un processus gaussien-markovien. Un processus de Wiener est en fait le processus limite quand $T \to 0$ de la marche aléatoire.

Il est également possible de montrer que le processus de Wiener est directement lié au bruit blanc. Soit $Z(t)$ le processus stochastique défini comme un bruit blanc.

$$E[Z(t)] = 0 \quad E[Z(t)Z(t)'] = \delta(t - \tau)$$

On définit le processus stochastique $W(t)$:

$$W(t) = \int_0^t Z(\xi)d\xi$$

alors $W(t)$ est un processus de Wiener :

- $E[W(t)] = 0$.
- $W(0)$: v.a. de valeur nulle.
- $W(t)$ est à ISI.
- $W(t)$ est normalement distribué

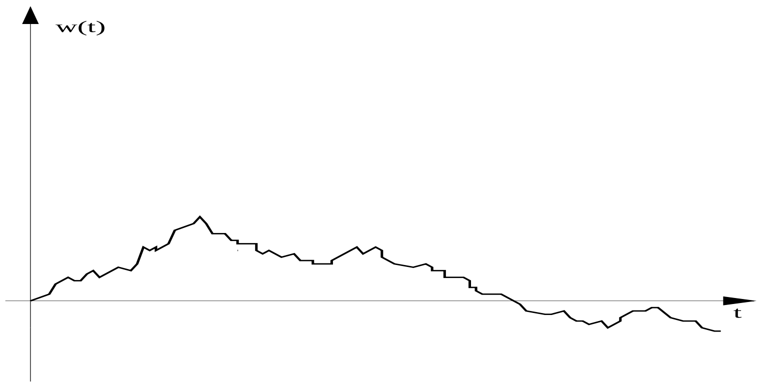


FIG. 2.3 – *Processus de Wiener*

2.8 Processus de Poisson

Les processus présentés dans les sections précédentes sont tous à états continus. Nous présentons maintenant une classe de processus très importante à états discrets.

Définition 2.8.1 *: processus de comptage*

Un processus $\{x(t)\ ,\ t \geq 0\}$ *est un processus de comptage si* $x(t)$ *représente le nombre total d'évènements intervenu aléatoirement dans* $(0\ ,\ t)$.

Propriété 2.8.1 *:*

1- $x(t) \geq 0$ *et* $x(0) = 0$

2- $x(t) \in \mathbb{N}$

3- $x(s) \leq x(t)$ *si* $s < t$

4- $x(t) - x(s)$ *est égal au nombre d'évènements intervenant dans* $(s\ ,\ t)$

Définition 2.8.2 *: processus de Poisson*

Un processus de comptage $\{x(t)\ ,\ t \geq 0\}$ *est un processus de Poisson d'intensité* λ *si :*

1- $x(0) = 0$

2- $x(t)$ *est à incréments indépendants (indépendance du nombre d'évènements dans* $T_1 \cap T_2 = \emptyset$*)*

3- Le nombre d'évènements dans un intervalle de longueur t *suit une distribution de Poisson de moyenne* λt

$$\forall\ s\ ,\ t > 0 \qquad P[x(t+s) - x(s) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad n = 0,\ 1, \cdots$$

Nota : ce processus est un modèle pour les arrivées d'appel à un central téléphonique, émission de particules d'une source radioactive, arrivées dans un magasin.

Une autre manière de définir un processus de Poisson est de définir un processus qui vérifie les propriétés suivantes.

Propriété 2.8.2 *:*

1- $x(t)$ *est à incréments stationnaires*

2- Moment $E[x(t)] = \lambda t$ *et variance* $var[x(t)] = \lambda t$

3- $\forall\ t > 0 \qquad P[x(t+\Delta t) - x(t) = 0] = 1 - \lambda\Delta t + o(\Delta t)$

4- Fonction d'autocorrélation $R_x(t,s) = \lambda \min(t,s) + \lambda^2 ts$

4- Le temps d'arrivée y *d'un évènement est une v.a. distribuée suivant la loi exponentielle :*

$$\forall\ t > 0 \quad P[y \leq t] = F(t) = 1 - e^{-\lambda t} \qquad p_y(t) = \lambda e^{-\lambda t}$$

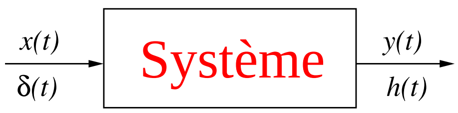
Chapitre 3

Systèmes linéaires et processus stochastiques : théorie fréquentielle

Dans ce chapître, nous étudions dans la première partie la transmission des signaux aléatoires par des systèmes linéaires temps-invariant dont le modèle est une fonction de transfert. Dans la deuxième partie, nous posons le problème inverse de la détermination de filtres linéaires dont la réponse à un bruit blanc a un spectre donné.

3.1 Filtrage linéaire des signaux aléatoires

Soit un système linéaire temps-invariant défini par sa réponse impulsionnelle $h(t)$ ou sa fonction de transfert $H(p) = \mathcal{L}[h(t)]$.



La réponse du système linéaire temps-invariant à un signal quelconque déterministe est donnée par :

$$y(t) = h(t) * x(t) = \int_{-\infty}^{t} h(\tau)x(t-\tau)d\tau$$

3.1.1 Moyenne et corrélation du signal de sortie

La moyenne $m_y(t)$ du signal de sortie peut se calculer par :

$$\begin{aligned} m_y(t) = E[y(t)] &= E\left[\int_{-\infty}^{t} h(\tau)x(t-\tau)d\tau\right] \\ &= \int_{-\infty}^{t} h(\tau)E[x(t-\tau)]d\tau \\ &= \int_{-\infty}^{t} h(\tau)m_x(t-\tau)d\tau = h(t) * m_x(t) \end{aligned}$$

De même, la corrélation $R_x(t_1, t_2) = E[y(t_1)y(t_2)]$ se calcule par :

$$R_x(t_1, t_2) = \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} h(\alpha)h(\beta)R_x(t_1-\alpha, t_2-\beta)d\alpha d\beta$$

Si le signal d'entrée $x(t)$ est stationnaire au sens large, on obtient alors :

$$m_y(t) = m_x\int_{-\infty}^{t} h(\alpha)d\alpha = m_x H(0)$$

De même $R_x(t_1, t_2)$ est une fonction de la différence temporelle $\tau = t_2 - t_1$.

$$R_y(\tau) = \int_{-\infty}^{\tau}\int_{-\infty}^{\tau} h(\alpha)h(\beta)R_x(\tau+\alpha-\beta)d\alpha d\beta$$

Si le signal d'entrée est stationnaire au sens large alors le signal de sortie est stationnaire au sens large.

3.1.2 Densité spectrale de puissance du signal de sortie

En appliquant la transformée de Fourier aux deux membres de l'équation précédente, la densité spectrale de puissance est obtenue :

$$\begin{aligned} \Psi_y(\omega) &= \int_{-\infty}^{\infty} R_y(\tau)e^{-j\omega\tau}d\tau \\ &= \int_{-\infty}^{\tau}\int_{-\infty}^{\tau}\int_{-\infty}^{\infty} h(\alpha)h(\beta)R_x(\tau+\alpha-\beta)e^{-j\omega\tau}d\tau d\alpha d\beta \\ &= |H(\omega)|^2\Psi_x(\omega) \end{aligned}$$

Ainsi, la densité spectrale de puissance du signal de sortie est égale à la densité spectrale de puissance du signal d'entrée multipliée par le carré du module de la réponse en fréquences du système. Si l'on souhaite calculer la corrélation du signal de sortie, il est plus aisé de calculer la densité spectrale de puissance du signal de sortie et de prendre la transformée de Fourier inverse.

$$R_y(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 \Psi_x(\omega) e^{j\omega\tau} d\omega$$

La puissance moyenne du signal de sortie est :

$$E[y^2(t)] = R_y(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 \Psi_x(\omega) d\omega$$

<u>Exemple :</u>

Soient le processus stochastique $x(t)$ SSL de moyenne 2 et de corrélation :

$$R_x(\tau) = \sigma^2 e^{-\dfrac{|\tau|}{\theta}}$$

et le système LTI de réponse impulsionnelle :

$$h(t) = 3e^{-2t}$$

On obtient alors :

- La moyenne $m_y(t)$:

$$m_y(t) = m_x(t)H(0) \quad H(p) = \frac{3}{p+2}$$
$$m_y(t) = 2\left(\frac{3}{2}\right) = 3$$

- La densité spectrale de puissance $\Psi_y(\omega)$:

$$\Psi_y(\omega) = |H(\omega)|^2 \Psi_x(\omega) \qquad \Psi_x(\omega) = \frac{2\sigma^2\theta}{\theta^2\omega^2 + 1}$$

$$\Psi_y(\omega) = \frac{18\sigma^2\theta}{(\theta^2\omega^2 + 1)(4 + \omega^2)}$$

3.2 Représentation spectrale des signaux aléatoires

L'utilisation de la transformée de Laplace bilatère permet une généralisation intéressante de la notion de densité spectrale de puissance à tout le plan complexe.

Définition 3.2.1 *:*
Un signal aléatoire $x(t)$ SSL peut être caractérisé par sa réponse fréquentielle ou spectre complexe $\Phi_x(p)$ défini comme la transformée de Laplace bilatère de sa fonction d'autocorrélation :

$$\Phi_x(p) = \mathcal{L}_2[R_x(\tau)] = \int_{-\infty}^{+\infty} R_x(\tau)e^{-\tau p}d\tau$$

Nota :
Les relations entre la densité spectrale de puissance et le spectre complexe sont données par :

$$\Phi_x(p) = \Phi_x(\omega)_{|\omega=-jp} \qquad \Phi_x(\omega) = \Phi_x(p)_{|p=j\omega}$$

Les résultats précédents concernant la transmission de la densité spectrale de puissance par des systèmes linéaires temps-invariant peuvent être étendus au spectre complexe.

Théorème 3.2.1 *:*
La sortie d'un système LTI défini par sa fonction de transfert $H(p)$ dont l'entrée est un signal aléatoire $x(t)$ SSL de spectre complexe $\Phi_x(p)$ est un signal aléatoire $y(t)$ de SSL de spectre complexe :

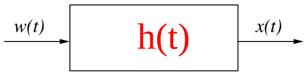
$$\Phi_y(p) = H(-p)\Phi_x(p)H(p)$$

La notion de spectre complexe permet d'aborder le problème difficile de la détermination des filtres formeurs associés à la donnée d'une densité spectrale de sa sortie.

Définition 3.2.2 *:*
Etant donné un signal aléatoire $x(t)$, un filtre formeur de $x(t)$ est le filtre linéaire stable à minimum de phase de fonction de transfert $H(p)$ tel que $x(t)$ est généré comme la réponse de ce filtre à un bruit blanc $w(t)$ stationnaire d'intensité unitaire.

On a donc les relations :

$$x(t) = \int_{-\infty}^{\infty} h(\tau)\delta(t-\tau)d\tau \quad E[x^2] = \int_{-\infty}^{\infty} h^2(t)dt$$



On obtient alors :

$$\Psi_x(p) = H(p)H(-p) \quad \Psi_x(\omega) = |H(\omega)|^2$$

Le problème de détermination du filtre formeur est connu également comme le problème de **factorisation spectrale**.

Problème 3.2.1 *:*
Etant donnée une fonction positive et paire délimitant une surface finie $\Psi_x(\omega)$, déterminer une fonction de transfert à minimum de phase $H(p)$ telle que :

$$\Psi_x(\omega) = |H(\omega)|^2$$

Théorème 3.2.2 *:*
Ce problème a une solution ssi $\Psi_x(\omega)$ vérifie la condition de bf Paley-Wiener :

$$\int_{-\infty}^{\infty} \frac{|\ln(\Psi_x(\omega))|}{1+\omega^2} d\omega < \infty$$

Nota :
Le problème de factorisation de $\Psi_x(\omega)$ n'est pas simple en général. Un cas particulier est celui des spectres rationnels.

Chapitre 4

Systèmes linéaires et processus gaussiens-markoviens

Dans ce chapître, nous étudions le processus stochastique défini comme l'état d'un système dynamique en temps discret ou en temps continu, perturbé par des entrées aléatoires. On s'intéressera exclusivement à établir les équations d'évolution des deux premiers moments (moyenne et covariance) caractérisant ces processus. Une attention toute particulière sera portée au cas où les entrées aléatoires perturbatrices sont des bruits blancs gaussiens.

4.1 Systèmes discrets

Soit le système récurrent :

$$x_{k+1} = F_k x_k + G_k w_k \quad , \quad x_0 \tag{4.1}$$

où $x_k \in \mathbb{R}^n$ est le vecteur d'état du système à un instant t_k repéré par l'indice k. $F_k \in \mathbb{R}^{nn}$, $G_k \in \mathbb{R}^{nm}$ sont les matrices dynamique et d'entrée du système. x_0 est la condition initiale, vecteur aléatoire de moyenne m_0 et de covariance P_0.
Enfin, $w_k \in \mathbb{R}^m,\ k = 0, 1, \cdots$ est une séquence de bruits stochastiques dont les deux premiers moments sont définis par μ_k, la moyenne et Q_{kl} la matrice de covariance.

Le système (4.1) est appelé **système dynamique stochastique en temps discret**. Si l'entrée aléatoire w_k n'est pas présente dans l'équation (4.1), on retrouve une équation récurrente classique qui a pour solution x_k. Dans le cas contraire, il est nécessaire de connaitre p_{x_k} à partir de laquelle on calcule $E[x_k]$ et P_{kl} qui donnent de l'information sur le comportement statistique de x_k.

Hypothèses 4.1.1 *:*
x_0 et $\{w_k,\ \forall\ k\}$ sont indépendants.

La solution de l'équation (4.1) s'écrit :

$$x_k = \Phi(k,0)x_0 + \sum_{i=0}^{k-1} \Phi(k,i+1)G_i w_i \quad , \quad k > 1 \tag{4.2}$$

où la matrice $\Phi(k,l)$ est la matrice de transition entre les instants l et k, définie par :

$$\begin{cases} \Phi(k,l) = F_{k-1}F_{k-2}\cdots F_{l+1}F_l \\ \\ \Phi(l,l) = \mathbf{1_n} \end{cases}$$

A partir de l'expression de x_k qui apparaît comme une combinaison linéaire des variables aléatoires x_0 et $(w_i,\ i = 0,\cdots,k-1)$, il est possible de déterminer les moments de ce vecteur aléatoire.

4.1.1 La moyenne ou espérance mathématique

A partir de (4.2), les matrices Φ et G étant à éléments déterministes et du fait de la linéarité de l'opérateur espérance, il vient :

$$E[x_k] = m_k = \Phi(k,0)m_0 + \sum_{i=0}^{k-1} \Phi(k,i+1)G_i \mu_i \tag{4.3}$$

Formellement équivalente à (4.2), cette expression montre que la moyenne du processus stochastique discret généré par (4.1) obéït à l'équation dynamique récurrente équivalente à (4.1).

$$\boxed{m_{k+1} = F_k m_k + G_k \mu_k \quad , \quad m_0}$$

Son obtention peut être effectuée directement à partir de (4.1), en appliquant l'opérateur espérance mathématique à cette expression.

4.1.2 La matrice de covariance

Soit :

$$P_{kl} = E[(x_k - m_k)(x_l - m_l)'] \tag{4.4}$$

Portant les expressions (4.2) et (4.3) dans (4.4), développant et utilisant le fait que x_0 et $(w_i,\ i = 0, \cdots, k-1)$ sont indépendants (annulation des termes croisés puisque $E[(x_0 - m_0)(w_i - \mu_i)'] = 0$), il vient :

$$P_{kl} = \Phi(k,0)P_0\Phi'(l,0) + \sum_{i=0}^{k-1}\sum_{j=0}^{l-1}\Phi(k,i+1)G_iQ_{ij}G_j'\Phi'(l,j+1) \tag{4.5}$$

qui est l'expression, dans le cas général, de la matrice de covariance.

Cette expression ne permet pas de rechercher une équation d'évolution récurrente "simple". La raison se trouve dans le second terme du second membre de (4.5) et notamment le fait qu'il y a un couplage dans le temps au niveau de la perturbation w (a priori $Q_{ij} \neq 0 \ \ \forall\ i,j$).

Pour poursuivre, on introduit une hypothèse supplémentaire.

Hypothèses 4.1.2 *:*

$\{w_k\}_k$ *est une séquence blanche, c'est à dire :*

$$E[(w_k - \mu_k)(w_l - \mu_l)'] = Q_k\delta_{kl} \tag{4.6}$$

où δ_{kl} *est le symbole de Kronecker*

$$\delta_{kl} = \begin{cases} 0 \ pour \ k \neq l \\ 1 \ pour \ k = l \end{cases}$$

Avec cette hypothèse (importante et intéressante du point de vue pratique) (4.5) se transforme en :

$$P_{kl} = \Phi(k,0)P_0\Phi'(l,0) + \sum_{i=0}^{min(k-1,l-1)}\Phi(k,i+1)G_iQ_iG_i'\Phi'(l,i+1) \tag{4.7}$$

et considérant la matrice de covariance P_k :

$$P_k = P_{kk} = E[(x_k - m_k)(x_k - m_k)']$$

on obtient :

$$P_k = \Phi(k,0)P_0\Phi'(k,0) + \sum_{i=0}^{k-1}\Phi(k,i+1)G_iQ_iG_i'\Phi'(k,i+1)$$

On vérifie qu'elle obéït à **l'équation récurrente de Lyapunov** suivante :

$$\boxed{P_{k+1} = F_kP_kF_k' + G_kQ_kG_k' \quad , \quad P_0}$$

On pourra vérifier que l'expression précédente peut être obtenue bien plus simplement à partir de (4.1). La dérivation faite ici a fait apparaître la nécéssité de l'hypothèse 2 (4.6), nécessité qui peut passer inaperçue en utilisant l'approche directe.

Avec l'hypothèse 2 et utilisant l'expression de P_k, il est possible d'établir que :

$$\boxed{P_{kl} = \begin{cases} \Phi(k,l)P_l \text{ pour } k > l \\ \\ P_k\Phi(l,k)' \text{ pour } k < l \end{cases}}$$

4.1.3 Nature du processus stochastique x_k

Jusqu'ici, nous nous sommes intéressés aux lois de variation des deux premiers moments associés au processus x_k sans connaître a priori la nature du processus stochastique. Dans cette section, sous des hypothèses données, nous précisons cette nature et le cas où la connaissance des deux premiers moments est suffisante.

Hypothèses 4.1.3 *:*

- $\{w_k\}_k$ *est une séquence blanche (indépendance de* w_k *et* w_l , $\forall\ k \neq l$*).*
- x_0 *et* $\{w_k,\ k = 0, 1, \cdots\}$ *sont indépendants*

Sous ces hypothèses,

$$\boxed{\text{Le processus stochastique } x_k \text{ est } \textbf{un processsus Markovien}.}$$

Démonstration:

En effet, à partir de (4.1), pour x_k donné, x_{k+1} ne dépend que de w_k qui, étant indépendant de $w_l,\ l < k$ et x_0, est donc indépendant de $x_{k-1}, x_{k-2}, \cdots, x_1$, donc :

$$p_{x_{k+1}/x_k,x_{k-1},\cdots,x_1,x_0}(\alpha_{k+1}/\alpha_k, \cdots, \alpha_0) = p_{x_{k+1}/x_k}(\alpha_{k+1}/\alpha_k)$$

Les équations récurrentes définies précédemment permettent de déterminer les deux premiers moments du processus stochastique x_k. Bien que permettant de répondre à nombre de questions pratiques, ces éléments ne suffisent pas seuls à caractériser complètement le processus stochastique, ce qui est le cas si on rajoute en plus l'hypothèse gaussienne.

Hypothèses 4.1.4 *:*

- $\{w_k\}_k$ *est une séquence blanche gaussienne.*

- x_0 *est une variable aléatoire gaussienne.*

Sous ces hypothèses,

$\boxed{\text{Le processus stochastique est } \textbf{un processus gaussien-markovien}.}$

Démonstration:

La propriété markovienne est établie ci-dessus. La propriété gaussienne découle très simplement du fait que toute combinaison linéaire de variables aléatoires gaussiennes est elle-même gaussienne. En effet,

$$x_1 = F_0 x_0 + G_0 w_0$$

x_1 dépend linéairement des vecteurs gaussiens x_0 et w_0, x_1 est donc un vecteur gaussien. Par induction, il est établi que x_k est gaussien.

Pour ce cas, la connaissance des deux premiers moments définis par les équations récurrentes revêt une importance toute particulière car elle permet de caractériser entièrement le processus stochastique du fait de son caractère gaussien.

4.1.4 Processus stationnaires

Dans le cas d'un système dynamique invariant dans le temps, les équations (4.1) sont données par :

$$x_{k+1} = F x_k + G w_k \quad , \quad x_0 \tag{4.8}$$

Deux hypothèses, intéressantes en pratique, sont faites.

Hypothèses 4.1.5 *:*

- *Le système (4.8) est asymptotiquement stable (les valeurs propres de F ont un module strictement inférieur à 1).*
- *La séquence blanche gaussienne $\{w_k\}_k$ est stationnaire. Cela signifie que moyenne et covariance Q_k sont indépendantes du temps :*

$$w_k \sim \mathcal{N}[\mu, Q]$$

Les équations d'évolution de la moyenne m_k et de la covariance P_k du processus stochastique x_k s'écrivent :

$$\left\{ \begin{array}{l} m_{k+1} = F m_k + G\mu \quad , \quad m_0 \\ \\ P_{k+1} = F P_k F' + G Q G' \quad , \quad P_0 \end{array} \right. \tag{4.9}$$

Alors, lorsque k croit indéfiniment, le processus stochastique x_k tend vers un processus gaussien de distribution stationnaire définie par :

$$m_k \to m = (\mathbf{1} - F)^{-1} G\mu$$

$$P_k \to P$$

où P est la solution symétrique définie positive de **l'équation matricielle algébrique de Lyapunov** :

$$\boxed{FPF' - P + GQG' = 0}$$

L'hypothèse de stabilité asymptotique sur F assure l'unicité de la solution et son caractère non-négatif. Ainsi, $P_{kl} = F^{k-l} P_l$, $(k > l)$ est uniquement fonction de l'intervalle $(k - l)$.

4.1.5 Fonction densité de probabilité de transition

Pour un processus markovien, la loi de distribution est entièrement caractérisée par p_{x_0} (*la distribution initiale*) et p_{x_{k+1}/x_k} (*la distribution de transition*). Dans le cas gaussien (x_{k+1} et x_k sont conjointement gaussiens) p_{x_{k+1}/x_k} définit également une distribution gaussienne. Pour sa caractérisation complète, il suffit de déterminer les deux premiers moments associés, à savoir :

- la moyenne conditionnelle :

$$\boxed{E[x_{k+1}/x_k] = m_{k+1/k}}$$

- la covariance conditionnelle :

$$\boxed{E[(x_{k+1} - m_{k+1/k})(x_{k+1} - m_{k+1/k})'/x_k] = P_{k+1/k}}$$

A partir de (4.1), il vient :

$$m_{k+1/k} = F_k x_k + G_k \mu_k \quad , \quad x_0$$

car

$$E[x_k/x_k] = m_{k/k} = x_k \quad et \quad E[w_k/x_k] = E[w_k] = \mu_k$$

De plus, $x_{k+1} - m_{k+1/k} = G_k(w_k - \mu_k)$ d'où

$$P_{k+1/k} = E[G_k(w_k - \mu_k)(w_k - \mu_k)' G_k'] = G_k Q_k G_k'$$

Dans le cas où cette matrice de covariance conditionnelle est inversible, l'expression analytique de la fonction densité de probabilité de transition peut être fournie :

$$\boxed{p_{x_{k+1}/x_k}(\alpha_{k+1}/\alpha_k) = \frac{1}{(2\pi)^{n/2}det^{1/2}[P_{k+1/k}]}exp[-1/2(\alpha_{k+1} - m_{k+1/k})'P_{k+1/k}^{-1}(\alpha_{k+1} - m_{k+1/k})]}$$

4.2 Systèmes continus

Dans le cas des systèmes dynamiques stochastiques à temps continu, il serait nécessaire afin de justifier tous les résultats, de développer la théorie des équations différentielles stochastiques. Les processus stochastiques ne sont pas des fonctions intégrables au sens courant de Riemann. L'intégrale de Riemann du calcul classique doit être modifiée afin de rentrer dans le cadre du calcul de Itô. Nous nous contenterons, dans le cadre de ce cours d'énoncer des résultats formellement étendus à partir de ceux du cas discret.

Soit le système dynamique à temps continu décrit par :

$$\dot{x}(t) = F(t)x(t) + G(t)w(t) \quad , \quad x(t_0) = x_0 \tag{4.10}$$

où $x \in \mathbb{R}^n$, $F \in \mathbb{R}^{nn}$, $G \in \mathbb{R}^{nm}$ sont définis classiquement.

Hypothèses 4.2.1 *:*

- x_0, *condition initiale, est un vecteur aléatoire de moyenne* m_0 *et de matrice de covariance* P_0.
- $w(t)$ *est un bruit, processus stochastique tel que :*

$$E[w(t)] = \mu(t) \text{ et } E[(w(t) - \mu(t))(w(\tau) - \mu(\tau))'] = Q(t, \tau)$$

- x_0 *et* $w(t)$ *sont indépendants.*

Pour une réalisation du bruit $w(t)$, la solution de (4.10) s'écrit :

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)G(\tau)w(\tau)d\tau \tag{4.11}$$

où $\Phi(t, t_0)$ est la matrice de transition du système, solution de l'équation différentielle matricielle :

$$\frac{d\Phi(t, t_0)}{dt} = F(t)\Phi(t, t_0) \quad , \quad \Phi(t_0, t_0) = \mathbf{1}$$

4.2.1 Espérance mathématique

Appliquant l'opérateur espérance mathématique à (4.11), permutant les opérateurs espérance et intégration, Φ et G étant des matrices à éléments déterministes, il vient :

$$E[x(t)] = m(t) = \Phi(t,t_0)m_0 + \int_{t_0}^{t} \Phi(t,\tau)G(\tau)\mu(\tau)d\tau \tag{4.12}$$

(4.12) est formellement identique à (4.11) montrant à l'évidence que la moyenne $m(t)$ obéït au système différentiel ordinaire :

$$\boxed{\dot{m}(t) = F(t)m(t) + G(t)\mu(t) \quad , \quad x_0}$$

Cette expression peut être immédiatement obtenue à partir de (4.10) par application de l'opérateur espérance mathématique en admettant que

$$E[\frac{dx}{dt}] = \frac{d}{dt}E[x]$$

4.2.2 Matrice de covariance

Par définition,

$$P(t,\tau) = E[(x(t) - m(t))(x(\tau) - m(\tau))']$$

Utilisant la solution (4.11) et du fait que x_0 et $w(t)$ sont indépendants, il vient :

$$P(t,\tau) = \Phi(t,t_0)P_0\Phi(\tau,t_0)' + \int_{t_0}^{t}\int_{t_0}^{\tau} \Phi(t,\theta_1)G(\theta_1)Q(\theta_1,\theta_2)G'(\theta_2)\Phi'(\theta_1,\theta_2)d\theta_1 d\theta_2$$

Expression générale complexe qui, comme dans le cas des systèmes discrets, est développée sous l'hypothèse :

Hypothèses 4.2.2 *:*
$w(t)$ est un bruit blanc:

$$Q(t,\tau) = Q(t)\delta(t-\tau)$$

alors :

$$P(t,\tau) = \Phi(t,t_0)P_0\Phi'(\tau,t_0) + \int_{t_0}^{min(t,\tau)} \Phi(t,\theta)G(\theta)Q(\theta)G'(\theta)\Phi'(t,\theta)d\theta \tag{4.13}$$

Considérons tout d'abord la matrice de covariance de $x(t)$:

$$\begin{aligned} P(t) &= P(t,t) = E[(x(t)-m(t))(x(t)-m(t))'] \\ P(t) &= \Phi(t,t_0)P_0\Phi'(t,t_0) + \int_{t_0}^{t} \Phi(t,\theta)G(\theta)Q(\theta)G'(\theta)\Phi'(t,\theta)d\theta \end{aligned}$$

Dérivant par rapport à t il vient :

$$\frac{dP(t)}{dt} = F(t)\left[\Phi(t,t_0)P_0\Phi'(t,t_0) + \int_{t_0}^{t} \Phi(t,\theta)G(\theta)Q(\theta)G'(\theta)\Phi'(t,\theta)d\theta\right] +$$
$$\left[\Phi(t,t_0)P_0\Phi'(\tau,t_0) + \int_{t_0}^{t} \Phi(t,\theta)G(\theta)Q(\theta)G'(\theta)\Phi'(t,\theta)d\theta\right] F'(t) + G(t)Q(t)G'(t)$$

soit **l'équation différentielle de Lyapunov** :

$$\boxed{\dot{P}(t) = F(t)P(t) + P(t)F'(t) + G(t)Q(t)G'(t) \quad , \quad P(t_0) = P_0}$$

On établit alors de façon similaire au cas discret :

$$\boxed{P(t,\tau) = \begin{cases} \Phi(t,\tau)P(\tau) & t \geq \tau \\ P(t)\Phi'(t,\tau) & t \leq \tau \end{cases}}$$

4.2.3 Nature du processus stochastique $x(t)$

On énonce sans les démontrer, les résultats similaires au cas discret.

Hypothèses 4.2.3 *:*
$w(t)$ est un bruit blanc indépendant de x_0.

Sous cette hypothèse,

Le processus stochastique $x(t)$ est **un processus markovien**.

Hypothèses 4.2.4 *:*

- *$w(t)$ est un bruit blanc gaussien.*
- *x_0 est un vecteur aléatoire gaussien.*
- *x_0 et $w(t)$ sont indépendants.*

Le processus stochastique $x(t)$ est **un processus gaussien - markovien.**

4.2.4 Processus stationnaires

On se place ici dans le cas de systèmes dynamiques invariants.

$$\dot{x}(t) = Fx(t) + Gw(t) \quad , \quad x_0 \tag{4.14}$$

Hypothèses 4.2.5 *:*

- *F et G sont des matrices constantes.*

- *F est stable asymptotiquement. Les valeurs prores de F ont une partie réelle strictement négative.*

- *$w(t)$ est un bruit blanc stationnaire décrit par :*

$$\mu(t) = \mu \quad Q(t) = Q$$

Les équations dynamiques des deux premiers moments deviennent :

$$\boxed{\dot{m}(t) = Fm(t) + G\mu \quad , \quad m_0}$$

$$\boxed{\dot{P}(t) = FP(t) + P(t)F' + GQG' \quad , \quad P_0}$$

ainsi :

$$\lim_{t\to\infty} m(t) = F^{-1}G\mu$$

$$\lim_{t\to\infty} P(t) = P$$

où P est la solution définie positive de **l'équation algébrique matricielle de Lyapunov** :

$$\boxed{FP + PF' + GQG' = 0}$$

De plus, $P(t,\tau) = e^{F(t-\tau)}P(\tau) = P(t-\tau)$, $(t > \tau)$. Donc, $x(t)$ tend asymptotiquement vers un processus stationnaire.

4.3 Processus générateur - représentation de markov

4.3.1 Définition

Etant donné un processus $s(t)$ (s_n) de moyenne nulle, on peut se poser le problème général de rechercher deux fonctions F et H et un processus $v(t)$ (v_n), de moyenne nulle tels que $v(t_1),\ v(t_2)$ $(v_{n_1},\ v_{n_2})$ soient non corrélés et tels que :

- Cas continu :

$$\dot{x}(t) = F(t)x(t) + v(t)$$

$$y(t) = H(t)x(t)$$

- Cas discret :

$$x_{n+1} = F_n x_n + v_n$$

$$y_n = H_n x_n$$

où $P_y(t,\tau) = P_s(t,\tau)$ $(P_y(k,l) = P_s(k,l))$.

Définition 4.3.1 *: processus générateur*

Le modèle d'état associé au processus $s(t)$ (s_n) est appelé **processus générateur** *du processus stochastique.*

La recherche d'un processus générateur dans le cas général est un problème complexe qui fait appel, pour sa résolution, à des techniques d'analyse harmonique ou spectrale qui sont hors du cadre de ce cours. Dans le cas où $s(t)$ (s_n) est un processus gaussien-markovien stationnaire de moyenne nulle, il est aisé de montrer que le processus générateur est caractérisé par :

- Cas continu :

$$F(t) = P_x(t,\tau)P_x^{-1}(t,\tau)$$

- Cas discret :

$$F_n = P_x(n+1,n)P_x^{-1}(n,n)$$

Définition 4.3.2 *: représentation markovienne*

Si un processus générateur existe alors on dit que $s(t)$ (s_n), est à **représentation markovienne**.

Si la densité spectrale fournit une représentation des signaux aléatoires stationnaires, le processus générateur associé à la représentation gaussienne-markovienne permet d'étudier les processus non stationnaires.

4.3.2 Quelques exemples

Dans cette section, nous donnons quelques exemples de processus générateurs associés à des processus aléatoires classiques. Une version continue et discrète est donnée.

4.3.3 Constante aléatoire

La constante aléatoire est une variable non dynamique d'amplitude aléatoire.

- Cas continu :

$$\dot{x}(t) = 0$$

- Cas discret :

$$x_{k+1} = x_k$$

La constante aléatoire peut être vue comme la sortie d'un intégrateur sans entrée mais avec une condition initiale aléatoire non nulle.

4.3.4 La marche aléatoire

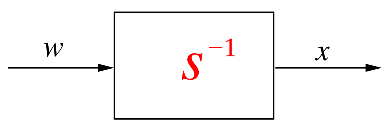
Ce processus a été défini au chapitre 2.

- Cas continu :

$$\dot{x}(t) = w \qquad E[w(t)w(\tau)] = V(t)\delta(t-\tau)$$

- Cas discret :

$$x_{k+1} = x_x + w_k \qquad E[w_k w_l] = V_k \delta_{kl}$$



4.3.5 Variable aléatoire exponentiellement corrélée

Une variable aléatoire dont la fonction d'autocorrélation est une exponentielle décroissante, est dite exponentiellement corrélée.

$$R_s(\tau) = \sigma^2 e^{-\beta|\tau|}$$

Ce type de processus est utilisé comme représentation des perturbations affectant les systèmes dynamiques. Cela fournit également une bonne approximation des signaux à bande limitée dont la densité spectrale de puissance est plate sur une bande passante finie.

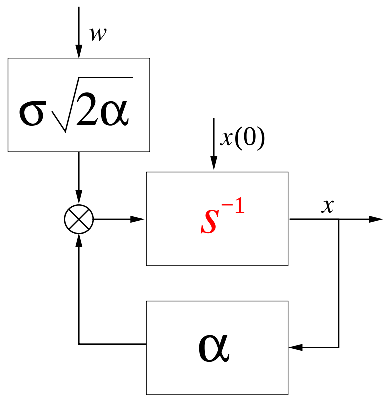
- Cas continu :

$$\dot{x}(t) = -\beta x(t) + w(t)$$

  où $w(t)$ est non corrélé, (bruit blanc par exemple).

- Cas discret :

$$x_{k+1} = e^{-\beta(t_{k+1}-t_k)} x_k + w_k$$



Bibliographie

[1] D. Arzelier, *Introduction à l'étude des processus stochastiques et au filtrage optimal*, Notes de cours, version 2, INSA, 1998.

[2] D. Arzelier, *Introduction à la théorie de l'estimation*, Notes de cours, version 1.2, ENSICA, 2000.

[3] J. Bernussou, *Processus stochastique - filtrage*, Cours ENSICA, Toulouse, 1994.

[4] D.R. Cox, H.D. Miller, *The theory of Stochastic Processes*, Chapamn et Hall, 1965.

[5] A. Gelb, *Applied Optimal Estimation*, The MIT Press, 1994.

[6] M.S. Grewal, A.P. Andrews, *Kalman filtering*, Prentice hall information and system siences series, Englewood Cliffs, New jersey, 1993.

[7] P.G. Hoel, S.C. Port, C.J. Stone, *Introduction to Probability Theory*, Haughton Milfilin Company, Boston, 1971.

[8] H.P. Hsu, *Communications analogiques et numériques*, Série Schaum, Mc Graw and Hill, 1994.

[9] H. Kwakernaak, R. Sivan, *Linear optimal control systems*, Wiley interscience, New York, 1976.

[10] A.H. Jazwinski, *Stochastic processes and filtering*, Academic press, New York, 1970.

[11] M. Labarrere, J.P. Krief, B. Gimonet, *Le fitrage et ses applications*, Cepadues editions, 1978.

[12] F.L. Lewis, *Applied optimal control and estimation*, Prentice hall, Englewood Cliffs, New Jersey, 1992.

Annexe A

Rappels de la théorie des probabilités

La théorie des probabilités est l'étude des lois régissant les phénomènes aléatoires de masse. Il ne faut pas réduire cette théorie à l'étude des méthodes de traitement des résultats des épreuves aléatoires et d'extraction des données nécessaires qui est appelée **statistique mathématique**.

Un phénomène aléatoire est un phénomène dans lequel le hasard intervient. L'exemple le plus simple de phénomène aléatoire est donné par les erreurs de mesure. Si l'on répète à de multiples reprises la mesure d'une même grandeur sous les mêmes conditions, les résultats obtenus ne seront jamais identiques. Le résultat de chaque mesure est entaché d'une erreur aléatoire et les résultats des différentes mesures contiennent différentes erreurs. On a ainsi une dispersion des résultats de mesure. Un autre exemple est fourni par la dispersion des obus tirés par un même canon visant la même cible. L'impossibilité de connaître avec une précision absolue les paramètres définissant l'atmosphère en tous les points constituant la trajectoire de l'obus conduit à cette dispersion aléatoire. Les pannes dans un dispositif technique quelconque sont à l'origine d'un autre type d'exemple de phénomènes aléatoires. En effet, il est impossible de prévoir si une panne aura lieu et dans l'affirmative, à quel instant précis. Dans la suite de ce chapître, nous allons rappeler quelques fondements de la théorie des probabilités en introduisant tout d'abord les notions et le vocabulaire élémentaire à l'aide d'une approche empirique. Partant de ces notions empiriques de probabilité, variable aléatoire..., nous déduirons en séquence, les notions abstraites par axiomatisation.

A.1 Approche empirique des phénomènes aléatoires

Pendant l'observation répétée de phénomènes aléatoires, certaines lois régissant ces phénomènes aléatoires peuvent être mises en évidence. L'étude de ces lois permet de "maitriser" le phénomène aléatoire étudié et d'en prévoir en quelque sorte les produits. Ces lois ne peuvent évidemment être déduites que si le phénomène aléatoire est observé un nombre suffisant de fois. Les phénomènes aléatoires pouvant être observés suffisamment souvent sont appelés **phénomènes aléatoires de masse**. La théorie des probabilités est donc la théorie des lois régissant les phénomènes aléatoires de masse. Son point de départ est constitué par des faits expérimentaux sur la base desquels seront formulés des concepts abstraits généralisant ces faits. Avant de définir ces concepts abstraits, il est nécessaire de détailler le vocabulaire que nous allons employer.

Définition A.1.1 *: épreuve*
L'observation d'un phénomène au cours de la réalisation d'un certain nombre de conditions et d'actions est appelée **épreuve**.

Une épreuve peut être caractérisée qualitativement ou quantitativement. Ainsi, les résultats produits par une épreuve peuvent ou non vérifier certaines propriétés.

Définition A.1.2 *: évènement*
Un évènement *est le fait que les résultats d'une épreuve vérifient une certaine propriété.*

Par exemple, le fait d'atteindre ou non une cible après un tir d'obus est un évènement.

Quantitativement, il peut être souhaitable de déterminer certaines grandeurs obtenues au cours du déroulement de l'épreuve.

Définition A.1.3 *: variable aléatoire*
Une variable aléatoire *est une grandeur choisie ou obtenue suite à une épreuve et pouvant prendre différentes valeurs (***ses réalisations***) non prévisibles.*

Des exemples simples de variables aléatoires sont donnés par le résultat d'une mesure ou le temps de fonctionnement d'un appareil quelconque sans panne. Un exemple typique d'évènement associé tout particulièrement à une variable aléatoire est défini par l'appartenance d'une variable aléatoire à un ensemble donné. Un autre élément quantitatif intéressant dans l'étude d'un phénomène aléatoire est la fréquence d'occurence d'un évènement lors de la répétition d'une épreuve donnée.

Définition A.1.4 *: fréquence d'un évènement*
La fréquence f *d'un évènement est définie par le quotient du nombre de réalisations d'un évènement par le nombre d'épreuves réalisées.*

Bien évidemment, afin de pouvoir faire des comparaisons entre différentes fréquences, il est nécessaire de supposer qu'une épreuve donnée peut être répétée indéfiniment. Par extension directe, il peut être nécessaire de calculer une fréquence d'un évènement A sous la condition qu'un évènement B a au préalable été réalisé.

Définition A.1.5 *: fréquence conditionnelle*
La fréquence conditionnelle *d'un évènement est le quotient du nombre de réalisations d'un évènement A, sachant qu'un évènement B s'est réalisé, par le nombre d'épreuves réalisées.*

Les propriétés des fréquences sont facilement déduites de leur définition.

Propriété A.1.1 *:*

- $0 \leq f \leq 1$.
- *La fréquence d'un évènement impossible est* 0.
- *La fréquence d'un évènement certain est* 1.
- $f(A \text{ ou } B) = f(A) + f(B)$ *si* A *et* B *sont incompatibles.*
- $f(A \text{ et } B) = f(A)f(B/A)$.

Expérimentalement, il est remarquable de constater que la valeur d'une fréquence d'un évènement se stabilise quand le nombre d'épreuves augmente. Il semble donc que la fréquence n'est plus un résultat aléatoire dépendant du nombre d'épreuves réalisées. Cette stabilité de la fréquence des évènements permet de supposer que chaque évènement est associé à un certain nombre, **la probabilité** de l'évènement. Cette probabilité est une caractéristique objective d'un évènement dans une épreuve donnée.

Associée à la fréquence conditionnelle, on peut également définir la notion abstraite de **probabilité conditionnelle**. Une fois définie cette notion, il est important, afin d'approfondir l'étude quantitative, de définir des moyens d'étudier la position des valeurs d'une variable aléatoire ainsi que leur dispersion lors de la tenue de n épreuves.

Définition A.1.6 *: échantillon*

Un échantillon *est l'ensemble des valeurs prises par une variable aléatoire dans une suite d'épreuves.*

Afin de caractériser la position des points expérimentaux d'un échantillon, on utilise habituellement la moyenne arithmétique des valeurs de la variable aléatoire.

Définition A.1.7 *: moyenne*
Supposons qu'une variable aléatoire X a pris au cours de n épreuves les valeurs $\alpha_1, \cdots, \alpha_n$ alors **sa moyenne** *est définie par :*

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}\alpha_i$$

De même, afin de caractériser la dispersion des valeurs d'une variable aléatoire, on définit la moyenne arithmétique des carrés des écarts des valeurs expérimentales de la variable aléatoire à la moyenne.

Définition A.1.8 *: variance*
Supposons qu'une variable aléatoire X a pris au cours de n épreuves les valeurs $\alpha_1, \cdots, \alpha_n$ alors **la variance** *est donnée par :*

$$v_x = \frac{1}{n}\sum_{i=1}^{n}(\alpha_i - \bar{x})^2$$

L'inconvénient majeur de la variance est qu'elle a la dimension du carré de la valeur de la variable aléatoire. Pour des raisons d'homogénéïté, on adopte comme mesure de dispersion l'écart quadratique moyen défini comme suit.

Définition A.1.9 *: écart-type*

$$\sigma_x = \sqrt{v_x} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\alpha_i - \bar{x})^2}$$

Lors de l'étude de plusieurs variables aléatoires liées au même évènement, il est également intéressant d'étudier la dépendance existant entre elles.

Définition A.1.10 *: covariance*
Soient deux variables aléatoires X et Y prenant respectivement au cours de n épreuves les valeurs $\alpha_1, \cdots, \alpha_n$ et $\beta_1, \cdots, \beta_n$. On définit **la covariance** *de X et Y par :*

$$C_{xy} = \frac{1}{n}\sum_{i=1}^{n}(\alpha_i - \bar{x})(\beta_i - \bar{y})$$

Cette grandeur entre en fait dans le calcul de la meilleur dépendance linéaire entre X et Y. Associée à la covariance, une mesure de dépendance entre X et Y est définie par la grandeur sans dimension suivante :

Définition A.1.11 *: coefficient de corrélation*

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y}$$

Nous sommes parvenus à la notion de probabilité en introduisant celle de fréquence. Toutefois, les propriétés des fréquences ne peuvent être démontrées une fois étendues aux probabilités. Il est donc nécessaire de les poser comme des axiomes définissant le concept abstrait de probabilité. On constatera dans le paragraphe suivant que nombre de notions définies empiriquement ont leur contrepartie abstraite servant à l'élaboration de la théorie des probabilités.

A.2 Théorie axiomatique des probabilités

A.2.1 Espace probabilisé

Il est nécessaire de définir en premier lieu le concept **d'espace probabilisé** associé à une expérience donnée et qui est le fondement de la théorie. Pour cela, on doit tout d'abord se donner un ensemble Ω, ensemble des résultats possibles à une expérience donnée, dont les éléments sont notés ω et sont exclusifs. Un **évènement**, la réalisation ou la non réalisation d'un phénomène, est défini comme un sous-ensemble de Ω. Alors, Ω est **l'évènement sûr** et $\emptyset$ est **l'évènement impossible**.

On note $\mathcal{B}$ une classe d'ensembles d'éléments ω (un ensemble d'évènements) munie de la structure algébrique de σ-algèbre (ou de tribu Borélienne).

Définition A.2.1 *: σ-algèbre*

*$\mathcal{B}$ est une σ-**algèbre** construite sur Ω si les propriétés suivantes sont satisfaites :*

Propriété A.2.1 *:*

$$
\begin{aligned}
&\bigcup_{i=1}^{\infty} A_i \in \mathcal{B} && \forall\ A_i \in \mathcal{B} \\
&\bigcap_{i=1}^{\infty} A_i \in \mathcal{B} && \forall\ A_i \in \mathcal{B} \\
&(\Omega - A) \in \mathcal{B} && \forall\ A \in \mathcal{B} \\
&\emptyset \in \mathcal{B} && \Omega \in \mathcal{B}
\end{aligned}
\tag{A.1}
$$

Définition A.2.2 *: mesure de probabilité*

Pour compléter la structure de l'espace probabiliste, on définit **une mesure de probabilité P**, *qui n'est rien d'autre qu'une fonction définie de* $\mathcal{B} \rightarrow [0 \quad 1]$, *satisfaisant les axiomes de probabilité suivants :*

Axiomes A.2.1 *:*

$$
\begin{aligned}
&P[A] \geq 0 \quad \forall\ A \in \mathcal{B} \\
&P[\Omega] = 1 \\
&Si\ A_i \bigcap A_j = \emptyset, \quad i \neq j \quad P[A_i \bigcup A_j] = P[A_i] + P[A_j]
\end{aligned}
\tag{A.2}
$$

Le dernier axiome s'étend à une séquence d'évènements A_i, mutuellement exclusifs deux à deux :

$$P[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} P[A_i]$$

A partir de ces axiomes, on déduit les résultats de base suivants :

Propriété A.2.2 *:*

$$
\begin{aligned}
&P[\emptyset] = 0 \\
&P[A] = 1 - P[\Omega - A] \\
&P[A_i \bigcup A_j] = P[A_i] + P[A_j] - P[A_i \bigcap A_j] \\
&A_1 \subset A_2 \Rightarrow P[A_1] \leq P[A_2]
\end{aligned}
\tag{A.3}
$$

Définition A.2.3 *: espace probabilisé*

Le triplet $(\Omega, \mathcal{B}, P)$ *est appelé* **espace probabilisé** *(ou encore probabiliste).*

Exemple 1 : le jeu de dé

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

ω, entier compris entre 1 et 6 (correspond au tirage de la face ω). Le tirage d'un nombre pair correspond à l'évènement $\{2, 4, 6\}$. Ainsi,

$$\mathcal{B}_1 = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$$

est une σ-algèbre alors que :

$$\mathcal{B}_2 = \{\emptyset, \{1\}, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$$

n'est pas une σ-algèbre car :

$$[1] \bigcup [2, 4, 6] \notin \mathcal{B}_2$$

De plus, il est aisé de voir que $P[\{2, 4, 6\}] = 0.5$.

A.2.2 Probabilités conditionnelles

A partir de l'espace probabilisé $(\Omega, \mathcal{B}, P[\bullet])$, on peut définir une nouvelle loi de probabilité sur la même tribu $P[\bullet/A]$, qui est, par définition, la probabilité d'évènements sachant que l'évènement A est réalisé.

Définition A.2.4 *: probabilité conditionnelle*

$P[\bullet/A]$, *appelée* **probabilité conditionnelle**, *est une application de* $\mathcal{B}$ *dans* $[0 \quad 1]$ *définie par :*

$$P[B/A] = \frac{P[B \bigcap A]}{P[A]} \qquad \forall\ A,\ B \in \mathcal{B} \tag{A.4}$$

Exemple 2 :

On suppose que la population d'une ville est masculine à 40% et féminine à 60%. On suppose de plus que 50% des hommes sont des fumeurs alors que 30% de femmes fument.

Trouver la probabilité qu'un fumeur soit un homme.

Soit H l'évènement correspondant à la sélection d'un homme et F l'évènement correspondant à la sélection d'une femme. De même, soit S l'évènement correspondant à la sélection d'un fumeur et N à celui d'un non fumeur.

Alors,

$$P[S/H] = 0.5 \quad P[S/F] = 0.3$$

$$P[H] = 0.4 \qquad P[F] = 0.6$$

On recherche $P[H/S]$. Par définition,

$$P[H/S] = \frac{P[H \bigcap S]}{P[S]}$$

De plus,

$$P[H \bigcap S] = P[H]P[S/H] = 0.2$$

Or,

$$S = (S \bigcap H) \bigcup (S \bigcap F)$$

soit

$$P[S] = P[S \bigcap H] + P[S \bigcap F]$$

et

$$P[S \bigcap F] = P[F][S/F] = 0.18$$

d'où :

$$P[S] = 0.2 + 0.18 = 0.38 \qquad P[H/S] = 0.2/0.38 = 0.53$$

De la définition de la probabilité conditionnelle et de l'exemple précédent, on déduit aisément **la formule de Bayes** :

$$\boxed{P[B/A] = P[A/B]\frac{P[B]}{P[A]}}$$

La probabilité conditionnelle permet d'introduire la notion d'indépendance statistique.

Définition A.2.5 *: indépendance*
A et B sont dits **indépendants** *si* $P[B/A] = P[B]$ *alors* $P[B \bigcap A] = P[A]P[B]$.

Remarques : deux évènements mutuellement exclusifs ne sont pas indépendants puisque

$$P[B/A] = \frac{P[B \bigcap A]}{P[A]} = 0$$

A.2.3 Variables aléatoires

Soit l'ensemble Ω, sur lequel on a défini une tribu $\mathcal{B}$ et une mesure de probabilité $P[\bullet]$, définissant ainsi un espace probabiliste $(\Omega, \mathcal{B}, P[\bullet])$.

Une variable aléatoire x permet d'assigner un nombre $x(\omega)$ à chaque résultat ω d'une expérience aléatoire.

Définition A.2.6 *:*
Une Variable aléatoire (v.a.) x est une fonction $x \; : \; \Omega \; \rightarrow \; \mathbb{R}$ satisfaisant :

1- L'ensemble $\{x(\omega) \leq \alpha\}$ est un évènement (un élément de la tribu $\mathcal{B}$) $\forall \; \in \mathbb{R}$.

2- Les probabilités des évènements $\{x(\omega) = \infty\}$ et $\{x(\omega) = -\infty\}$ sont nulles :

$$P[\{x(\omega) = \infty\}] = 0 \qquad\qquad P[\{x(\omega) = -\infty\}] = 0$$

Variables aléatoires discrètes

On définit un ensemble discret qui peut être fini ou infini et dénombrable $\mathcal{C} = \{\alpha_1, \alpha_2, \cdots\} \subset \mathbb{R}$. Si la variable aléatoire prend ses valeurs dans cet ensemble, elle sera dite discrète.

Définition A.2.7 *: variable aléatoire réelle discrète*
A tout point $\omega \in \Omega$, on associe un nombre réel $x(\omega)$, par la fonction, appelée **variable aléatoire discrète**, *$x \; : \; \Omega \; \rightarrow \; \mathcal{C}$ et possédant les propriétés définies précédemment à la définition A.2.6.*

L'ensemble $\{\omega \in \Omega \;\; : \;\; x(\omega) = \alpha_i\}$ est un évènement pour lequel il est possible de définir une valeur de la probabilité associée.

Définition A.2.8 *: distribution*
La fonction $p_x(\alpha) \; : \; \mathcal{C} \rightarrow [0 \;\; 1]$ telle que :

$$\boxed{p_x(\alpha) = P[\{\omega \in \Omega \; : \; x(\omega) = \alpha\}]}$$

s'appelle **la distribution de probabilité** *de la v.a. x.*

Cette fonction décrit comment les probabilités sont distribuées sur l'ensemble des valeurs possibles de x.

$$p_x(\alpha_j) = P[x = \alpha_j] = \sum_{i=1} P[x = \alpha_i]\delta_{i-j} = \sum_{i=1} p_i \delta_{i-j} \tag{A.5}$$

<u>Exemple</u> : jeu de dé

On définit :

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

et une v.a. :

$$x \; : \; \Omega \to \mathbb{R}$$
$$x(i) = i \qquad \forall \; i \in \Omega$$

La distribution de probabilité de la v.a. discrète associée au jeu de dé précédent est représentée à la figure ci-dessous.

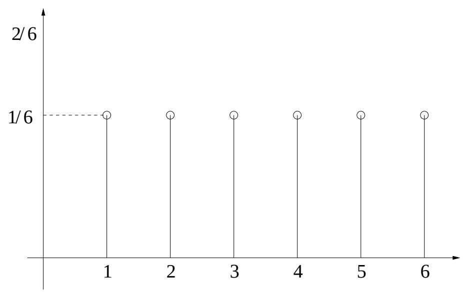


FIG. A.1 – *Exemple de distribution discrète*

Propriété A.2.3 *:*

- $0 \leq p_x(\alpha) \leq 1 \quad \forall \; \alpha \in \mathcal{C}$
- $\{\alpha \; : \; p_x(\alpha) \neq 0\}$ *est un sous-ensemble fini ou infini dénombrable de* $\mathbb{R}$
- $\sum_{\alpha_i} p_x(\alpha_i) = 1$
- *Soit* $\mathcal{D}$ *une collection de résultats possibles pour* x *:*

$$P[x(\omega) \in \mathcal{D}] = \sum_{x \in \mathcal{D}} p_x(\alpha)$$

Pour une variable aléatoire discrète $x \; : \; \Omega \; \to \; \mathcal{C}$, on définit **la fonction densité** :

$$f_x(\alpha) = \sum_{i=1} P[x = \alpha_i] \delta(\alpha - \alpha_i) = \sum_{i=1} p_i \delta(\alpha - \alpha_i) \qquad \text{(A.6)}$$

Exemple : jeu de dé (suite)

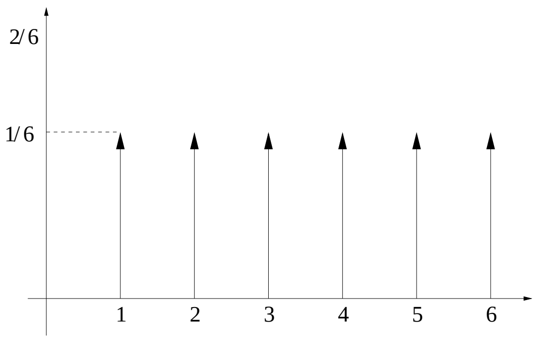


FIG. A.2 – *Exemple de densité discrète*

Définition A.2.9 *:*
Etant donnée une v.a. discrète $x(\omega) \ : \Omega \rightarrow \mathcal{C}$*,*

$$F_x \ : \ \Omega \rightarrow \mathcal{C}$$
$$F_x(\alpha) = P[\{x(\omega) \leq \alpha\}] = \sum_{\beta=-\infty}^{\alpha} p_x(\beta)$$

est appelée **fonction de répartition discrète de x**

Exemple : jeu de dé (suite)
La fonction de répartition $F_x(\alpha)$ a l'allure suivante :

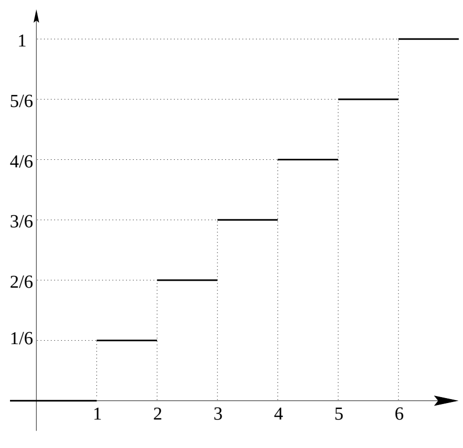


FIG. A.3 – *Exemple de fonction de répartition*

Propriété A.2.4 *:*

- $F_x(\alpha) \geq 0 \quad \forall \ \alpha \in \mathbb{C}.$

- $F_x(\infty) = 1$ *et* $F_x(-\infty) = 0$.
- *Si* $\alpha_1 < \alpha_2$ *alors* $F_x(\alpha_1 \leq F_x(\alpha_2)$.
- $P[x > \alpha] = 1 - F_x(\alpha)$.
- F_x *est continue à droite :*

$$F_x(\alpha) = F_x(\alpha^+)$$

- $P[\alpha_1 < x \leq \alpha_2] = F_x(\alpha_2) - F_x(\alpha_1)$.
- $P[x = \alpha] = F_x(\alpha) - F_x(\alpha^-)$.

Variable aléatoire continue

Définition A.2.10 *:*
La variable aléatoire $x : \Omega \to \mathbb{R}$ *est dite continue de* **densité de probabilité** $p_x(\alpha)$ *telle que :*

$$P[\alpha_1 \leq x \leq \alpha_2] = \int_{\alpha_1}^{\alpha_2} p_x(\alpha) d\alpha \tag{A.7}$$

Propriété A.2.5 *:*

- $p_x(\alpha) \geq 0 \; \forall \; \alpha \in \mathbb{R}$.
- $\displaystyle\int_{-\infty}^{+\infty} p_x(\alpha) d\alpha = 1$.
- $P[\alpha \leq x \leq \alpha + d\alpha] = p_x(\alpha) d\alpha$.

De manière identique à ce qui a été fait dans le cas des variables aléatoires discrètes, il est possible de définir la notion de fonction de répartition.

Définition A.2.11 *:*
Soit x *une variable aléatoire continue de densité* $p_x(\alpha)$ *alors la fonction :*

$$\begin{aligned} &F_x \; : \; \Omega \to \mathcal{C} \\ &F_x(\alpha) = P[\{x(\omega) \leq \alpha\}] = \int_{-\infty}^{\alpha} p_x(u) du \end{aligned}$$

est appelée **fonction de répartition** *de la variable aléatoire* x.

Propriété A.2.6 *:*

- $F_x(\alpha) \geq 0 \quad \forall\, \alpha \in \mathbb{R}$
- $F_x(\alpha)$ *est une fonction non décroissante.*
- $F_x(-\infty) = 0 \qquad F_x(+\infty) = 1$
- $P[\alpha_1 < x(\omega) \leq \alpha_2] = F(\alpha_2) - F(\alpha_1)$
- $F_x(\alpha) = \displaystyle\int_{-\infty}^{\alpha} p_x(\xi)d\xi.$
- *En tout point où* $F_x(\alpha)$ *est dérivable, il vient* $p_x(\alpha) = \dfrac{d}{d\alpha}F_x(\alpha).$

Exemple :

Soit l'expérience de lancer d'une fléchette sur une cible circulaire de rayon R et centrée en 0. On suppose que le lanceur est suffisamment habile pour être toujours dans la cible. Soit x la v.a. associée à la distance de la fléchette à l'origine de la cible. On a alors

$$\begin{cases} \alpha < 0 & P[x \leq \alpha] = F_x(\alpha) = 0 \\ 0 \leq \alpha \leq R & P[x \leq \alpha] = F_x(\alpha) = \frac{\pi\alpha^2}{\pi R^2} = \frac{\alpha^2}{R^2} \\ R < \alpha & P[x \leq \alpha] = F_x(\alpha) = 1 \end{cases}$$

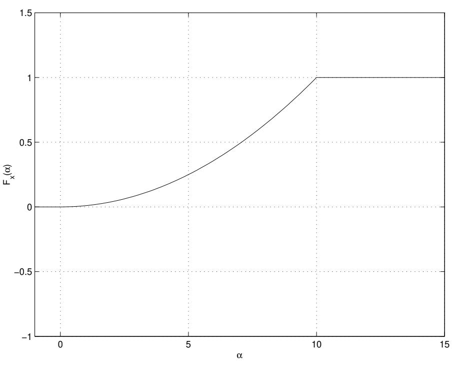


FIG. A.4 – *Exemple de fonction de répartition continue* ($R = 10$)

La fonction densité de probabilité de la v.a. continue associée au jeu de fléchette précédent est représentée à la figure ci-dessous.

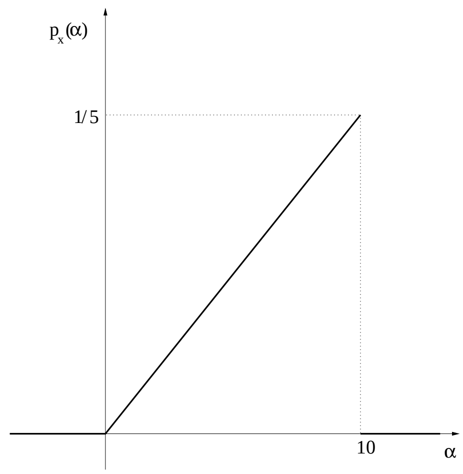


FIG. A.5 – *Exemple de fonction densité*

<u>Nota :</u> une v.a. continue est associée le plus souvent à des mesures de quantités physiques telles que coordonnées spatiales, poids, temps, température, tension, ... alors que les v.a. discrètes permettent plutôt de dénombrer des objets ou des évènements.

Définition A.2.12 *:*

Une v.a. x est continue si sa fonction de répartition $F_x(\alpha)$ est une fonction continue. De plus,

$$P[x = \alpha] = 0 \qquad -\infty < \alpha < +\infty$$

A.2.4 Propriétés statistiques des v.a.

Par souci de concision et sauf nécessité, il est fait abandon de la notation $x(\omega)$ au profit de x pour désigner la variable aléatoire.

Habituellement, lorsque l'on cherche la valeur moyenne d'une quantité aléatoire, on effectue N mesures et si l'on trouve n_i fois la quantité α_i, on calcule la moyenne par :

$$M = \frac{1}{N} \sum_i n_i \alpha_i$$

Ainsi, par analogie, on remplace la fréquence $\frac{n_i}{N}$ par $P_i = P_i[x = \alpha_i]$.

Définition A.2.13 *: moyenne*

On définira la moyenne ou moment du premier ordre par :

- *Cas discret :*

$$\boxed{m_1 = \sum_{\alpha_i} P_i \alpha_i = \sum_{\alpha_i} P_i[x = \alpha_i]\alpha_i}$$

- *Cas continu :*

$$\boxed{m_1 = E[x] = \int_{-\infty}^{+\infty} \alpha p_x(\alpha) d\alpha}$$

L'opérateur $E[\bullet]$ est appelé **opérateur espérance mathématique**. Il représente une moyenne pondérée des valeurs possibles de la variable aléatoire; le poids associé à la i^{eme} valeur est sa probabilité P_i.

Propriété A.2.7 *:*
Soit y, variable aléatoire définie par $y = f(x)$. Sa moyenne peut être calculée de deux manières :

- *Cas discret :*

$$E[y] = \sum_{\alpha} f(\alpha) p_x(\alpha) = \sum_{\alpha} f(\alpha) P[x = \alpha]$$

- *Cas continu :*

$$E[y] = \int_{-\infty}^{+\infty} \beta p_y(\beta) d\beta = \int_{-\infty}^{+\infty} f(\alpha) p_x(\alpha) d\alpha$$

- *Linéarité de l'opérateur espérance mathématique :*

$$\forall\ a, b \in \mathbb{R} \qquad E[ag(x) + bf(x)] = aE[g(x)] + bE[f(x)]$$

- *Soient x, y deux variables aléatoires indépendantes :*

$$E[xy] = E[x]E[y]$$

Définition A.2.14 *: moment d'ordre k*
Le moment d'ordre k *est :*

- *Cas discret :*

$$\boxed{m_k = E[x^k] = \sum_{\alpha} \alpha^k p_x(\alpha)}$$

- *Cas continu :*

$$\boxed{m_k = E[x^k] = \int_{-\infty}^{+\infty} \alpha^k p_x(\alpha) d\alpha}$$

Définition A.2.15 *: moment centré d'ordre k*
Le moment centré d'ordre k *est :*

- *Cas discret :*

$$\boxed{\mu_k = E[(x - E[x])^k] = \sum_{\alpha} (\alpha - E[x])^k p_x(\alpha)}$$

- *Cas continu :*

$$\boxed{\mu_k = E[(x - E[x])^k] = \int_{-\infty}^{+\infty} (\alpha - E[x])^k p_x(\alpha) d\alpha}$$

Noter que $E[x]$ est une constante, moyenne de la variable aléatoire x. En général, la connaissance du plus grand nombre de moments entraine une meilleure information sur la distribution de la variable aléatoire.

Définition A.2.16 *: la variance*
Le moment centré d'ordre 2 s'appelle **la variance** *qui est définie par :*

$$\boxed{var(x) = E[(x - E[x])^2] = \sigma(x)^2}$$

où $\sigma(x)$ *est appelé* **l'écart-type**.

La variance et l'écart-type donnent une information sur la dispersion de la variable aléatoire autour de sa moyenne.
Noter que $var(x) = E[x^2] - E^2[x]$.

Définition A.2.17 *: la covariance*
Etant données deux variables aléatoires x *et* y*, la quantité*

$$cov(x, y) = E[(x - E[x])(y - E[y])]$$

s'appelle **la covariance** *de* x *et* y*.*
De plus,

$$\boxed{var(x + y) = var(x) + var(y) + 2cov(x, y)}$$

Propriété A.2.8 *:*
Etant données deux variables aléatoires x et y indépendantes alors

$$cov(x,y) = 0 \qquad var(x+y) = var(x) + var(y)$$

Définition A.2.18 *: coefficient de corrélation*
Etant données deux variables aléatoires x et y, la quantité

$$\boxed{\rho(x,y) = \frac{cov(x,y)}{\sqrt{var(x)var(y)}}}$$

s'appelle **le coefficient de corrélation** *de x et y.*

C'est une mesure du degré de dépendance entre x et y. Toutefois, $\rho(x,y)$ indique que les deux v.a. sont **décorrélées** mais cela ne signifie pas obligatoirement qu'elles sont indépendantes.

A.2.5 Exemples de lois de distribution

Dans cette section, nous présentons deux exemples typiques de lois de distribution qui seront particulièrement utilisées lors de ce cours.

La loi normale

Définition A.2.19 *: loi normale*
La variable aléatoire associée x est dite **normale**, **gaussienne**, *ou* **normalement distribuée** *si sa fonction densité de probabilité s'écrit :*

$$\boxed{p_x(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2}\frac{(\alpha-m)^2}{\sigma^2}\right]}$$

Elle dépend de deux paramètres m et σ^2 qui sont, en fait, respectivement la moyenne et la variance de la variable aléatoire :

$$\begin{array}{l} E[x] = m \\ var(x) = E[(x-m)^2] = \sigma^2 \end{array}$$

Ces deux paramètres définissent complètement la loi de distribution.

<u>Remarques :</u>

- Pour $m = 0$, la loi normale est dite centrée.

- Pour $m = 0, \ \sigma = 1$, la loi normale est dite réduite (standard).

$$\boxed{p_x(\alpha) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\alpha^2}{2}}}$$

La fonction de répartition d'une loi normale réduite est tabulée avec :

$$\Phi(\alpha) = F_x(\alpha)_{|x \sim \mathcal{N}[0,1]} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{\frac{-v^2}{2}} dv$$

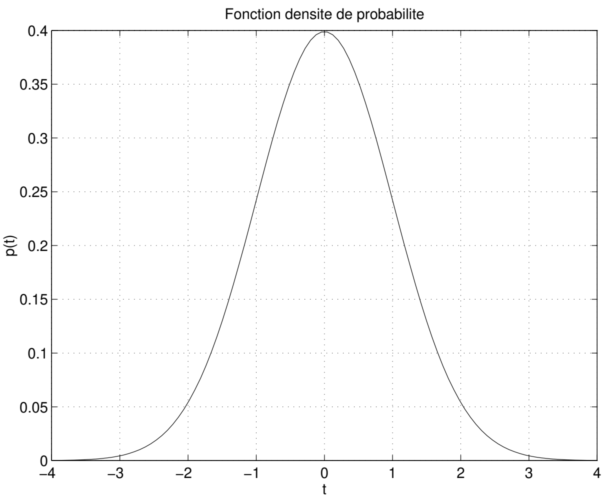


FIG. A.6 – *Loi normale standard*

De même dans le cas général, la fonction de répartition ne peut être calculée facilement et doit donc être évaluée numériquement.

Elle est donnée par :

$$F_x(\alpha) = P[x \leq \alpha] = \int_{-\infty}^{\alpha} \frac{1}{\sigma\sqrt{2\pi}} exp[-\frac{(t-m)^2}{2\sigma^2}] dt$$

qui par le changement de variables $u = \frac{(t-m)}{\sigma}$ se ramène à :

$$F_x(x) = \int_{-\infty}^{\frac{(\alpha-m)}{\sigma}} \frac{1}{\sqrt{2\pi}} exp[-\frac{u^2}{2}] du = \Phi\left(\frac{\alpha - m}{\sigma}\right)$$

qui peut, elle-même, se ramener à l'intégrale évaluée à l'aide de tables :

$$Erf(\frac{x}{\sqrt{2}}) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+x} e^{-\frac{u^2}{2}} du$$

En effet,

$$\Phi(\alpha) = \begin{cases} 0.5(1 + \mathrm{erf}(\alpha/\sqrt{2}) \\ 0.5(1 - \mathrm{erf}(|\alpha|/\sqrt{2}) \end{cases}$$

La loi de distribution normale est en pratique très importante. Elle permet de modéliser un grand nombre de phénomènes aléatoires pratiques. Toutefois, dans la plus part des applications, les variables aléatoires ont une loi de distribution approximativement normale. De plus, le théorème de la limite centrale établit que de nombreux phénomènes aléatoires complexes tendent à être bien représentés par une loi normale.

Théorème A.2.1 *:*
Pour n variables aléatoires indépendantes $x_1, \cdots, x_n$ de moyenne m_i et de variance σ_i^2 alors :

$$y_n = \frac{\displaystyle\sum_{i=1}^{n}(x_i - m_i)}{\sqrt{\sum i = 1^n \sigma_i^2}} \quad \xrightarrow{n\to\infty} \quad y \sim (0, 1)$$

Une somme normalisée d'un grand nombre de variables aléatoires uniformément petites et négligeables tend vers une variable aléatoire gaussienne.

D'autres exemples de lois continues sont résumées dans le tableau de la page suivante.

Nom	expression de la densité
Lognormal	$\frac{1}{\sqrt{2}\sigma\alpha}\exp[-\dfrac{(\ln(\alpha) - \mu)^2}{2\sigma^2}]$
Rayleigh	$\alpha\exp[-\alpha^2/2]$
Maxwell	$\alpha^2\exp[-\alpha^2/2]$
Beta	$\alpha^b(1 - \alpha)^c$
Gamma	$\alpha^n\exp[-\alpha]$
Laplace	$\exp[-\|\alpha\|]$
Cauchy	$\frac{1}{1+\alpha^2}$

La loi de Poisson

Définition A.2.20 *: Loi de Poisson*

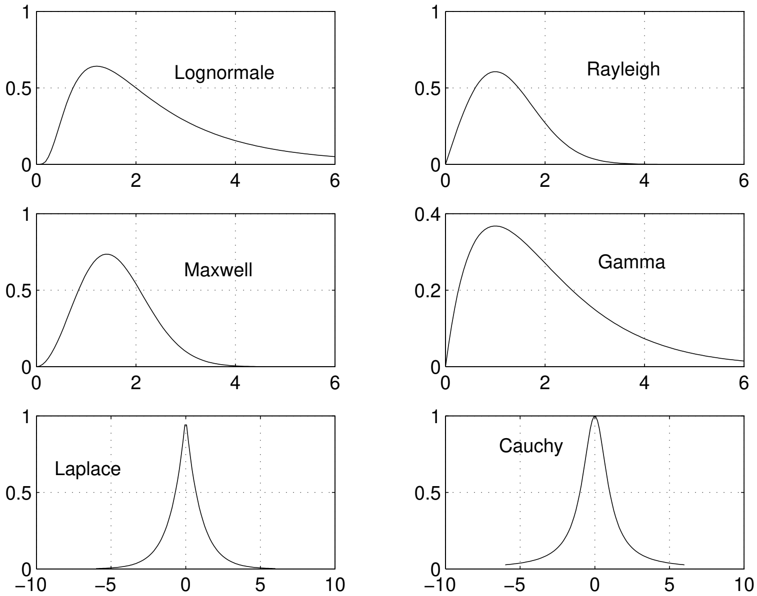


FIG. A.7 – *Quelques lois de distribution*

Une v.a. x est distribuée de manière poissonienne avec un paramètre a si elle prend les valeurs $0, 1, \cdots, n$ avec :

$$p_x(k) = P[x = k] = e^{-a}\frac{a^k}{k!} \qquad k = 0, 1, \cdots$$

La densité de probabilité est alors donnée par :

$$f_x(\alpha) = \sum_k P[x = k]\delta(\alpha - k) = e^{-a}\sum_{k=0}^{\infty}\frac{a^k}{k!}\delta(\alpha - k)$$

Il est à noter que :

$$\frac{P[x = k-1]}{P[x = k]} = \frac{k}{a}$$

A.3 Vecteurs aléatoires

Nous ne présenterons dans cette partie que le cas des vecteurs aléatoires continus, le cas discret pouvant s'en déduire aisément.

A.3.1 Loi de distribution conjointe

Définition A.3.1 *: loi de distribution conjointe*

Des variables aléatoires $x_1,\ x_2,\ \cdots, x_n$ ont **une loi de distribution conjointe** *si elles sont définies par rapport au même univers Ω.*

Elles peuvent être caractérisées via :

- leur fonction de répartition conjointe :

$$\boxed{F_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n) = P[x_1 < \alpha_1,\cdots,x_n < \alpha_n]}$$

  qui est la probabilité de l'évènement défini comme l'ensemble des éléments ω qui satisfont :

$$x_1(\omega) < \alpha_1 \ et \ \cdots \ et \ x_n(\omega) < \alpha_n \Leftrightarrow \omega \in \{x_1(\omega) < \alpha_1\}\bigcap\cdots\bigcap\{x_n(\omega) < \alpha_n\}$$

- leur fonction densité de probabilité conjointe :

$$p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n)$$

  définie par :

$$\boxed{F_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n) = \int_{-\infty}^{\alpha_1}\cdots\int_{-\infty}^{\alpha_n} p_{x_1,\cdots,x_n}(\xi_1,\cdots,\xi_n)d\xi_1\cdots d\xi_n}$$

  c'est à dire en tous les points où la fonction de répartition est dérivable par rapport à tous ses arguments :

$$p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n) = \frac{\partial^n F_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n)}{\partial\alpha_1\cdots\partial\alpha_n}$$

En dimension 2, ces concepts sont facilement illustrés. La fonction de répartition est donc la probabilité pour que le point de coordonnées $(x_1(\omega),\ x_2(\omega))$ soit à l'intérieur du rectangle défini par $(\alpha_1,\ \alpha_2)$.

<u>Exemple :</u>

Une loi dont l'utilisation est particulièrement répandue est la loi uniforme.

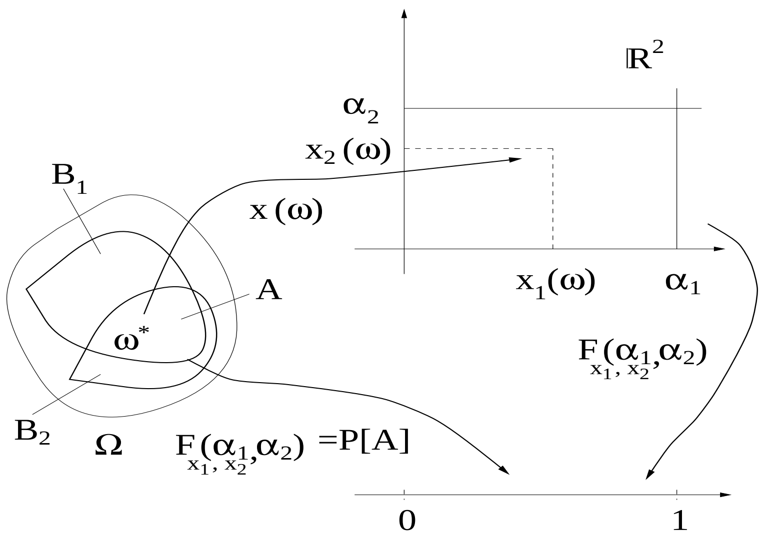


FIG. A.8 – *Distribution conjointe en deux dimensions*

Définition A.3.2 *: loi uniforme*
On appelle **loi uniforme** *sur* $[a\ \ b]$*, une loi de probabilité définie par :*

$$F_x(\alpha) = \begin{cases} 0 \text{ } pour \text{ } \alpha < a \\ \\ \dfrac{\alpha - a}{b-a} \text{ } pour \text{ } a \leq \alpha < b \\ \\ 1 \text{ } pour \text{ } \alpha \geq b \end{cases}$$

ou

$$p_x(\alpha) = \begin{cases} \dfrac{1}{b-a} \text{ } pour \text{ } a \leq \alpha < b \\ \\ 0 \text{ } ailleurs \end{cases}$$

Soient x et y deux variables aléatoires uniformément distribuées sur $[0\ \ T] \times [0\ \ T]$, alors :

$$P[x < T/2, y < T/2] = 1/4$$

$$p_{x,y}(\alpha,\beta) = \begin{cases} 1/T^2 \quad sur \text{ } [0\ \ T] \times [0\ \ T] \\ 0 \quad ailleurs \end{cases}$$

Propriété A.3.1 *:*

- *La fonction de répartition est une fonction non décroissante de chacun de ses arguments :*

$$p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n) \geq 0$$

- *Ce qui suit est exprimé pour $n = 2$ mais à extension immédiate pour $n > 2$.*

  -

$$F(-\infty,\alpha_2) = F(\alpha_1,-\infty) = F(-\infty,-\infty) = 0$$

  -

$$F(+\infty,+\infty) = 1 \Leftrightarrow \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_{x_1,x_2}(\alpha_1,\alpha_2)d\alpha_1 d\alpha_2 = 1$$

  -

$$F_{x_1,x_2}(\alpha_{11},\alpha_2) - F_{x_1,x_2}(\alpha_{12},\alpha_2) = P[\alpha_{12} \leq x_1 < \alpha_{11}, x_2 < \alpha_2]$$

A.3.2 Loi de distribution marginale

Si l'on s'intéresse à un sous-ensemble $x_1, \;\cdots, x_m\;,\; m < n$ des variables aléatoires précédentes distribuées conjointement, on définit **la loi de distribution marginale** à partir de la distribution conjointe :

- la fonction de répartition marginale :

$$\boxed{F_{x_1,\cdots,x_m}(\alpha_1,\cdots,\alpha_m) = F_{x_1,\cdots,x_m,\cdots,x_n}(\alpha_1,\cdots,\alpha_m,+\infty,\cdots,+\infty)}$$

- La fonction densité de probabilité marginale:

$$\boxed{p_{x_1,\cdots,x_m}(\alpha_1,\cdots,\alpha_m) = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n)d\alpha_{m+1}\cdots d\alpha_n}$$

- De plus,

$$p_x(\alpha) = \frac{\partial F(\alpha,\infty)}{\partial\alpha} \qquad p_y(\beta) = \frac{\partial F(\infty,\beta)}{\partial\beta}$$

Exemple :
Toujours sur le même exemple, les densités marginales sont :

$$p_x(\alpha) = \int_{-\infty}^{+\infty} p_{x,y}(\alpha,\beta)d\beta = \int_0^T (1/T^2)d\beta = 1/T \quad pour\ \alpha \in [0\quad T]$$

$$p_x(\alpha) = 0 \quad pour\ \alpha \notin [0\quad T]$$

On montre de la même manière que :

$$p_y(\beta) = 1/T \quad pour\ \beta \in [0\quad T]$$

$$p_y(\beta) = 0 \quad pour\ \beta \notin [0\quad T]$$

et donc, $p_{x,y}(\alpha,\beta) = p_x(\alpha)p_y(\beta)$, ce qui implique que x et y sont indépendantes.

A.3.3 Propriétés statistiques

Les moments sont bien sûr définis à partir des lois de distribution conjointes ou marginales précédentes et des fonctions associées.

- au premier ordre (moyenne des variables aléatoires x_i) :

$$\boxed{E[x_i] = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \alpha_i p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n)d\alpha_1\cdots d\alpha_n \qquad i = 1,\cdots,n}$$

- ordre quelconque :

$$\boxed{E[x_1^{k_1}\cdots x_n^{k_n}] = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \alpha_1^{k_1}\cdots\alpha_n^{k_n} p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n)d\alpha_1\cdots d\alpha_n}$$

- et les moments centrés correspondants :

$$\boxed{E[(x_1-m_1)^{k_1}\cdots(x_n-m_n)^{k_n}] = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (\alpha_1-m_1)^{k_1}\cdots(\alpha_n-m_n)^{k_n} p_{x_1,\cdots x_n}(\alpha_1,\cdots,\alpha_n)d\alpha_1\cdots d\alpha_n}$$

Exemple : suite de l'exemple sur la loi uniforme

$$E[x] = E[y] = T/2$$

Notation matricielle :

Quand on traite des vecteurs aléatoires, il est plus pratique d'utiliser une notation matricielle permettant d'obtenir des écritures plus concises.

Ainsi, notant :

$$X = [x_1 \ x_2 \ \cdots \ x_n]' \qquad \alpha = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]'$$

On définit alors :

$$F_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n) = F_X(\alpha)$$

et

$$p_{x_1,\cdots,x_n}(\alpha_1,\cdots,\alpha_n) = p_X(\alpha)$$

De plus,

$$E[X] = [E[x_1] \ E[x_2] \ \cdots \ E[x_n]]' = [m_1 \ m_2 \ \cdots \ m_n]' = M_X$$

Définition A.3.3 *: matrice de covariance*

La matrice de covariance Q_X *est la matrice dont les éléments* q_{ij} *sont égaux à* $cov(x_i, x_j)$*, soit :*

$$Q_X = E[(X - M_X)(X - M_X)']$$

Noter que Q_X est symétrique, semi-définie positive.

Définition A.3.4 *: matrice de corrélation*

La matrice $C_X = E[XX']$ *est dite* **matrice de corrélation** *(matrice des moments d'ordre deux non centrés)*

C'est également une matrice symétrique, semi-définie positive et l'on a la relation suivante :

$$\boxed{Q_X = C_X - M_X M_X'}$$

Exemple :

Soient les deux variables aléatoires x, y définies par :

$$x = sin(2\pi t) \qquad y = cos(2\pi t)$$

où t est uniformément distribuée sur $[0 \ \ 1]$. On peut calculer :

$$E[x] = E[y] = E[xy] = 0$$

donc $cov(x, y) = 0$.

A.3.4 Un théorème important

Soit $X = [x_1, \cdots, x_n]'$ un vecteur aléatoire dont la fonction densité de probabilité conjointe est connue :

$$p_{x_1,x_2,\cdots,x_n}(\alpha_1, \alpha_2, \cdots, \alpha_n) = p_X(\alpha)$$

Soit $Y = f(X)$ où f est une application de $\mathbb{R}^n \to \mathbb{R}^n$ inversible et continûment différentiable. Le théorème suivant permet d'exprimer la fonction densité de probabilité de Y en fonction de celle de X.

Théorème A.3.1 *:*
Sous les hypothèses précédentes concernant f et f^{-1}

$$\boxed{p_Y(\beta) = p_X[f^{-1}(\beta)]|det(J)|}$$

où J est la matrice jacobienne de f^{-1} :

$$J = \left[\frac{\partial f^{-1}(\beta)}{\partial \beta}\right]$$

Exemple :

Soit $y = x + z$ où $x,\ z$ sont deux variables aléatoires indépendantes. On cherche la densité de probabilité de y connaissant celles de x et z.

Posant :

$$Y = \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} X$$

où $X = \begin{bmatrix} x \\ z \end{bmatrix}$, alors $A^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ et l'on peut écrire :

$$p_{y,z}(\alpha, \beta) = p_{x,z}(\alpha - \beta, \beta)$$

d'où :

$$p_y(\alpha) = \int_{-\infty}^{+\infty} p_x(\alpha - \beta) p_z(\beta) d\beta$$

A.4 Distributions conditionnelles

A.4.1 Définition

Il a été vu précédemment qu'étant donnés deux évènements $A_1,\ A_2$ définis en relation avec une expérience aléatoire caractérisée par le triplet

$\{\Omega, \mathcal{B}, P[\bullet]\}$, on détermine la probabilité conditionnelle de A_1 par rapport à A_2 par :

$$P[A_1/A_2] = \frac{P[A1 \bigcap A_2]}{P[A_2]}$$

$P[A_1/A_2]$ satisfait tous les axiomes des probabilités de telle sorte que l'on peut caractériser une expérience aléatoire par $(\Omega, \mathcal{B}, P[\bullet/A_2])$ où l'on s'intéresse tout particulièrement aux évènements pour lesquels A_2 est réalisé.

<u>Exemple :</u>

Reprenons l'exemple du jeu de dé avec $\Omega = \{1, 2, 3, 4, 5, 6\}$. Soit l'évènement $A = \{2, 4, 6\} = \{\text{Tirage pair}\}$, alors nous pouvons calculer les probabilités conditionnelles suivantes :

$$P[1/A] = P[3/A] = P[5/A] = 0 \qquad P[2/A] = P[4/A] = P[6/A] = 1/3$$

Soient deux variables aléatoires $x,\ y$ distribuées conjointement sur l'univers Ω. Etant données la distribution conjointe et la densité de x et y, on recherche la distribution et la densité de x étant donnée une réalisation β de y. Soient les évènements:

$$A_1 = \{\omega \in \Omega\ :\ x(\omega) < \alpha\} \quad A_2 = \{\omega \in \Omega\ :\ \beta \leq y(\omega) < \beta + d\beta\}$$

ainsi $P[A_1/A_2]$ représente la fonction de répartition de x conditionnellement à A_2 et notée :

$$F_{x/y}(\alpha/\beta) = \frac{P[x < \alpha\ ,\ \beta \leq y < \beta + d\beta]}{P[\beta \leq y < \beta + d\beta]} = \frac{F_{x,y}[\alpha\ ,\ \beta + d\beta] - F_{x,y}[\alpha\ ,\ \beta]}{F_y[\beta + d\beta] - F_y[\beta]}$$

Par passage à la limite $(d\beta \to 0)$ et sous l'hypothèse de dérivabilité des fonctions de répartition au point considéré, il vient :

$$\boxed{F_{x/y}(\alpha/\beta) = \frac{\partial F_{x,y}(\alpha, \beta)/\partial\beta}{\partial F_y(\beta)/\partial\beta} = \frac{\displaystyle\int_{-\infty}^{\alpha} p_{x,y}(\xi, \beta)d\xi}{p_y(\beta)}}$$

qui est donc la fonction de répartition de la variable aléatoire x conditionnellement à la réalisation β de la variable aléatoire y. La dérivée par rapport à α, fournit la fonction densité de probabilité conditionnelle :

$$\boxed{p_{x/y}(\alpha/\beta) = \tfrac{p_{x,y}(\alpha,\beta)}{p_y(\beta)}}$$

Ces notions de distributions conditionnelles sont très importantes car elles introduisent la notion de dépendance stochastique entre deux variables aléatoires.

A.4.2 La rêgle de Bayes

Ainsi, par raison de symétrie,

$$p_{x,y}(\alpha,\beta) = p_{x/y}(\alpha/\beta)p_y(\beta) = p_{y/x}(\beta/\alpha)p_x(\alpha)$$

d'où l'on tire **la rêgle de Bayes :**

$$\boxed{p_{x/y}(\alpha/\beta) = \frac{p_{y/x}(\beta/\alpha)p_x(\alpha)}{p_y(\beta)}}$$

A.4.3 Propriétés statistiques

L'opérateur $E[x/y]$ est l'espérance mathématique conditionnelle définie, classiquement, à partir de la fonction densité de probabilité conditionnelle par :

$$\boxed{E[x/y] = \int_{-\infty}^{+\infty} \alpha p_{x/y}(\alpha/\beta)d\alpha}$$

L'opérateur espérance mathématique conditionnelle permet la définition des moments conditionnels et notamment :

- l'espérance conditionnelle :

$$\boxed{m_{x/y} = E[x/y]}$$

- la covariance conditionnelle :

$$\boxed{cov(x_i, x_j/\beta) = E[(x_i - m_{x_i/y})(x_j - m_{x_j/y})/\beta]}$$

Il est à noter que $p_{x/y}(\alpha/\beta)$ dépend des réalisations de y; c'est donc elle-même une variable aléatoire. Ceci est également vrai pour la moyenne conditionnelle et tout moment conditionnel. On peut ainsi appliquer à ces diverses variables aléatoires, l'opérateur espérance mathématique qui consiste à calculer la moyenne de ces variables aléatoires sur toutes les réalisations β de la variable aléatoire y.
Ainsi :

$$\boxed{\begin{array}{c} E[p_{x/y}(\alpha/\beta)] = p_x(\alpha) \\ \\ E[E[x/y]] = E(x) \end{array}}$$

Développés pour le cas scalaire, les résultats et définitions ci-dessus s'étendent directement au cas multidimensionnel, où x, y sont des vecteurs.

A.4.4 Propriété d'indépendance

Des variables aléatoires (ou vecteurs) à distribution conjointe, sont dites indépendantes si l'une quelconque des propriétés suivantes est satisfaite :

$$\boxed{\begin{array}{l} F_{x,y}(\alpha, \beta) = F_x(\alpha) F_y(\beta) \\ \\ p_{x,y}(\alpha, \beta) = p_x(\alpha) p_y(\beta) \\ \\ E[f(x)g(y)] = E[f(x)]E[g(y)] \end{array}}$$

Dans ce cas,

$$\boxed{p_{x/y}(\alpha/\beta) = p_x(\alpha)}$$

<u>Nota :</u> deux variables aléatoires indépendantes sont décorrélées alors que l'inverse n'est pas forcément vrai.

$$\boxed{\text{Indépendance} \Rightarrow \text{décorrélation.}}$$

A.5 Vecteurs gaussiens

A.5.1 Définition

Soit Z un vecteur de variables aléatoires décomposable en deux sous-vecteurs $X \in \mathbb{R}^n$, $Y \in \mathbb{R}^m$, $Z' = [X' \;\; Y']$. Z est gaussien, ses composantes $z_i, \; i = 1, 2, \cdots n + m$ sont conjointement distribuées suivant une loi normale ou gaussienne.

Définissant les deux premiers moments m_Z et Q_Z, la fonction densité de probabilité s'écrit :

$$\boxed{p_Z(\gamma) = \left[(2\pi)^{\frac{m+n}{2}} (det(Q_Z))^{1/2} \right]^{-1} e^{\left[-1/2 (\gamma - m_Z)' Q_Z^{-1} (\gamma - m_Z) \right]}}$$

Correspondant à la partition de Z en X et Y, m_X et m_Y désignent les vecteurs espérance mathématique (moyenne) de X et Y respectivement, la matrice de covariance Q_Z se partitionne en :

$$Q_Z = \begin{bmatrix} Q_X & Q_{XY} \\ Q_{YX} & Q_Y \end{bmatrix} \qquad\qquad Q_Z > 0$$

où :

$$Q_X = E[(X - m_X)(X - m_X)'] \quad Q_Y = E[(Y - m_Y)(Y - m_Y)']$$

$$Q_{XY} = E[(X - m_X)(Y - m_Y)']$$

Important : les deux premiers moments définissent entièrement la loi de distribution normale ou gaussienne.
Notation :

$$Z \sim \mathcal{N}[m_Z, Q_Z]$$

A.5.2 Propriétés

Théorème A.5.1 *:*
Si Z est gaussien alors X et Y sont également gaussiens.

Démonstration : si X et Y sont décorrélés $\Leftrightarrow cov(X,Y) = E[(X - m_X)(Y - m_Y)'] = 0$ alors, il vient aisément :

$$p_{X,Y}(\alpha, \beta) = p_X(\alpha)p_Y(\beta)$$

c'est à dire, X et Y sont indépendantes.

Théorème A.5.2 *:*
Deux vecteurs gaussiens non corrélés sont indépendants.

De plus,

Théorème A.5.3 *:*
Si Z est gaussien, $W = CZ + A$ est également gaussien tel que :

$$\boxed{W \sim \mathcal{N}[A + Cm_Z, CQ_ZC']}$$

A et C sont respectivement un vecteur et une matrice de dimensions appropriées et d'éléments déterministes.

Théorème A.5.4 *:*
Si X et Y sont deux vecteurs conjointement normalement distribués alors la fonction densité de probabilité conditionnelle $p_{X/Y}$ est elle-même normale et définie par les deux paramètres

- *moyenne conditionnelle* $E[X/Y] = m_X + Q_{XY}Q_Y^{-1}(\beta - m_Y)$

- *covariance conditionnelle* $Q_{X/Y} = Q_X - Q_{XY}Q_Y^{-1}Q_{YX}$

Il est à remarquer que la matrice de covariance conditionnelle ne dépend pas de β. Ceci est une propriété importante du cas gaussien, propriété qui se concrétisera par la suite dans le problème du filtrage des systèmes linéaires excités par des bruits gaussiens.
Un autre résultat important est contenu dans la proposition suivante.

Proposition A.5.1 *:*
$V = X - E[X/Y]$ *est un vecteur aléatoire gaussien indépendant de* Y*.*

Démonstration : V apparaît comme une combinaison linéaire de vecteurs gaussiens, c'est donc un vecteur gaussien. Pour l'indépendance vis-à-vis de Y, il suffit de démontrer que V et Y sont décorrélés :

$$\begin{aligned} cov(V, Y) &= E[(X - E[X/Y])(Y - E[Y])'] \\ &= E[(X - E[X] - Q_{XY}Q_Y^{-1}(Y - E[Y])(Y - E[Y])'] \\ &= Q_{XY} - Q_{XY}Q_Y^{-1}Q_Y = 0 \end{aligned}$$

Annexe B

Rappels sur la transformée de Fourier

B.1 Définition

Définition B.1.1 *:* **transformée de Fourier**

La transformée de Fourier *d'une fonction $x(t)$ est une fonction de la pulsation définie par,*

$$\boxed{\mathcal{F}[x(t)] = F(\omega) = \int_{-\infty}^{+\infty} e^{-j\omega t} x(t) dt} \tag{B.1}$$

La transformation réciproque se définit par,

$$\boxed{x(t) = \mathcal{F}^{-1}[F(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{j\omega t} F(\omega) d\omega} \tag{B.2}$$

On démontre que pour que la transformée de Fourier et la transformation réciproque associée existent, il faut que $x(t)$ appartienne à l'espace des fonctions de carré sommable noté $\mathcal{L}_2$. Cela signifie en fait que $x(t)$ et sa transformée de Fourier sont à énergie finie.

B.2 Propriétés de la transformée de Fourier

Etant donnée la transformée de Fourier d'une fonction $x(t)$, on peut distinguer les parties réelles et les parties imaginaires, qui sont respectivement,

$$\begin{aligned} R(\omega) &= \int_{-\infty}^{+\infty} x(t)cos(\omega t)dt \\ I(\omega) &= \int_{-\infty}^{+\infty} x(t)sin(\omega t)dt \end{aligned} \tag{B.3}$$

B.2.1 Linéarité

$$\mathcal{F}[\alpha x(t) + \beta y(t)] = \alpha X(\omega) + \beta Y(\omega) \tag{B.4}$$

B.2.2 Dérivation

$$\mathcal{F}[\frac{dx(t)}{dt}] = j\omega X(\omega) \tag{B.5}$$

B.2.3 Décalage temporel

$$\mathcal{F}[x(t \pm T)] = e^{\pm j\omega T} X(\omega) \tag{B.6}$$

B.2.4 Convolution

$$\mathcal{F}^{-1}[X(\omega)Y(\omega)] = \int_{-\infty}^{+\infty} x(t-\tau)y(\tau)d\tau = \int_{-\infty}^{+\infty} x(\tau)y(t-\tau)d\tau \tag{B.7}$$

B.2.5 Relation de Parseval

$$\int_{-\infty}^{+\infty} x(t)^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |X(\omega)|^2 d\omega \tag{B.8}$$

B.3 Table de transformée

$x(t)$	$x(\omega)$
$\delta(t)$	1
$\delta(t-t_0)$	$e^{-j\omega t_0}$
1	$2\pi\delta(\omega)$
$u(t)$	$\pi\delta(\omega)+\frac{1}{j\omega}$
$sgn(t)$	$\frac{2}{j\omega}$
$\frac{1}{\pi t}$	$-jsign(\omega)$
$e^{j\omega_0 t}$	$2\pi\delta(\omega-\omega_0)$
$\cos\omega_0 t$	$\pi[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]$

$\sin \omega_0 t$	$-j\pi[\delta(\omega-\omega_0)-\delta(\omega+\omega_0)]$
$e^{-at}u(t) \quad a>0$	$\dfrac{1}{j\omega+a}$
$te^{-at}u(t) \quad a>0$	$\dfrac{1}{(j\omega+a)^2}$
$e^{-a\|t\|} \quad a>0$	$\dfrac{2a}{\omega^2+a^2}$
$e^{-t^2/(2\sigma^2)}$	$\sigma\sqrt{2\pi}e^{-\sigma^2\omega^2/2}$
$p_a(t)=\begin{cases}1 & \|t\|<a\\0 & \|t\|>a\end{cases}$	$2a\dfrac{\sin\omega a}{\omega a}$
$\dfrac{\sin at}{\pi t}$	$p_a(\omega)=\begin{cases}1 & \|\omega\|<a\\0 & \|\omega\|>a\end{cases}$
$x(t)=\begin{cases}1-\dfrac{\|t\|}{a} & \|t\|<a\\0 & \|t\|>a\end{cases}$	$a\left[\dfrac{\sin\omega a/2}{\omega a/2}\right]^2$
$\displaystyle\sum_{n=-\infty}^{\infty}\delta(t-nT)$	$\omega_0\displaystyle\sum_{n=-\infty}^{\infty}\delta(\omega-n\omega_0) \quad \omega_0=2\pi/T$