

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

NOTES DE COURS EN

Modèles Stochastiques



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

OLIVIER CLOUX

AUTOMNE 2017



Module 0: Table des Matières

1 Variables Aléatoires	4
1.1 Bases de proba	4
1.1.1 Axiomes de probabilités	4
1.1.2 Probabilité conditionnelle	4
1.1.3 Indépendance	5
1.2 Variable aléatoire	5
1.2.1 Fonction de répartition	5
1.2.2 Densité de probabilité	6
1.2.3 Fonction d'une variable aléatoire	7
1.2.4 Espérance d'une fonction d'une v.a.	7
1.2.5 Transformée	8
1.2.6 Inégalités	10
1.3 Variables aléatoires	10
1.3.1 Bernoulli	10
1.3.2 Binomiale	11
1.3.3 Géométrique	11
1.3.4 Poisson	12
1.3.5 Uniforme	12
1.3.6 Gaussienne	12
1.3.7 Exponentielle (Lambda)	13
1.3.8 Gamma	13
2 Vecteurs Aléatoires	14
2.1 Paires de Variables aléatoires	14
2.1.1 Fonction de répartition jointe et marginale	14
2.1.2 Densité de probabilité jointe et marginale	15
2.1.3 Variables aléatoires indépendantes	15
2.1.4 Fonction de deux variables aléatoires	16
2.1.5 Espérance d'une fonction de deux variables	17
2.1.6 Transformées (fonction caractéristique jointe)	17
2.1.7 Densité de probabilité conditionnelle	17
2.1.8 Espérance conditionnelle	17
2.2 Vecteurs aléatoire de dimension n	18
2.2.1 V.a. gaussiennes jointes/multivariées	18
2.3 Convergence d'une suite de variables aléatoires	18
2.3.1 Convergence presque sûre	18
2.3.2 Convergence en distribution et Théorème Central Limite	18
2.3.3 Convergence en probabilité et loi faible des grands nombres	18
3 Processus Stochastiques	19
3.1 Processus stochastiques à temps continu	19
3.1.1 Définition	19
3.1.2 Exemples	19
3.1.3 Stationarité	21
3.1.4 Ergodisme	21
3.1.5 Densité spectrale	21
3.1.6 Exemples	22

3.1.7	Ergodisme	22
3.2	Application 1 : Distorsion de fréquence en radiocommunication avec un mobile	24
3.3	Application 2 : Taux d'erreur d'un signal PAM	24
4	Processus stochastiques à temps discret : analyse au second ordre	24
5	Processus de Poisson	24
5.1	Poisson sur \mathbb{R}^+	24
5.1.1	Processus de comptage	24
5.1.2	Définition et loi de probabilité d'un processus de Poisson	24
5.1.3	Propriétés	26
5.1.4	Séquence des temps d'arrivée et inter-arrivées	27
5.2	Processus et bruit impulsif de Poisson sur \mathbb{R}	27
5.2.1	Processus de Poisson sur \mathbb{R}	27
5.2.2	Bruit impulsif de Poisson (Shot noise)	27
5.3	Application : rapport signal/bruit d'un photodétecteur (fibre optique)	29
6	Chaînes de Markov à temps discret	30
6.1	Chaînes de Markov à temps discret	30
6.1.1	Probabilités d'état et de transition	30
6.1.2	Classification des états	31
6.1.3	Des chaînes récurrentes classiques	32
6.1.4	Temps d'atteinte	33
6.1.5	Chaînes réversibles	35
7	Chaînes de Markov à temps continu	35
7.1	Introduction	35
7.1.1	Première définition	35
7.1.2	Distribution des temps de séjour	35
7.1.3	Chaînes de Markov induites	36
7.1.4	Équations de Kolmogorov	36
8	Files d'attente	36
A	Preuves	36
A.1	Chapitre 1	36
A.1.1	Fonction d'une variable aléatoire	36
A.1.2	Existence de la fonction caractéristique	37
A.1.3	Moment d'ordre k avec la fonction caractéristique	37
A.1.4	Inégalité de Markov	38
A.1.5	Inégalité de Chebychev	38
A.1.6	Limite d'une v.a. binomiale vers Poisson	38
A.2	Chapitre 2	38
A.2.1	La somme de deux v.a. indépendantes est le produit de convolution	38
A.3	Chapitre 3	39
A.4	Chapitre 4	39
A.5	Chapitre 5	39

A.5.1 Les deux définitions du processus de Poisson homogène sont équivalentes	39
--	----

Module 1: Variables Aléatoires

1.1 Bases de proba

1.1.1 Axiomes de probabilités

ζ_i le résultat de la i -ème expérience. $A = \{\zeta_1, \zeta_2, \zeta_{17}\}$ un événement et Ω l'événement certain (\emptyset l'événement impossible). L'espace de probabilité tel que nous le notons est

$$(\Omega, P(\cdot))$$

Exemple 1. Ω peut être l'ensemble des résultats d'un jet de dé, donc $\Omega = \{1, 2, 3, 4, 5, 6\}$

Il en résulte les axiomes des probabilités :

- $P(A) \geq 0$
- $P(\Omega) = 1$
- Dans le cas où Ω est un ensemble *fini*, si $A \cap B = \emptyset$ alors $P(A \cup B) = P(A) + P(B)$
- Dans le cas où Ω est un ensemble *infini*, si A_1, A_2, \dots est une suite d'événements telle que $A_i \cap A_j = \emptyset \forall i \neq j$ alors $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Et de ceci en suivent les corollaires :

- $P(\bar{A}) = 1 - P(A)$ où $\bar{A} = \Omega \setminus A$
- $P(A) \leq 1$
- $P(\emptyset) = 0$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Si $A \subseteq B$ alors $P(A) \leq P(B)$

1.1.2 Probabilité conditionnelle

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Théorème des probabilités totales : Soit $\{A_1, A_2, \dots, A_n\}$ une partition de Ω (donc $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$ et $A_i \cap A_j = \emptyset \forall i \neq j$). Alors

$$\begin{aligned} P(B) &= P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_n) \\ &= \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i) \end{aligned} \quad (1.1)$$

Exemple 2. Soit Ω la météo, partitionnée par $A_1 = \text{Soleil}$ et $A_2 = \text{Pluie}$. Nous cherchons la probabilité de venir en vélo. En sachant que $P(\text{soleil}) = 0.7$ et $P(\text{pluie}) = 0.3$, $P(\text{vélo}|\text{soleil}) = 0.8$ et $P(\text{vélo}|\text{pluie}) = 0.3$.
Alors $P(\text{vélo}) = P(\text{vélo}|\text{soleil})P(\text{soleil}) + P(\text{vélo}|\text{pluie})P(\text{pluie})$

Règle de Bayes Soit à nouveau $\{A_1, \dots, A_n\}$ une partition de Ω . Alors

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)} \quad (1.2)$$

1.1.3 Indépendance

A et B sont **indépendants** si $P(A \cap B) = P(A)P(B)$

Corollaire : $P(A|B) = P(A)$ et $P(B|A) = P(B)$

De même, trois événements sont indépendants si non seulement $P(A \cap B) = P(A)P(B)$ (et pareil pour toute combinaison de A, B, C) mais également $P(A \cap B \cap C) = P(A)P(B)P(C)$

1.2 Variable aléatoire

Une variable aléatoire (ou **v.a.**) est une fonction X qui map à chaque résultat ζ un réel $X(\zeta)$.

$$X : \Omega \mapsto \mathbb{R}$$

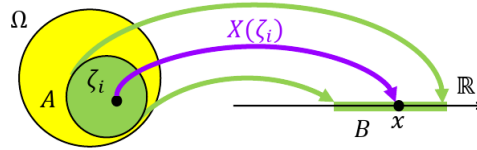


Figure 1: Représentation d'une variable aléatoire

Comme visible à la figure 1, nous pouvons établir la bijection $B = \{X(\zeta) : \zeta \in A\}$ et $A = \{\zeta : X(\zeta) \in B\}$ avec évidemment $P(A) = P(B)$

1.2.1 Fonction de répartition

Toute v.a. est caractérisée par sa fonction de répartition (ou Cumulative Distribution Function, CDF) ; cette fonction représente la probabilité que X prenne une valeur inférieure ou égale à un réel x

$$F_X(x) = P(X \leq x) = P(A) \quad (1.3)$$

Cette fonction a plusieurs propriétés :

P1. $0 \leq F_X(x) \leq 1$

P2. $\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow +\infty} F_X(x) = 1$

P3. $a < b \rightarrow F_X(a) \leq F_X(b)$

P4. $P(a < X \leq b) = F_X(b) - F_X(a)$

P5. $F_X(x)$ est continue à droite, i.e. $F_X(x) = \lim_{\substack{\epsilon \rightarrow 0 \\ \epsilon > 0}} (F_X(x + \epsilon)) = F_X(x^+)$

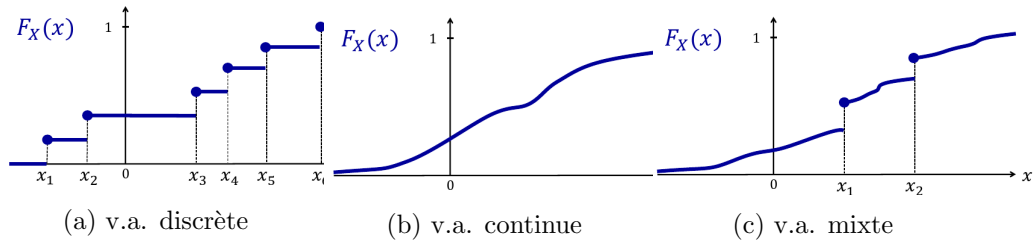


Figure 2: Différents types de support d'une variable aléatoire

La variable aléatoire X peut généralement prendre un ensemble de valeurs ; cet ensemble est appelé **le support** S_X , défini comme

$$S_X = \{X(\zeta) | \zeta \in \Omega\} \quad (1.4)$$

Les valeurs que peut prendre la variable aléatoire X peuvent être

- Discrètes (2a) : dénombrables (pas forcément fini). E.g. : des jets de dé. Alors $S_X = \{x_1, x_2, x_3, \dots\}$. Nous avons alors que $P(X = x_i) = F_X(x_i) - F_X(x_{i-1})$
- Continues (2b) : infini et indénombrable. E.g. : des mesures infiniment précises de température. Alors $P(X = x) = 0$ ¹
- Mixte (2c) : infini mais pas continu

1.2.2 Densité de probabilité

Aussi appelée Probability Density Function (PDF). Définie comme

$$f_X(x) = \frac{dF_X}{dx}(x) \quad (1.5)$$

On distinguera des propriétés différentes si le support est discret ou continu. Dans le cas d'une v.a. **continue**, elles sont ainsi :

P1. $f_X(x) \geq 0$

P2. $\int_{-\infty}^{\infty} f_X(x) dx = F_X(\infty) - F_X(-\infty) = 1 - 0 = 1$

P3. $P(X \leq a) = F_X(a) = \int_{-\infty}^a f_X(x) dx$

P4. $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx$, ce qui implique que $P(x < X \leq x + \Delta x) = f_X(x)\Delta x$ lorsque $\Delta x \rightarrow 0$

Pour un support **discret**, nous utilisons les impulsions de Dirac $\delta(\cdot)$ afin de réécrire l'équation 1.5 comme

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

avec $p_i = P(X = x_i)$. Ceci donne lieu à des modifications des propriétés :

¹ Bien que cela semble contre-intuitif, le nombre de valeurs est infini, donc chaque valeur exacte a une probabilité $\frac{1}{\infty}$ d'apparaître, soit nulle

$$P1. p_i \geq 0$$

$$P2. \sum_i p_i = 1$$

$$P3. P(X \leq a) = F_X(a) = \sum_{x_i \leq a} p_i$$

$$P4. P(X = x_i) = F_X(x_i) - F_X(x_i^-) = p_i$$

1.2.3 Fonction d'une variable aléatoire

Prenons $g(\cdot)$ une fonction continue, et X une v.a. continue. Alors

$$Y = g(X)$$

désigne aussi une v.a. continue². Nous cherchons à trouver $f_Y(y)$ en n'ayant que $Y = g(X)$. Deux méthodes s'offrent à nous :

Nous illustrons la première avec un exemple. Prenons la fonction $Y = g(X)$, pour X une variable aléatoire continue. Par exemple, nous pourrions prendre $g(X) = X^2$. Alors

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \text{ si } y > 0 \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \text{ si } y > 0 \end{aligned}$$

De ceci on extrait la dérivée :

$$\begin{aligned} f_Y(y) &= \frac{dF_Y(y)}{dy} \\ &= \left(\frac{1}{2\sqrt{y}} \right) \frac{dF_X(\sqrt{y})}{d(\sqrt{y})} - \left(\frac{-1}{2\sqrt{y}} \right) \frac{dF_X(-\sqrt{y})}{d(-\sqrt{y})} \\ &= \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}} \text{ si } y > 0 \end{aligned}$$

Cette opération, longue et complexe, est propice à des erreurs. Une autre technique plus simple consiste à passer par la densité de probabilité. Nous passons par 3 étapes :

1. Prendre $y = g(x)$ et calculer toutes les racines $x_i(y)$
2. Pour chacune de ces dérivées, calculer $|g'(x_i)|$
3. finalement $f_Y(y) = \sum_i \frac{f_X(x_i)}{|g'(x_i)|}$

Une preuve complète se trouve à l'annexe [A.1.1](#)

1.2.4 Espérance d'une fonction d'une v.a.

Soit la fonction $g(X)$ où X est une v.a. de densité de probabilité $f_X(x)$. Nous définissons **l'espérance** et *la variance* de $g(X)$ (ou X) comme

- Nous définissons *l'espérance* ci-dessous (si l'intégrale converge) :

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (1.6)$$

² Pas en toute rigueur mathématique, car il faudrait des vérifications en plus (c.f. le polycopié) mais nous considérerons que c'est toujours le cas.

- Pour le cas particulier où $g(X) = X^n$, l'espérance s'appelle alors le **moment d'ordre n**³ :

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx \quad (1.7)$$

- Finalement, le moment d'ordre 1 est appelé la **moyenne**:

- Cas continu :

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f(x) dx \quad (1.8)$$

- Cas discret : comme $P(X = x_i) = p_i$ alors :

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x \left(\sum_i p_i \delta(x - x_i) \right) dx \\ &= \sum_i p_i \int_{-\infty}^{\infty} x \delta(x - x_i) dx = \sum_i p_i x_i \end{aligned} \quad (1.9)$$

- **Variance** :

- Cas continu :

$$VAR[X] = \sigma_X^2 = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x) dx \quad (1.10)$$

- Cas discret :

$$\begin{aligned} E[(X - \mu_X)^2] &= \int_{-\infty}^{\infty} (x - \mu_X)^2 \left(\sum_i p_i \delta(x - x_i) \right) dx \\ &= \sum_i p_i \int_{-\infty}^{\infty} (x - \mu_X)^2 \delta(x - x_i) dx \\ &= \sum_i p_i (x_i - \mu_X)^2 \end{aligned}$$

- **Écart-type** : σ_X est simplement la racine carrée de la variance.

Exemple 3. Prenons

$$g(X) = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n$$

Dans ce cas, on aura que

$$E[g(X)] = a_0 + a_1 E[X] + a_2 E[X^2] + \dots + a_n E[X^n]$$

1.2.5 Transformée

Si nous appliquons l'espérance à la fonction $g(X) = e^{j\omega X}$ ⁴, nous obtenons

la **fonction caractéristique** $\Phi_X(\omega)$:

$$\Phi_X(\omega) = E[e^{j\omega X}] = \int_{-\infty}^{\infty} f_X(x) e^{j\omega x} dx \quad (1.11)$$

Nous notons qu'il s'agit du complexe conjugué de la transformée de Fourier. Nous pouvons lui définir les propriétés suivantes :

³ Attention, le moment pourrait ne pas exister selon $f_X(x)$

⁴ Avec $j = \sqrt{-1}$

- P1. Elle existe toujours (voir A.1.2) P3. $E[X^k] = j^{-k} \left[\frac{d^k \Phi_X(\omega)}{d\omega^k} \right]_{\omega=0}$
- P2. $f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_X(\omega) e^{-j\omega x} d\omega$ P4. $|\Phi_X(\omega)| \leq \Phi_X(0) = 1$

Par la propriété P2., nous voyons que les fonctions $f_X(x)$ et $\Phi_X(x)$ caractérisent tout deux totalement la v.a. et peuvent être utilisées de manière équivalente (pas de perte d'information).

La propriété P3. quant à elle démontre l'intérêt de la fonction caractéristique : il devient bien plus simple de calculer les moments d'ordre élevés une fois que l'on a la fonction caractéristique, car nous n'avons à faire que des dérivées (et non plus des intégrales). La preuve de cette propriété se trouve à l'annexe A.1.3

Il nous est également possible de généraliser la transformée de Fourier à la transformée de Laplace. Nous choisissons pour cela un nombre $s \in \mathbb{C}$ dont la partie complexe peut être nulle, contrairement au nombre purement imaginaire $j\omega$. Nous étendons donc la définition de *fonction caractéristique* à celle de

Fonction génératrice de moment (\simeq transformée de Laplace):

$$\hat{\Phi}_X(s) = E[e^{sX}] = \int_{-\infty}^{\infty} f_X(x) e^{sX} dx \quad (1.12)$$

Avec comme propriété élémentaire que $\hat{\Phi}_X(s = j\omega) = \Phi_X(\omega)$

On utilise aussi parfois la **fonction génératrice de cumulant**⁵ (ou *seconde fonction caractéristique*), qui n'est que le logarithme de la fonction caractéristique :

$$\Psi_X(\omega) = \ln \Phi_X(\omega) = \sum_{i=1}^{\infty} K_i \frac{(j\omega)^i}{i!} \quad (1.13)$$

avec K_i les cumulants : $E[X^1] = K_1$,
 $E[X^2] = K_2 + K_1^2$,
 $E[X^3] = K_3 + K_2^2 + K_1^3, \dots$

Enfin la **fonction génératrice de probabilité** (\equiv espérance de z^X) :

$$G_X(z) = \sum_{i=0}^{\infty} z^i P(X = i) = \sum_{i=0}^{\infty} z^i p_i = E[z^X]$$

Avec les propriétés suivantes :

- P1. Il s'agit, au signe de l'exposant près, à la transformée en z (Z-transform) de la fonction p_k
- P2. Ne s'applique qu'aux v.a. discrètes prenant des valeurs uniformément espacées ($S_X = \{0, 1, 2, 3, \dots\} = \mathbb{N}$) :
- P3. On peut retrouver les probabilités p_k en dérivant $G_X(z)$:

$$p_k = P(X = k) = \frac{1}{k!} \left[\frac{d^k G_X(z)}{dz^k} \right]_{z=0}$$

⁵ Nous ne l'utiliserons pas, elle n'est là qu'à titre d'information

P4. En dérivant par rapport à z (et en évaluant à $z = 1$) on retrouve les moments factoriaux de X (desquels on retrouve tous les moments de X) :

$$E[X(X-1)(X-2)\dots(X-k+1)] = \left. \frac{d^k G_X(z)}{dz^k} \right|_{z=0}$$

P5. $G_X(1) = \sum_{k=0}^{\infty} p_k = 1$

1.2.6 Inégalités

Il arrive souvent que l'on ne connaisse pas complètement la densité de probabilité $f_X(x)$ d'une v.a. Les inégalités suivantes permettent de borner une probabilité impliquant cette v.a. :

Tout d'abord, nous avons **l'inégalité de Markov** :

$$P(X \geq a) \leq \frac{E[X]}{a}, \quad X \geq 0, a > 0 \quad (1.14)$$

Cette inégalité permet de donner une borne *supérieure* (faible certes) sans connaître complètement la densité de probabilité $f_X(x)$ d'une v.a.

Une seconde inégalité intéressante et celle de **Tchebychev** :

$$P(|X - \mu_X| \geq a) \leq \frac{\sigma_X^2}{a^2} \quad (1.15)$$

Celle-ci nous donne une borne *inférieure*, à nouveau sans connaître la densité de probabilité. Les preuves de ces deux inégalités se trouvent aux annexes [A.1.4](#) et [A.1.5](#) respectivement.

1.3 Variables aléatoires

Nous allons évoquer rapidement ici les deux types de variables aléatoires :

- Discrètes : Bernoulli, binomiale, géométrique, de Poisson.
- Continues : uniforme, gaussienne, exponentielle, Gamma, Erlang, Chi-carré

1.3.1 Bernoulli

Un simple test, un jet à pile ou face avec un unique paramètre p , la probabilité que l'événement se passe. Nous créons la *fonction indicatrice d'un événement* A :

$$1_{\{\zeta \in A\}} = I_A(\zeta) = \begin{cases} 0, & \text{si } \zeta \notin A \\ 1, & \text{si } \zeta \in A \end{cases} \quad (1.16)$$

Exemple 4. Si nous souhaitons obtenir un chiffre pair lors d'un lancer de dé, alors :

$$1_{\{\zeta \in A\}} = I_A(\zeta) = \begin{cases} 0, & \text{si } \zeta = 1, 3, 5 \\ 1, & \text{si } \zeta = 2, 4, 6 \end{cases}$$

- $S_X = \{0, 1\}$
- $\begin{cases} P(X = 0) = 1 - p \\ P(X = 1) = P(A) \end{cases}$
- $\mu_X = E[X] = 0(1 - p) + 1 \cdot p = p$
- $\begin{cases} \sigma^2 = E[X^2] - E[X]^2 \\ = E[X^2] - \mu_X^2 \\ = (0^2(1 - p) + 1^2p) - p^2 \\ = p - p^2 = \mathbf{p(1 - p)} \end{cases}$
- $G_X(z) = 1 - p + pz$

1.3.2 Binomiale

En tirant cette même pièce non pas une mais plusieurs fois, nous obtenons une variable aléatoire binomiale : on répète l'expérience de Bernoulli n fois de manière indépendantes. X compte le nombre de succès obtenus :

$$X = I_1 + I_2 + \dots + I_n$$

avec comme support $S_X = \{0, 1, 2, \dots, n\}$ (donc entre aucun succès et que des succès). Nous calculons la probabilité d'un nombre exact de réussites comme :

$$p_k = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (1.17)$$

Avec évidemment $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. De plus, nous trouvons les valeurs suivantes :

- $\mu_X = E[X] = \sum_{k=0}^n k \binom{n}{k} p^k (1 - p)^{n-k} = \mathbf{np}$
- $\sigma_X^2 = E[X^2] - E[X]^2 = \mathbf{np(1 - p)}$
- $G_X(z) = \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} z^k = \sum_{k=0}^n (pz)^k (1 - p)^{n-k} = (\mathbf{pz + (1 - p)})^n$

1.3.3 Géométrique

Cette variable aléatoire compte le nombre d'expériences de Bernoulli (toujours indépendantes) jusqu'au premier succès. Le support est cette fois infini : $S_X = \{0, 1, 2, \dots\}$, et la probabilité d'un nombre précis d'échecs. Deux versions : soit on compte le nombre d'essais avec le premier succès, ou le nombre d'échecs.

Version 1 : X = nombre d'échecs avant le 1^{er} succès.

- $S_X = \{0, 1, 2, \dots\}$
- $E[X] = \frac{1-p}{p}$
- $P(X = k) = p(1 - p)^k$
- $\sigma_X^2 = \frac{1-p}{p^2}$
- $G_X(z) = \sum_{k=0}^{\infty} p(1 - p)^k z^k = p \sum_{k=0}^{\infty} ((1 - p)z)^k = \frac{\mathbf{p}}{\mathbf{1 - (1 - p)z}}$

Version 2 : X = nombre d'essais avant le 1^{er} succès.

- $S_X = \{1, 2, 3, \dots\}$
- $E[X] = \frac{1}{p}, \quad \sigma_X^2 = \frac{1-p}{p^2}$
- $P(X = k) = p(1 - p)^{k-1}$
- $G_{X'}(z) = \frac{pz}{1 - (1 - p)z}$

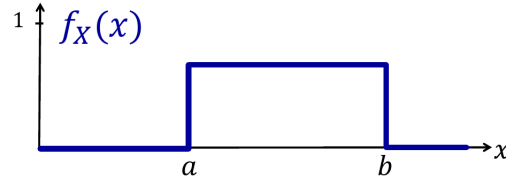


Figure 3: Variable aléatoire uniforme continue

1.3.4 Poisson

Ici, nous comptons le nombre de d'arrivées de clients dans un système pendant une unité de temps, ces arrivées étant soumises à certaines hypothèses. Le seul paramètre est μ (on note "Poisson(μ)") et la probabilité d'avoir un nombre d'événements précis est

- $S_X = \{0, 1, 2, \dots\} = \mathbb{N}$
- $\mu_X = E[X] = \mu$
- $P(X = k) = \frac{\mu^k}{k!} e^{-\mu}$
- $\sigma_X^2 = \mu$
- $G_X(z) = \sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} z^k = e^{-\mu} \sum_{k=0}^{\infty} \frac{(\mu z)^k}{k!} = e^{-\mu} e^{\mu z} = e^{\mu(z-1)}$

Une propriété intéressante des v.a. de Poisson est d'être la limite d'une v.a. binomiale $B(n, p)$ lorsque $n \rightarrow \infty$ et $p \rightarrow 0$. Précisément, soit $X \sim B(n, p)$ et $Y \sim \text{Poisson}(\mu)$. Alors avec $n \rightarrow \infty$ et $p \rightarrow 0$ (tout en gardant np fini) :

$$p_k = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \rightarrow \frac{\mu^k}{k!} e^{-\mu} = P(Y = k)$$

avec $\mu = np$. Une preuve se trouve à l'annexe [A.1.6](#).

1.3.5 Uniforme

Nous arrivons dans les variables aléatoires continues. Une **variable aléatoire continue** sur un intervalle $[a, b] = S_X$ est définie par sa densité de probabilité :

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{sinon} \end{cases} \quad (1.18)$$

Avec les propriétés suivantes, facilement calculées :

- $\mu_X = E[X] = \frac{a+b}{2}$
- $\sigma_X^2 = E[X^2] - \mu_X^2 = \int_a^b x^2 \frac{1}{b-a} dx - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$
- $\Phi_X(\omega) = \int_a^b \frac{1}{b-a} e^{j\omega x} dx = \frac{e^{j\omega b} - e^{j\omega a}}{(b-a)j\omega}$

1.3.6 Gaussienne

La **variable aléatoire gaussienne** est très importante (pensez à la répartition de vos notes après un examen), car elle approxime très bien la somme d'un grand nombre de v.a. dont la densité de probabilité n'est pas totalement connue (très fréquent en

pratique). Elle est bien connue avec sa forme de cloche, le pic se trouvant à la valeur la plus probable. Ses propriétés sont les suivantes :

P1. Ses paramètres sont (μ, σ)

$$P3. F_X(x) = \frac{1}{2\pi} \int_{-\infty}^{-\frac{x-\mu}{\sigma}} e^{-t^2/2} dt$$

$$P2. f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$P4. \mu_X = \mu, \quad \sigma_x^2 = \sigma^2$$

$$P5. \Phi_X(\omega) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) e^{j\omega x} dx = e^{j\mu\omega - \frac{\sigma^2\omega^2}{2}}$$

Regardons un instant ces propriétés. Nous voyons via **P1.** et **P4.** que les paramètres passés à la v.a. ne sont autres que la moyenne (le pic que la cloche) et la variance (l'écartement de la cloche). La propriété **P3.** quant à elle est un peu différente : l'intégrale telle quelle ne peut se résoudre de manière analytique (sans des paramètres fixés). Il faut faire l'intégrale de manière numérique pour trouver la fonction de répartition, et elle peut ne pas exister.

Une v.a. normale est en fait une v.a. gaussienne normalisée (i.e. $\mu = 0$ et $\sigma = 1$).

1.3.7 Exponentielle (Lambda)

La **variable aléatoire exponentielle** mesure le temps s'écoulant entre 2 événements successifs dans un processus de Poisson de taux λ . Nous listons à nouveau ses propriétés :

- $S_X = [0, \infty[$
- Paramétrisée par $\lambda \in \mathbb{R}_+^*$
- $f_X(x) = \lambda e^{-\lambda x}$
- $F_X(x) = 1 - e^{-\lambda x}$
- $\mu_X = \frac{1}{\lambda}$
- $\sigma_X^2 = \frac{1}{\lambda^2}$
- $\Phi_X(\lambda) = \int_0^\infty \lambda e^{-\lambda x} e^{j\omega x} dx = \frac{\lambda}{\lambda - j\omega}$
- Elle est sans mémoire :

$$P(X > t + T | X > t) = P(X > t)$$

1.3.8 Gamma

Avant de parler de la v.a. gamma, nous devons parler de la fonction $\Gamma(\cdot)$, définie pour $u > 0$:

$$\Gamma(u) = \int_0^\infty x^{u-1} e^{-x} dx \quad (1.19)$$

Qui a des propriétés aux allures bien étranges :

- $\Gamma(1/2) = \sqrt{\pi}$
- $\Gamma(u+1) = u\Gamma(u)$
- $\Gamma(u) = (u-1)!$ si $u \in \mathbb{N}$

Avec ceci nous sommes en mesure de définir la **variable aléatoire Gamma** $G(\lambda, \alpha)$:

$$P1. S_X =]0, \infty[$$

$$P4. \mu_X = \frac{\alpha}{\lambda}$$

$$P2. \text{Paramétrisée par } \alpha, \lambda > 0$$

$$P5. \sigma_X^2 = \frac{\alpha}{\lambda^2}$$

$$P3. f_X(x) = \frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

$$P6. \Phi_X(\omega) = \frac{1}{(1-j\omega/\lambda)^\alpha}$$

La v.a. Gamma est très générale, il n'est donc pas surprenant que d'autres variables aléatoires découlent de Gamma. Par exemple, la v.a. $G(\lambda, m)$ avec $m \in \mathbb{N}$ est une **variable aléatoire d'Erlang**, de paramètres (λ, m) . Comme m est entier, **P3.** devient

$$f_X(x) = \frac{\lambda(\lambda x)^{m-1} e^{-\lambda x}}{(m-1)!} \quad (1.20)$$

Un cas particulier de la v.a. d'Erlang est lorsque $m = 1$, on obtient alors une variable *v.a. exponentielle*.

Un autre cas particulier, et lorsque $\lambda = 1/2$ et $\alpha = \nu/2$ avec $\nu \in \mathbb{N}$, nous obtenons une **variable aléatoire Chi-carré** $\chi^2(\nu)$ à ν degrés de liberté. **P3.** devient alors

$$f_X(x) = \frac{x^{(\nu-2)/2} e^{-x/2}}{2^{\nu/2} \Gamma(\nu/2)} \quad (1.21)$$

Module 2: Vecteurs Aléatoires

2.1 Paires de Variables aléatoires

Maintenant que nous avons bien défini le concept de variable aléatoire, il est facile d'étendre la définition à une paire de v.a.

2.1.1 Fonction de répartition jointe et marginale

Nous avons précédemment la probabilité pour une variable $F_X(x) = P(X \leq x)$. Nous définissons maintenant la **fonction de répartition jointe** :

$$F_{XY}(x, y) = P(\{X \leq x\} \cap \{Y \leq y\}) = P(X \leq x, Y \leq y)$$

Avec comme propriétés :

- $F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = 0$
- $F_{XY}(\infty, \infty) = 1$

Similairement, nous avons la **fonction de répartition marginale** :

$$F_X(x) = P(X \leq x) \quad F_Y(y) = P(Y \leq y)$$

Qui est exactement le cas du chapitre précédent, à une variable. Elle est cependant définissable depuis le cas à deux variables, via ses propriétés :

P1. $0 \leq F_{X,Y} \leq 1$

P2. $\lim_{x \rightarrow -\infty} F_{XY}(x, y) = \lim_{y \rightarrow -\infty} F_{XY}(x, y) = 0$ et $\lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} (F)_{XY}(x, y) = 1$

P3. Les fonction de répartition marginales peuvent être obtenues depuis la fonction de répartition jointe :

$$\begin{aligned} \lim_{y \rightarrow \infty} F_{XY}(x, y) &= P(X \leq x, Y \leq \infty) = F_X(x) \\ \lim_{x \rightarrow \infty} F_{XY}(x, y) &= P(X \leq \infty, Y \leq y) = F_Y(y) \end{aligned}$$

2.1.2 Densité de probabilité jointe et marginale

La **densité de probabilité jointe** est similaire :

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \quad (2.1)$$

$$f_{XY}(x, y) = \sum_{i,j} p_{ij} \delta(x - x_i) \delta(y - y_i) \quad (2.2)$$

En utilisant les impulsions de Dirac, nous trouvons les propriétés suivantes pour le cas discret :

$$P4. \sum_{i,j} P(X = x_i, Y = y_i) = 1$$

$$P5. P(X \leq a, Y \leq b) = \sum_{x_i \leq a, y_j \leq b} P(X = x_i, Y = y_i) = F_{XY}(a, b)$$

Et similairement pour le cas continu :

$$P6. \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1$$

$$P7. P(X \leq a, Y \leq b) = F_{XY}(a, b) = \int_{-\infty}^a \int_{-\infty}^b f_{XY}(x, y) dx dy$$

$$P8. P(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{XY}(x, y) dx dy$$

Ceci implique que $P(x < X \leq x + \Delta x, y < Y \leq y + \Delta y) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} (f_{XY}(x, y) \Delta x \Delta y)$

Nous pouvons de la même manière définir la **densité de probabilité marginale**, cas continu ou discret. Pour le cas *continu* :

$$\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} = \frac{dF_{XY}(x, +\infty)}{dx} = \frac{d}{dx} \left[\int_{-\infty}^x \int_{-\infty}^{\infty} f_{XY}(\xi, \eta) d\xi d\eta \right] \\ &= \int_{-\infty}^{\infty} f_{XY}(x, \eta) d\eta \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(\xi, y) d\xi \end{aligned} \quad (2.3)$$

En *discret*, les choses sont légèrement plus simples grâce aux impulsions de Dirac :

$$\begin{aligned} f_X(x) &= \sum_{i,j} P(X = x_i, Y = y_i) \delta(x - x_i) \\ f_Y(y) &= \sum_{i,j} P(X = x_i, Y = y_i) \delta(y - y_i) \\ P(X \leq a) &= \sum_{x_i \leq a} \sum_j P(X = x_i, Y = y_j) \\ P(X = a) &= \sum_j P(X = a, Y = y_j) \end{aligned}$$

2.1.3 Variables aléatoires indépendantes

Deux variables aléatoires X, Y sont indépendantes si

$$F_{XY}(x, y) = F_X(x) F_Y(y)$$

d'où

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

Dans le cas discret, nous avons la même chose avec les probabilités :

$$P(X = a, Y = b) = P(X = a) P(Y = b) \quad (2.4)$$

Il devient donc facile de voir la propriété suivante :

- Si X et Y sont des v.a. indépendantes, alors $g(X)$ et $g(Y)$ sont également indépendants : X

2.1.4 Fonction de deux variables aléatoires

Soient X, Y des variables aléatoires, et $g_1(\cdot), g_2(\cdot)$ des fonction continues. Alors nous définissons, en fonction de la densité jointe des v.a. X, Y :

$$Z = g_1(X, Y) \quad W = g_2(X, Y)$$

Soient (z, w) une paire de valeurs particulières que peut prendre la paire (Z, W) et soient $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$ les m racines respectives du système

$$z = g_1(x, y) \quad w = g_2(x, y)$$

Rappelons que le jacobien de la transformation $(g_1(\cdot), g_2(\cdot))$ est la déterminant :

$$J(x, y) = \det \begin{bmatrix} \frac{\partial g_1(x, y)}{\partial x} & \frac{\partial g_2(x, y)}{\partial x} \\ \frac{\partial g_1(x, y)}{\partial y} & \frac{\partial g_2(x, y)}{\partial y} \end{bmatrix} \quad (2.5)$$

On peut alors montrer que

$$f_{ZW}(z, w) = \sum_i \frac{f_{XY}(x_i, y_i)}{|J(x_i, y_i)|} \quad (2.6)$$

Fonctions linéaires Un cas particulier à prendre en compte est celui des fonctions linéaires : Soient g_1, g_2, Z, W définis comme précédemment, avec la particularité que nous imposons g_i linéaire, autrement dit

$$\begin{aligned} z = g_1(x, y) &= a_{11}x + a_{12}y + b_1 \\ w = g_2(x, y) &= a_{21}x + a_{22}y + b_2 \end{aligned}$$

Ce qui peut s'écrire sous forme vectorielle :

$$\mathbf{z} = A\mathbf{x} + \mathbf{b}$$

En admettant que $\det(A) \neq 0$, la solution de cette équation est unique et donnée par

$$\mathbf{x} = A^{-1}(\mathbf{z} - \mathbf{b})$$

Le jacobien de (x, y) devient alors

$$J(x, y) = \det(A) = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Par conséquent, si le déterminant est non nul, la densité jointe de $\mathbf{Z} = (Z, W)$ peut s'exprimer en fonction de celle de $\mathbf{X} = (X, Y)$:

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{f_{\mathbf{X}}(A^{-1}(\mathbf{z} - \mathbf{b}))}{|\det(A)|} \quad (2.7)$$

Exemple : somme de deux variables aléatoires indépendantes Prenons le cas de $Z = X + Y$ avec $X \perp Y$. À l'annexe A.2.1 se trouve la preuve complète. Notons simplement que cela revient au résultat final

$$f_Z = f_X \star f_Y \quad (2.8)$$

C'est à dire le produit de convolution des densités de probabilité.

2.1.5 Espérance d'une fonction de deux variables

Nous pouvons définir un concept important, la **covariance** :

espérance

$$\text{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = \dots = E[XY] - E[X]E[Y] \quad (2.9)$$

Également le **coefficient de corrélation** :

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (2.10)$$

2.1.6 Transformées (fonction caractéristique jointe)

??

Rappel de la semaine passée

Soient X, Y des v.a. Nous avons que

$$\begin{aligned} F_{XY}(x, y) &= P(X \leq x, Y \leq y) \longrightarrow F_X(x) = P(X \leq x) = F_{XY}(x, +\infty) \\ f_{XY}(x, y) &= \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) \longrightarrow f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy \\ \Phi_{XY}(\omega_1, \omega_2) &= E[e^{j(\omega_1 X + \omega_2 Y)}] = \iint_{-\infty}^{\infty} f_{XY}(x, y) e^{j(\omega_1 x + \omega_2 y)} dx dy \longrightarrow \Phi_X(\omega) = E[e^{j\omega x}] = \Phi_{XY}(\omega, 0) \end{aligned}$$

2.1.7 Densité de probabilité conditionnelle

Deux cas sont possibles comme toujours. Pour le cas *discret* :

faire les deux cas

2.1.8 Espérance conditionnelle

$$E[X|y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx \quad (2.11)$$

$$E[X|y] = \sum_i x P(X = x_i | Y = y) \quad (2.12)$$

Notons que $E[X|y]$ est une fonction déterministe de y . On pourrait dire que $E[X|y] = g(y)$, ce qui implique que $E[X|Y] = g(Y)$ est une v.a. ! Cela revient à ne pas figer Y , et l'on peut ainsi calculer l'espérance de $g(Y)$, et donc de $E[X|Y]$:

$$E_Y[E_X[X|Y]] = \int_{-\infty}^{\infty} E[X|y] f_Y(y) dy \quad (2.13)$$

moar

2.2 Vecteurs aléatoire de dimension n

Nous pouvons élargir tout ce que nous avons vu à $n \geq 2$

$$F_{X_1 X_2 X_3 \dots X_n}(x_1, x_2, x_3, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n) \quad (2.14)$$

moar

2.2.1 V.a. gaussiennes jointes/multivariées

do

$$\Sigma = E[XX^T] - \mu\mu^T = \begin{bmatrix} \sigma_1^2 & Cov[X_1, X_2] \cdots \\ Cov[X_2, X_1] & \sigma_2^2 & \\ \vdots & & \ddots \end{bmatrix} \quad (2.15)$$

Voyons le cas lorsque $m = 2$:

$$\begin{aligned} \underline{X} &\sim \mathcal{N}(\mu, \Sigma), \quad \underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \\ \Sigma &= \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \rightarrow \det \Sigma = \sigma_1^2\sigma_2^2(1 - \rho^2) \rightarrow \Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1 - \rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \\ f_{X_1 X_2}(x_1, x_2) &= \frac{1}{\sqrt{(2\pi)^2 \det \Sigma}} \exp \left(\frac{-1}{2(1 - \rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right) = k \\ \Rightarrow \underbrace{\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2}_{K} &= -2(1 - \rho^2) \ln \left(k 2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2} \right) \end{aligned}$$

Pour tout $|\rho| < 1$: le plan de hauteur k qui coupera la cloche de la Gaussienne, nous pouvons déterminer la forme du cercle/ellipse qui en découle. Par exemple, pour :

- $\rho = 0$: $\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 = K$, ce qui, pour $\sigma_1 > \sigma_2$ donne une ellipse centrée en (μ_1, μ_2) et avec les demis-rayons de taille σ_1 et σ_2 . Si en revanche $\sigma_1 = \sigma_2 = \sigma$, nous obtenons un cercle, le centre est toujours le même.
- $\rho \neq 0$ nous obtenons une ellipse légèrement tournée. Plus ρ tend vers 1, plus l'ellipse s'aplatit et s'approche de la droite $y = x$.

propriétés

2.3 Convergence d'une suite de variables aléatoires

2.3.1 Convergence presque sûre

2.3.2 Convergence en distribution et Théorème Central Limite

2.3.3 Convergence en probabilité et loi faible des grands nombres

$$\forall \epsilon > 0, P\{|X_n - X| > \epsilon\} \rightarrow 0 \text{ pour } n \rightarrow \infty$$

La loi faible des grands nombres : Soit X_1, X_2, \dots, X_n une suite de v.a. i.i.d. de moyenne μ et de variance σ^2 . Alors en définissant

$$Z_n = \frac{1}{n} \sum_{i=1}^n X_i$$

Nous avons

$$\forall \epsilon > 0, P\{|Z_n - \mu| > \epsilon\} \rightarrow 0 \text{ pour } n \rightarrow \infty$$

Module 3: Processus Stochastiques

3.1 Processus stochastiques à temps continu

3.1.1 Définition

Nous connaissons les variables aléatoires, qui mappaient un événement à une valeur (réelle souvent). Par exemple des probabilités sur l'âge d'une personne, ou sur la température. Les processus stochastiques rajoutent une dimension temporelle à cela. Avant on avait $X(\zeta) \rightarrow x$ avec $\zeta \in \Omega$ et $x \in \mathbb{R}$. Nous avons maintenant le **processus stochastique** $X(\zeta, t) \rightarrow x \in \mathbb{R}$. Nous pouvons échantillonner le processus $X(t, \zeta)$ aux temps t_1, t_2, \dots, t_n afin d'avoir n variables aléatoires : $X_1 = X(t_1, \zeta)$, $X_2 = X(t_2, \zeta)$, ..., $X_n = X(t_n, \zeta)$. Le processus stochastique est spécifié par la collection de toutes les fonctions de répartition jointes du nième ordre obtenues pour tout $n \in \mathbb{N}_0 = \mathbb{N} \setminus \{0\}$ et tous les instants $t_1, \dots, t_n \in \mathbb{R}$

Cependant ici nous ne parlerons que des deux premiers ordres, donc la moyenne

$$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x f_X(x; t) dx$$

Notons qu'il ne s'agit plus ici d'une valeur fixe mais d'une fonction : la moyenne en fonction du temps

Exemple 5. Précédemment, nous avons l'espérance de la température, mais ici nous avons la moyenne de température à chaque instant de la journée : la moyenne à midi est différente de celle à 15h.

Pour l'ordre 2, nous avons la *fonction d'auto-corrélation* :

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1)X(t_2)}(x_1, x_2; t_1, t_2) dx_1 dx_2$$

Et de manière similaire la *fonction d'auto-covariance* :

$$C_X(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))] = R_X(t_1, t_2) - \mu_X(t_1)\mu_X(t_2)$$

3.1.2 Exemples

Sinusoïde à phase aléatoire

$$X(t) = a \sin(2\pi f_0 t + \Phi(\zeta))$$

Avec a, f_0 constants et $\Phi(\zeta)$ une v.a. de Bernoulli :

$$\Phi(\zeta) = \begin{cases} 0 & \text{a.p. } 1/2 \\ \pi & \text{a.p. } 1/2 \end{cases}$$

Autrement dit avec probabilité égale, soit $\Phi = 0$ et $X(t) = a \sin(2\pi f_0 t)$ ou $\Phi = \pi$ et $X(t) = a \sin(2\pi f_0 t + \pi) = -a \sin(2\pi f_0 t)$. Nous avons donc un graph $t \rightarrow X(t)$, sur lequel nous avons deux sinusoïdes : un sinus et un sinus inversé.

Prenons nos deux ordres :

$$\begin{aligned}\mu_X(t) &= E[X(t)] = E[X(t)|\Phi = 0]P(\Phi = 0) + E[X(t)|\Phi = \pi]P(\Phi = \pi) \\ &= \frac{1}{2} [a \sin(2\pi f_0 t) - a \sin(2\pi f_0 t)] = 0\end{aligned}$$

Similairement, la fonction d'auto-corrélation :

$$\begin{aligned}R_X(t_1, t_2) &= E[X(t_1)X(t_2)] \\ &= E[X(t_1)X(t_2)|\Phi = 0]\frac{1}{2} + E[X(t_1)X(t_2)|\Phi = \pi]\frac{1}{2} \\ &= \frac{1}{2}a^2 \sin(2\pi f_0 t_1) \sin(2\pi f_0 t_2) + \frac{1}{2}a^2 \sin(2\pi f_0 t_1) \sin(2\pi f_0 t_2) \\ &= a^2 (\cos(2\pi f_0(t_1 - t_2)) - \cos(2\pi f_0(t_1 + t_2)))\end{aligned}$$

Nouveau Phi Reprenons l'exemple précédent avec $\Phi \sim U(0, 2\pi)$ (donc $x = \frac{1}{2\pi}$ entre 0 et 2π , 0 partout ailleurs). La moyenne :

$$\mu_X(t) = E[X(t)] = \int_0^{2\pi} a \sin(2\pi f_0 t + \phi) \frac{1}{2\pi} d\phi = 0$$

Processus gaussien $X(t)$ est un processus gaussien si chaque v.a. $X(t_1), X(t_2), \dots, X(t_n)$ sont des v.a. gaussiennes jointes. Comme vu au module 2 (voir ???), il est donc entièrement spécifié par sa moyenne et sa fonction de covariance. Rappelons sa densité jointe de probabilité : ??? avec ?????

rajouter ref

Processus Markovien Un processus $X(t)$ tel que pour tout

mettre le bordel

Processus à accroissements indépendants

compléter

Processus de Wiener (pas utile pour le moment)

Processus de Poisson Rappel sur Poisson : $X \sim P(\lambda)$ discrète. Nous avons $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$. Nous calculons sa moyenne :

rajouter détails

$$\mu_X(t) = \lambda t$$

Martingale

Modulation d'amplitude de pulse (PAM) Un des signaux les plus importants en communication numérique : une suite de "pulses" $\{g(t - kT), k \in \mathbb{Z}\}$, modulée en amplitude par la suite de variables aléatoires $\{A_k, k \in \mathbb{Z}\}$ et T une constante (\simeq période)

$$X(t) = \sum_{k=-\infty}^{\infty} A_k g(t - kT) \quad (3.1)$$

Nous faisons plusieurs hypothèses :

hypothèses

3.1.3 Stationarité

Stationnarité au sens strict et au sens large

Un processus stochastique $X(t)$ est **stationnaire** (*au sens strict*) (SSS) si et seulement si toutes ses propriétés stochastiques sont indépendantes de l'origine des temps. Donc $\forall n \in \mathbb{N}_0, t_1, t_2, \dots, t_n \in \mathbb{R}$ et $c \in \mathbb{R}$

$$F_{X(t_1), \dots, X(t_n)}(x_1, \dots, x_n; t_1, \dots, t_n) = F_{X(t_1+c), \dots, X(t_n+c)}(x_1, \dots, x_n; t_1 + c, \dots, t_n + c)$$

$$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x f_{X(t)}(x, t) dx = \int_{-\infty}^{\infty} x f_{X(t+c)}(x; t+c) dx = E[X(t+c)] = \mu_X(t+c)$$

$$R_X(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1)X(t_2)}(x_1, x_2; t_1, t_2) dx_1 dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1+c)X(t_2+c)}(x_1, x_2; t_2, t_2) dx_1 dx_2$$

Si la moyenne ne dépend pas du temps, et la fonction d'auto-corrélation ne dépend que de la différence entre les deux instants auxquels elle est évaluée (ci-dessus) alors elle est dit **stationnaire au sens large** (WSS). Notons que $SSS \rightarrow WSS$.

Exemples

- $X(t) = A \cos(2\pi f_0 t) + B \sin(2\pi f_0 t)$ avec $A \perp B$, $\mu_A = \mu_B = 0, \sigma_A^2 = \sigma_B^2 = 1$. Nous trouvons que $\mu_{X(t)} = 0 = \mu_X$ et $R_X(t_1, t_2) = \cos(2\pi f_0(t_1 - t_2)) = R_X(t_2 - t_1)$, donc il est WSS
- Même exemple, mais avec $A \sim \mathcal{N}(0, 1)$ et $B \sim \text{unif}[-\sqrt{3\pi}, \sqrt{3\pi}]$. Si $t = 0$, on a $X(t = 0) = A$ mais pour $t = \frac{1}{4f_0} \rightarrow X(t = \frac{1}{4f_0}) = B$. Donc $P(X(0) \leq x) \neq P(X(\frac{1}{4f_0}) \leq x)$ donc il n'est pas WSS.
- $X(t) = a \cos(2\pi f_0 t + \Phi)$, $\Phi \sim \text{unif}[0, 2\pi]$. on calcule $\mu_X(t) = 0$ et $R_X(t_1, t_2) = \frac{a^2}{2} \cos(2\pi f_0(t_1 - t_2))$

3.1.4 Ergodisme

La semaine prochaine. Voir 3.1.7.

3.1.5 Densité spectrale

Soit $X(t)$ WSS (donc $R_X(t, t - \tau) = R_X(\tau)$). Alors on définit

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-2\pi j f \tau} d\tau \quad (3.2)$$

Bien définie si

$$\int_{-\infty}^{\infty} |R_X(\tau)| d\tau < \infty$$

Il est de plus possible d'obtenir la fonction d'auto-corrélation par la transformée inverse :

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f \tau} df \quad (3.3)$$

En particulier, la puissance des moyennes de $X(t)$ est :

$$E[X^2(t)] = R_X(0) = \int_{-\infty}^{\infty} S_X(f) df \quad (3.4)$$

Les jolies propriétés suivantes :

P1. Si $X(t)$ est WSS, alors $R_X(\tau)$ est paire :

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

P2. $S_X(f) \geq 0$

P3. La fonction $h(f) = S_X(f)/R_X(0) \geq 0$ est une densité de probabilité, car

$$\int_{-\infty}^{\infty} h(f) df = \frac{1}{R_X(0)} \int_{-\infty}^{\infty} S_X(f) df = \frac{R_X(0)}{R_X(0)} = 1$$

3.1.6 Exemples

- Sinusoïde
- Bruit blanc : processus stochastique WSS dont la densité spectrale de puissance est constante : $\frac{N_0}{2}$ pour toutes les fréquences :

$$S_N(f) = \frac{N_0}{2}$$

3.1.7 Ergodisme

Un processus qui va explorer toutes les valeurs que l'espace des valeurs peut prendre. Si nous prenons un processus $X(t) = \{x(t), t \in \mathbb{R}\}$ wss. Sa moyenne est bien entendu définie par $\mu_X = E[x(t, \zeta)] = \sum_i x_i P(X = i)$. Cette moyenne peut être estimée par la moyenne *temporelle* d'une de ces réalisations $x(t)$:

$$\langle x(t) \rangle_T = \frac{1}{T} \int_0^T X(t) dt$$

Au lieu de prendre une

définition

Exemple 6. Prenons un sac de monnaie (alliage métal-nickel) mais certaines biaisées : certaines ont un côté face en plomb, et certaines l'ont en aluminium. 0 si face, 1 si pile.

Là deux manières de faire : si on prend une pièce (disons plombée) et la tirens plusieurs fois, on aura un graphe biaisé (plus souvent pile, donc plus souvent 1). Si on tire souvent une pièce en aluminium, le graphe sera biaisé inversement. Mais ce qui est intéressant est qu'en tirant plus de fois cette même pièce, on n'aura pas un indicateur plus stable, la pièce reste fautive. Si la même pièce est jouée successivement à pile ou face, on aura $\langle x \rangle_T \neq \mu_X$. En d'autres termes, en prenant une trajectoire "alu" on aura un biais, en prenant une trajectoire "plomb" elle sera aussi biaisée (inversement). Il est donc intéressant de prendre une pièce au hasard à chaque tirage.

Notons que $\langle x \rangle_T$ est défini comme l'intégrale d'une v.a., c'est donc aussi une v.a. Nous remarquons aussi que

$$E[\langle x \rangle_T] = E\left[\frac{1}{T} \int_0^T x(t) dt\right] = \frac{1}{T} \int_0^T E[x(t)] dt = \frac{1}{T} \mu_X T = \mu_X$$

En continuant avec notre sac, nous trouvons

$$E[\langle x \rangle_T] = \underbrace{E[\langle x \rangle_T | Pb]}_{9/10} \underbrace{P(Pb)}_{1/2} + \underbrace{E[\langle x \rangle_T | Al]}_{1/10} \underbrace{P(Al)}_{1/2} = 1/2 = \mu_X$$

Notons également que

$$\lim_{T \rightarrow \infty} \text{VAR}[\langle x(t) \rangle_T] = \lim_{T \rightarrow \infty} E[\langle x(t) \rangle_T - \mu_X]^2 = 0$$

???

Exemples

- Signal P.A.M : sa définition est à l'équation 3.1. Notons le cas particulier $A_k \in \{-1, 1\}$ i.i.d. et $D \sim U[0, T]$. Nous calculons

$$\mu_X(t) = E\left[\sum_{k=-\infty}^{+\infty} A_k g(t - kT - D)\right] = \sum_{k=-\infty}^{+\infty} \underbrace{E[A_k]}_0 E[g(t - kT - D)] = 0$$

$$R_x(t_1, t_2) = E\left[\sum_{k=-\infty}^{+\infty} A_k g(t_1 - kT - D) \sum_{l=-\infty}^{+\infty} A_l g(t_2 - lT - D)\right] \quad (3.5)$$

$$= \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} E[A_k A_l] E[g(t_1 - kT - D) g(t_2 - lT - D)] \quad (3.6)$$

$$\forall l \neq k : E[A_k A_l] = E[A_k] E[A_l] = 0 \quad k = l : E[A_k A_l] = E[A_k^2] = 1 \quad (3.7)$$

Prenons la seconde espérance (en prenant le cas particulier où $l = k$) :

$$E[g(t_1 - kT - D) g(t_2 - lT - D)] = \int_0^T g(t_1 - kT - s) g(t_2 - lT - s) \frac{1}{T} ds \quad (3.8)$$

Le tout vaut donc

$$\frac{1}{T} \sum_{k=-\infty}^{+\infty} \int_0^T g(t_1 - kT - s) g(t_2 - kT - s) ds$$

On place le changement de variable $s' = t_1 - kT - s$ donc $s = t_1 - kT - s'$ on trouve

$$= \frac{1}{T} \sum_{k=-\infty}^{\infty} \int_{t_1 - (k+1)T}^{t_1 - kT} g(s') g(s' + t_2 - t_1) ds' = \frac{1}{T} \int_{-\infty}^{\infty} g(s') g(s' + t_2 - t_1) ds'$$

3.2 Application 1 : Distorsion de fréquence en radiocommunication avec un mobile

Pas fait

3.3 Application 2 : Taux d'erreur d'un signal PAM

Module 4: Processus stochastiques à temps discret : analyse au second ordre

Module 5: Processus de Poisson

5.1 Poisson sur \mathbb{R}^+

5.1.1 Processus de comptage

Un processus stochastique $\{N(t), t \in \mathbb{R}^+\}$ (donc à temps continu positif) est appelé un **processus de comptage** si $N(t)$ représente le nombre d'événements se déroulant dans l'intervalle de temps $[0, t]$ avec les propriétés suivantes :

- Processus à temps continu et à valeurs entières non négatives
- $N(0) = 0$ (totalement arbitraire)
- $t_1 < t_2$ alors $N(t_1) \leq N(t_2)$
- Le nombre d'événements dans l'intervalle $]t_1, t_2]$ est donné par $N(t_2) - N(t_1)$

A ce processus de comptage sont associés **la séquence des temps d'arrivée** $\{S(n), n \in \mathbb{N}_0\}$ qui décrit le temps où se produit l'arrivée du n ème événements, et **la séquence des temps entre arrivées** $\{T(n), n \in \mathbb{N}\}$ qui décrit l'intervalle de temps séparant l'arrivée du n ème et du $(n + 1)$ ème événement.

Ces processus sont liés par les relations suivantes :

$$N(t) \geq n \iff S(n) \leq t \quad (5.1)$$

$$S(n) = \sum_{m=0}^{n-1} T(m) \quad (5.2)$$

$$T(n) = S(n+1) - S(n) \quad (5.3)$$

On peut redéfinir $S(n)$:

$$N(t) = \max\{n \in \mathbb{N}_0 : S(n) \leq t\} \quad (5.4)$$

$$S(n) = \inf\{t \in \mathbb{R}^+ : N(t) = n\} \quad (5.5)$$

5.1.2 Définition et loi de probabilité d'un processus de Poisson

Une première définition du processus de Poisson homogène est la suivante :

Définition 1 Un processus de Poisson homogène d'intensité $\lambda > 0$ est un processus de comptage qui satisfait aux trois hypothèses suivantes :

Processus Stochastique à valeurs réelle et temps réel	Processus Stochastique à valeurs réelle et temps discret
$F_{X_1, \dots, X_m}(x_1, \dots, x_m; n_1, \dots, n_m)$	$P(X(n_1) \leq x_1; \dots; X(n_m) \leq x_m)$

H1. Le processus est à accroissement indépendants : des événements se déroulant dans des intervalles de temps disjoints sont indépendants, i.e.

$$P(\{N(t+T)-N(t) = n_0\} \cap \{N(t) = n_1\}) = [P(N(t+T)-N(t) = n_0)P(N(t) = n_1)]$$

H2. Le processus est homogène dans le temps, i.e. ses accroissements sont stationnaires : le nombre d'événements se déroulant dans un intervalle de temps ne dépend que de la longueur.

$$P(N(t+T) - N(t) = n_0) = P(N(T) = n_0)$$

H3. La probabilité que deux événements ou plus se produisent dans un petit intervalle de temps Δt est négligeable par rapport à la probabilité qu'il n'y ait qu'un seul événement. Plus précisément :

$$\begin{aligned} P(N(\Delta t) = 1) &= \lambda \Delta t + o(\Delta t) \\ P(N(\Delta t) \geq 2) &= o(\Delta t) \end{aligned}$$

Une manière alternative de définir le processus :

Définition 2 Un processus de Poisson homogène d'intensité $\lambda > 0$ est un processus de comptage qui satisfait aux deux hypothèses suivantes :

1. Le processus est à accroissements indépendants
2. Le nombre d'événements se déroulant dans un intervalle de temps de longueur T suit une loi de poisson, i.e. pour tout t , $T \geq 0$ et $n \in \mathbb{N}$

$$P(N(t+T) - N(t) = n) = \frac{(\lambda T)^n}{n!} e^{-\lambda T}$$

Ces deux définitions sont parfaitement équivalentes, comme prouvé à l'annexe [A.5.1](#).

Il est intéressant de calculer leur transformée en Z :

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} &= -\lambda p_n(t) + \lambda p_{n-1}(t) + \frac{o(\Delta t)}{\Delta t} \sum_{k=0}^n p_{n-k}(t) \\ z^n \frac{dp_n(t)}{dt} &= -\lambda z^n p_n(t) + \lambda z^n p_{n-1}(t) \\ z^0 \frac{dp_0(t)}{dt} &= -\lambda z^0 p_0(t) \\ \sum_{m=0}^{\infty} z^n \frac{dp_m(t)}{dt} &= -\lambda \sum_{m=0}^{\infty} z^m p_m(t) + \lambda z \sum_{m=1}^{\infty} z^{m-1} p_{m-1}(t) - \lambda \sum_{m=0}^{\infty} z^m p_m(t) + \lambda z \sum_{l=0}^{\infty} z^l p_l(t) \\ \frac{d}{dt} \sum_{n=0}^{\infty} z^n p_n(t) &= \lambda(z-1) \sum_{n=0}^{\infty} p_n(t) \\ \frac{d}{dt} G_{N(t)}(z; t) &= \lambda(z-1) G_{N(t)}(z; t) \rightarrow G_{N(t)} = G_N(z; 0) e^{\lambda t(z-1)} = e^{\lambda t(z-1)} \end{aligned} \quad n \geq 1$$

5.1.3 Propriétés

Ainsi $N(t)$ est une v.a. de Poisson, dont la fonction génératrice est calculée ci-dessus. Sa moyenne est

$$\mu_N(t) = E[N(t)] = \lambda t$$

Avec une variance

$$\sigma_N^2(t) = E[N^2(t)] - \mu_N^2(t) = \lambda t$$

Sa fonction d'auto-corrélation est

$$\begin{aligned}
 R_N(t_1, t_2) &= E[N(t_1)N(t_2)] \neq E[N(t_1)]E[N(t_2)] && \text{car } \neq \\
 &= E[N(t_1)[N(t_2) - N(t_1) + N(t_1)]] \\
 &= E[N(t_1)(N(t_2) - N(t_1)) + N^2(t_1)] \\
 &= E[N(t_1)(N(t_2) - N(t_1))] + E[N^2(t_1)] \\
 &= \underbrace{E[N(t_1)]}_{\lambda t_1} \underbrace{E[N(t_2) - N(t_1)]}_{\lambda(t_2 - t_1)} + \underbrace{E[N^2(t_1)]}_{\lambda^2 t_1^2 + \lambda t_1} \\
 &= \lambda^2 t_1 t_2 - \lambda^2 t_1^2 + \lambda^2 t_1^2 + \lambda t_1 && = \lambda t_1 + \lambda^2 t_1 t_2 \quad t_1 \leq t_2 \\
 & && = \lambda t_2 + \lambda^2 t_1 t_2 \quad t_2 \leq t_1 \\
 &= \lambda^2 t_1 t_2 + \lambda \min\{t_1, t_2\}
 \end{aligned}$$

La *superposition* de M processus de Poisson indépendants $N_i(t)$ de taux λ_i , $1 \leq i \leq M$ est encore un processus de Poisson $N(t)$ de taux $\lambda = \sum_{i=1}^M \lambda_i$

5.1.4 Séquence des temps d'arrivée et inter-arrivées

La *séquence des temps d'arrivée* $\{S(n), n \in \mathbb{N}_0\}$ suit une loi d'Erlang :

add

La *séquence des temps entre arrivées* $\{T(n), n \in \mathbb{N}\}$ est une suite de v.a. exponentielles i.i.d. :

$$\begin{aligned}
 P(T(0) > t) &= P(N(t) = 0) = e^{-\lambda t} \\
 F_{T(0)}(t) &= P(T(0) \leq t) = 1 - e^{-\lambda t} \\
 f_{T(0)}(t) &= \lambda e^{-\lambda t} && t \geq 0 \\
 P(T(n) > t) &= P(S(n+1) - S(n) > t) \\
 &= P(S(n+1) - S(n) > t | S(n) = s) \\
 &= P(N(t+s) - N(s) = 0 | S(n) = s) \\
 &= P(N(t+s) - N(s) = 0) \\
 &= P(N(t) = 0) = e^{-\lambda t}
 \end{aligned}$$

5.2 Processus et bruit impulsif de Poisson sur \mathbb{R}

5.2.1 Processus de Poisson sur \mathbb{R}

On décide maintenant de regarder les temps négatifs (on étend notre définition). Logiquement, les valeurs $N(t)$ $t < 0$ seront négatives. Nous pouvons donc redéfinir :

$$T(n) = S(n+1) - S(n) \quad T(0) = S(1) - S(0) \quad (5.6)$$

5.2.2 Bruit impulsif de Poisson (Shot noise)

Le processus stochastique

$$X_h(t) = \sum_{n=-\infty}^{\infty} h(t - S(n)) \quad (5.7)$$

avec $h(t)$ une fonction déterministe, et $\{S(n), n \in \mathbb{Z}\}$. $X_h(t)$ est appelé **bruit impulsif de Poisson**. Dans le cas particulier où

$$h(t) = 1_{\mathbb{R}_0^+}(t) = \begin{cases} 1 & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases}$$

Alors ce processus est le processus de Poisson défini plus tôt.

Nous prenons la dérivée, grâce à Dirac :

$$X_\delta(t) = \sum_{n=-\infty}^{\infty} \delta(t - S(n))$$

et en appliquant $h(t)$

$$X(t) = (h \star X_\delta)(t)$$

Mettons cela formel. Pour tout processus $Y(t)$ de moyenne $\mu_Y(t)$ et de fonction d'auto-corrélation $R_X(t_1, t_2)$. Alors :

$$Y'(t) = \frac{dY}{dt}(t)$$

Alors

$$\begin{aligned} \mu_{Y'}(t) &= E[Y'(t)] = \frac{d\mu_Y}{dt}(t) \\ R_{Y'}(t_1, t_2) &= E[Y'(t_1)Y'(t_2)] = \frac{\partial^2 R_Y}{\partial t_1 \partial t_2}(t_1, t_2) \end{aligned}$$

Dans le cas particulier $h(t) = \delta(t)$:

$$Y_\delta(t) = \frac{dY_{1_{R_0^+}}}{dt}(t) = \frac{dN}{dt}(t)$$

Et pour $Y(t) = N(t)$:

$$\begin{aligned} E[N(t)] &= \lambda t \rightarrow E\left[\frac{d}{dt}N(t)\right] = \lambda \\ R_N(t_1, t_2) &= \lambda^2 t_1 t_2 + \lambda \min\{t_1, t_2\} \\ \rightarrow R_{N'}(t_1, t_2) &= \frac{\partial^2}{\partial t_1 \partial t_2} [\lambda^2 t_1 t_2 + \lambda \min\{t_1, t_2\}] = \lambda^2 + \underbrace{\lambda \frac{\partial^2}{\partial t_1 \partial t_2} \min\{t_1, t_2\}}_{\delta(t_1 - t_2)} \end{aligned}$$

Alors $X_\delta(t) = N'(t) = \frac{dN}{dt}(t) = \lambda^2 + \lambda \delta(t_1 - t_2)$. Donc $X_\delta(t)$ est WSS, donc on peut calculer sa densité spectrale de puissance :

$$S_{X_\delta}(f) = \lambda^2 \delta(f) + \lambda$$

Rappel, début de semaine 7 Nous partons des temps d'arrivée d'un processus de Poisson de taux λ caractérisée par $\{S(n), n \in \mathbb{Z}\}$. A chaque arrivée nous avons une impulsion, en général qui décroît en forme e^{-x} . Ensuite nous avons x la somme de ces courbes :

$$x_h(t) = \sum_{n=-\infty}^{\infty} h(t - S(n))$$

Nous avons ensuite trouvé le théorème de Campbell :

$$\mu_{x_h} = \lambda \int_{-\infty}^{\infty} h(t) dt \quad \sigma_{x_h}^2 = \lambda \int_{-\infty}^{\infty} h^2(t) dt$$

Revenons à nos moutons

5.3 Application : rapport signal/bruit d'un photodétecteur (fibre optique)

Un signal est converti, grâce à un laser, en un signal optique. La théorie quantique prédit que le processus des arrivées des photons est un processus de Poisson. Mais les charges mettent un temps à parcourir la zone de déplétion de la diode, si bien que nous avons un bruit (impulsif de Poisson) :

$$I_i(t) = e \sum_n h(t - S'(n))$$

Nous avons aussi un courant d'obscurité $I_d(t)$:

$$I_d(t) = e \sum_n h(t - S''(n))$$

Le nombre moyen de paires de charges est proportionnel à la puissance optique incidente $P_s(t)$. Avec $h\nu$ la constante de Planck et μ la fréquence optique :

$$\lambda_i = \frac{\eta P_s}{h\nu}$$

Le courant d'obscurité a une intensité λ_d . Comme les deux processus de Poisson sont indépendants, nous leur somme est un processus de Poisson d'intensité $\lambda = \lambda_i + \lambda_d$ avec $\{S(n)\} = \{S'(n)\} \cup \{S''(n)\}$. Le courant total est converti en tension :

$$V(t) = g(I_i(t) + I_d(t)) = ge \sum_n h(t - S(n))$$

A cette tension s'ajoute une tension de bruit thermique $V_{th}(t)$ (à moyenne nulle et variance σ_{th}^2) :

$$V_{out}(t) = ge \sum_n h(t - S(n)) + V_{th}(t)$$

Le rapport signal sur bruit de la photodiode est défini par (avec P_s constant) :

$$SNR = \frac{E^2[V(t)|\lambda_d = 0]}{VAR[V_{out}(t)]} \quad (5.8)$$

Grâce à Campbell nous trouvons

$$E[V(t)|\lambda_d = 0] = ge\lambda_i \int_{-\infty}^{\infty} h(s) ds = \frac{\eta ge P_s}{h\nu} \int_{-\infty}^{\infty} h(s) ds \quad (5.9)$$

$$VAR[V_{out}(t)] = g^2 e^2 \left(\frac{\eta P_s}{h\nu} + \lambda_d \right) \int_{-\infty}^{\infty} h^2(s) ds + \sigma_{th}^2 \quad (5.10)$$

Ce qui donne :

$$SNR_{PIN} = \frac{\left(\left(\frac{\eta P_s}{h\nu} \right) \int_{-\infty}^{\infty} h(s) ds \right)^2}{\left(\frac{\eta P_s}{h\nu} + \lambda_d \right) \int_{-\infty}^{\infty} h^2(s) ds + \frac{\sigma_{th}^2}{g^2 e^2}} \quad (5.11)$$

Module 6: Chaines de Markov à temps discret

Un processus est dit *Markovien* si l'évolution du processus ne dépend que de sa valeur actuelle et non des valeurs passées. Plus précisément, si $X(t)$ est à valeur discrète, il est Markovien et est appelé *chaîne de Markov* ssi pour toute suite d'instantanés $t_1 < t_2 < \dots < t_k < t_{k+1}$ et toute suite de valeurs $x_1, x_2, \dots, x_k, x_{k+1}$:

$$P(X(t_{k+1})|X(t_1) = x_1, \dots, X(t_k) = x_k) = P(X(t_{k+1})|X(t_k) = x_k) \quad (6.1)$$

Toutes les valeurs que $X(t)$ peut prendre est appelé *l'espace d'état*. Si l'espace d'état est discret, il s'agit d'une **chaîne de Markov** (à temps discret ou continu, en fonction de t), sinon on parle d'un **processus de Markov**, ssi

grosses formules

6.1 Chaînes de Markov à temps discret

Ici, l'espace d'état \mathcal{S} est dénombrable, fini ou non.

6.1.1 Probabilités d'état et de transition

La probabilité que $X(n)$ soit dans l'état $i \in \mathcal{S}$ est une des **probabilités d'état** $\pi_i(n) = P(X(n) = i)$ avec comme propriété intéressante que

$$\sum_{i \in \mathcal{S}} \pi_i(n) = 1 \quad (6.2)$$

Alors que les probabilités $p_{ij}(n) = P(X(n+1) = j | X(n) = i)$ sont appelés les **probabilités de transition à une étape** de l'état i à j au temps n , qui satisfait la relation

$$\sum_{j \in \mathcal{S}} p_{ij}(n) = 1 \quad (6.3)$$

Dès maintenant, nous supposons qu'elles ne dépendent pas du temps n (on dit que la chaîne est **homogène**). On peut aussi les écrire sous la forme d'une **matrice de transition** ou **matrice stochastique** :

$$P = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0i} & \cdots \\ p_{10} & p_{11} & \cdots & p_{1i} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ p_{i0} & p_{i1} & \cdots & p_{ii} & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix} \quad (6.4)$$

Dont la somme des éléments d'une ligne valent toujours 1.

Il est ensuite intéressant de calculer les **probabilités de transition à 2 étapes** de l'état i à l'état j :

$$p_{ij}^{(2)} = P(X(n+1) = j | X(n) = i)$$

compléter

6.1.2 Classification des états

Classe / chaîne irréductible

- Un état j est **accessible** à partir de i si $\exists m \in \mathbb{N}_0 : p_{ij}^{(m)} > 0$, donc il existe une suite de transition (possibles) pour passer de i à j .
- Deux états i, j **communiquent** s'ils sont *accessibles* l'un à l'autre, que l'on note $i \leftrightarrow j$.
- Notons que cette relation ci-dessus est réflexive, symétrique et transitive. Elle définit donc des **classes d'équivalence** : l'ensemble de tous les états qui communiquent entre eux
- Une chaîne de Markov qui ne comporte qu'une seule classe est dite **irréductible**.

États (a)périodiques Un état i est dit **périodique** s'il ne peut être visité qu'à des instants multiples de d , avec d est le plus grand entier avec cette propriété. Autrement dit

$$p_{ii}^{(n)} = 0 \quad \text{si } n \text{ n'est pas multiple de } d$$

Notons qu'il n'est pas obligé qu'il soit visité à tous les instants multiples de d . De plus, tous les états d'une classe ont la même période d

États récurrents/transitoires Soit T_i le **temps de premier passage** de l'état i , défini par

$$T_i = \inf\{n \in \mathbb{N}_0 : X(n) = i\} \quad (6.5)$$

Si l'ensemble est vide, alors $T_i = \infty$. Si T_i est fini, le processus passera un jour ou l'autre par cet état. Inversement, si T_i est infini le processus n'atteindra jamais l'état i .

Définissons maintenant f_i la probabilité que le processus retourne après un certain temps à l'instant i :

$$f_i = P(X(n) = i \text{ pour un certain } n \in \mathbb{N}_0 | X(0) = i) \quad (6.6)$$

Ce qui engendre

$$f_i = \sum_{m \in \mathbb{N}_0} P(T_i = m | X(0) = i) = P(T_i < \infty | X(0) = i)$$

Ainsi, un état i est **récurrent** si $f_i = 1$, et il est **transitoire** si $f_i < 1$

Nous pouvons deviner si un état est récurrent ou transitoire directement depuis les probabilités $p_{ii}^{(n)}$ en introduisant une variable aléatoire V_i , comptant le nombre de passages par l'état i . Notons $1_{\{X(n)=i\}}$ la fonction indicatrice de l'état i , qui vaut 1 lorsque $X(n) = i$, 0 sinon. Alors

$$V_i = \sum_{n=0}^{\infty} 1_{\{X(n)=i\}} \quad (6.7)$$

De quoi on calcule

$$\begin{aligned} E[V_i | X(0) = i] &= \sum_{n=0}^{\infty} E[1_{\{X(n)=i\}} | X(0) = i] \\ &= \sum_{n=0}^{\infty} P(X(n) = i | X(0) = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)} \end{aligned}$$

Également :

$$\begin{aligned} P(V_i < \infty | X(0) = i) &= \dots \\ &= \sum_{n=0}^{\infty} (1 - f_i) p_{ii}^{(n)} \end{aligned}$$

Or, comme $P(V_i < \infty | X(0) = i) \leq 1$, on a que si i est transitoire alors $f_i < 1$ et par notre dernier résultat :

$$\sum_{n=0}^{\infty} \leq \frac{1}{1 - f_i}$$

dans le cas contraire, si i est récurrent, alors $f_i = 1$, ce qui implique que $P(V_i < \infty | X(0) = i) = 0$ et donc V_i est presque sûrement infini. Le théorème suivant en découle :

Proposition 2 L'état i est *récurrent* si et seulement si

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$$

et *transitoire* si et seulement si

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$$

Finalement, nous faisons une distinction supplémentaire entre les états récurrents, lorsque l'on considère l'espérance du temps de premier retour. Si i est récurrent, l'espérance s'écrit

$$E[T_i | X(0) = i] = \sum_{m \in \mathbb{N}_0 \cup \infty} m P(T_i = m | X(0) = i) = \sum_{m \in \mathbb{N}_0} m P(T_i = m | X(0) = i)$$

Si l'espérance est finie, l'état est **récurrent positif**, sinon il est **récurrent nul**.

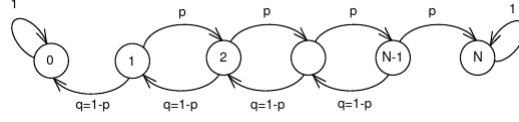
Une chaîne de Markov homogène irréductible, apériodique et dont tous les états sont récurrents positifs est dite **ergodique** (C'est un ergodisme en distribution, pas seulement en moyenne)

espace d'état fini

6.1.3 Des chaînes récurrentes classiques

Chaîne de Markov à deux états Nous avons 2 états : 0 ou 1. La matrice de transition est

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \quad (6.8)$$


 Figure 4: Ruine du joueur sur N états, à deux barrières absorbantes

Nous pouvons calculer facilement la matrice de transition à l'étape n :

$$P^n = \frac{1}{p+q} \begin{bmatrix} q & p \\ q & p \end{bmatrix} + \frac{(1-p-q)^n}{p+q} \begin{bmatrix} p & -p \\ -q & q \end{bmatrix} \quad (6.9)$$

La distribution initiale de probabilité est donc :

$$\begin{aligned} \pi(0) &= [\pi_0(0) \ \pi_1(0)] \rightarrow \pi(n) = \pi(0)P^n \\ &= \left[\underbrace{\frac{1}{p+q} + \frac{(1-p-q)^n}{p+q}(p\pi_0(0) - q\pi_1(0))}_{\pi_0(n)} \quad \underbrace{\frac{p}{p+q} + \frac{(1-p-q)^n}{p+q}(-p\pi_0(0) + q\pi_1(0))}_{\pi_1(n)} \right] \end{aligned}$$

Toute distribution initiale de probabilité $\pi(0)$ tend vers la solution stationnaire, à savoir

$$\pi^* = \begin{bmatrix} \frac{q}{p+q} & \frac{p}{p+q} \end{bmatrix} \quad (6.10)$$

Marche aléatoire unidimensionnelle sur un nombre fini d'états (aussi dit *a barrière bloquante*)

Ce qui est intéressant ici, est que tous les états sont récurrents positifs (pour tout $p, q < 1$), c'est à dire que même si on a une probabilité énorme d'aller vers la droite, il existe encore une possibilité de retourner à 0

6.1.4 Temps d'atteinte

Nous avons défini plus tôt le temps de premier retour à l'état i :

$$T_i = \inf\{n \geq 1 : X(n) = i\}$$

Nous définissons maintenant le **temps d'atteinte** de l'ensemble d'états $\mathcal{A} \in \mathcal{S}$:

$$H_{\mathcal{A}} = \inf\{n \in \mathbb{N} : X(n) \in \mathcal{A}\} \quad (6.11)$$

Notons que si $X(n) \notin \mathcal{A} \ \forall n \in \mathbb{N}$, on pose $H_{\mathcal{A}} = \infty$. Nous nous intéressons pour commencer aux probabilités que, partant d'un état $i \in \mathcal{S}$ on atteigne le sous-ensemble \mathcal{A} :

$$h_{i\mathcal{A}} = P(H_{\mathcal{A}} < \infty | X(0) = i) = P(\exists n \in \mathbb{N} : X(n) \in \mathcal{A} | X(0) = i) \quad (6.12)$$

Exemple de la Ruine du joueur sur $\{0, 1, \dots, N\}$ Nous avons un graph à $N+1$ états, comme représenté à la Figure 4. Nous commençons par imaginer que la

probabilité $p = \frac{1}{2}$ (jeu pile/face). Nous calculons immédiatement, pour $\mathcal{A} = \{0\}$

$$h_{i0} = \begin{cases} 1 & i = 0 \\ 0 & i = N \end{cases}$$

Pour calculer tous les cas intermédiaires, nous appliquons le théorème suivant :

Théorème 5: Le vecteur des probabilités $\mathbf{h}_{\mathcal{A}} = [h_{i\mathcal{A}}, i \in \mathcal{A}]$ est la solution minimale non négative du système d'équations linéaire

$$\begin{aligned} h_{i\mathcal{A}} &= 1 & \text{si } i \in \mathcal{A} \\ h_{i\mathcal{A}} &= \sum_{j \in \mathcal{S}} p_{ij} h_{j\mathcal{A}} & \text{si } i \notin \mathcal{A} \end{aligned} \quad (6.13)$$

Démonstration : voir annexe todo

Ainsi, pour tous les $1 \leq i \leq N - 1$ nous avons

$$h_{i0} = \sum_{j \in \mathcal{S}} p_{ij} h_{j0} = p \cdot h_{i+1,0} + q \cdot h_{i-1,0} \iff p \cdot h_{i+1,0} - h_{i,0} + q \cdot h_{i-1,0} = 0$$

En posant α, β les racines distinctes de $px^2 - x + q = 0$ nous avons

$$h_{i,0} = K_a \alpha^i + K_b \beta^i$$

Nous trouvons les racines :

$$\alpha, \beta = \frac{1 \pm \sqrt{1 - 4(p(1 - p))}}{2p} = \frac{1 \pm (1 - 2p)}{2p} = \left\{ 1, \frac{1 - p}{p} \right\} = \{\alpha, \beta\}$$

Si $p \neq 1/2$.

Rajouter conclusion

Ruine du joueur sur \mathbb{N} Ici nous ne pouvons plus terminer à N , car notre adversaire a une fortune illimitée.

rajouter ?

Nous nous intéressons d'autre part au temps moyen que prend $X(n)$ pour atteindre le sous-ensemble \mathcal{A} :

$$\mu_{i\mathcal{A}}^H = E[H_{\mathcal{A}} | X(0) = i] = \sum_{m \in \mathbb{N} \cup \{\infty\}} m P(H_{\mathcal{A}} = m | X(0) = i)$$

Que l'on peut expliciter comme :

$$\mu_{i\mathcal{A}}^H = \sum_{m \in \mathbb{N}} m P(X(m) \in \mathcal{A} \text{ et } X(n) \notin \mathcal{A} \text{ pour } 0 \leq n \leq m - 1 | X(0) = i) \quad (6.14)$$

Ainsi, nous avons le théorème suivant :

Théorème Le vecteur des temps moyen d'atteinte $\boldsymbol{\mu}_{\mathcal{A}}^H = [\mu_{i\mathcal{A}}^H, i \in \mathcal{S}]$ est la solution minimale non négative du système d'équations linéaire

$$\begin{aligned} \mu_{i\mathcal{A}}^H &= 0 & \text{si } i \in \mathcal{A} \\ \mu_{i\mathcal{A}}^H &= 1 + \sum_{j \notin \mathcal{A}} p_{ij} \mu_{j\mathcal{A}}^H & \text{si } i \notin \mathcal{A} \end{aligned} \quad (6.15)$$

6.1.5 Chaînes réversibles

Définition Supposons une chaîne de Markov ergodique à l'état stationnaire, et qu'à partir d'un temps n nous considérons la séquence d'états $X(n), X(n-1), X(n-2), \dots, X(0)$ c'est à dire la chaîne originelle en remontant le temps (**backward chain**). C'est également une chaîne de Markov car nous pouvons montrer que $P(X(n) = j | X(n+1) = i, X(n+1) = k, \dots) = P(X(n) = j | X(n+1) = i)$. De plus, en utilisant Bayes on peut calculer les probabilités de transitions de la chaîne "renversée" :

$$\tilde{p}_{ij} = \frac{p_{ji}\pi_j^*}{\pi_i^*} \quad (6.16)$$

Module 7: Chaînes de Markov à temps continu

7.1 Introduction

7.1.1 Première définition

Nous considérons maintenant des temps continu, c'est à dire $\{x(t), t \in \mathbb{R}^+\}$. Cet ensemble est une chaîne de Markov si, $\forall t_0 < t_1 < t_2 < \dots t_n \in \mathbb{R}^+ \forall i_0, i_1, \dots, i_n \in \mathcal{S}_X$ nous avons

$$P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}, \dots, X(t_0) = i_0) = P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}) \quad (7.1)$$

et que le processus est homogène pour tout $s, t \in \mathbb{R}^+$ et $i, j \in \mathcal{S}$

$$P(X(t+s) = j | X(s) = i) = p_{ij}(t) = P(X(t) = j | X(0) = i)$$

Les probabilités $p_{ij}(t)$ sont posées dans une matrice de transition $P(t)$, et les probabilités d'état $\pi_i(t) = P(X(t) = i)$ sont posées dans le vecteur d'état $\boldsymbol{\pi}(t) = [\pi_0(t) \ \pi_1(t) \ \pi_2(t) \ \dots]$.

A cette chaîne sont associés des processus stochastiques, comme pour les processus de comptage, Tout d'abord la *séquence des temps de transition* (jump times) $\{S(n), n \in \mathbb{N}\}$ avec $0 \leq S(0) < S(1) < \dots < S(n) < S(n+1) < \dots$ où $S(n)$ est le temps où se produit la n ème transition d'un état à l'autre. Ensuite la *séquence des temps de séjour* (holding times) $\{T(n), n \in \mathbb{N}\}$ qui décrit l'intervalle de temps pendant lequel le processus reste dans un même état (donc temps entre deux transitions successives). Les trois processus X, S, T sont liés :

liaison

7.1.2 Distribution des temps de séjour

Nous montrons facilement, en utilisant les propriétés d'un processus Markovien, que les temps de séjour sont une suite de variables aléatoires indépendantes (pas identiquement distribuées), avec une distribution exponentielle $\exp(\nu_i)$, avec

$$\nu_i = \frac{1}{E[T_i(n)]}$$

Soit l'inverse de la durée moyenne pendant laquelle le processus va séjourner en i .

7.1.3 Chaînes de Markov induites

Maintenant que nous avons le temps moyen que nous restons à chaque état, reste à déterminer la probabilité d'atteindre un autre état. En effet, chaque fois que la chaîne quitte l'état i , elle doit entrer en état $j \in \mathcal{S}$ avec une probabilité de \hat{q}_{ij} qui vérifie trivialement

$$\hat{q}_{ii} = 0 \quad \sum_{j \in \mathcal{S}} \hat{q}_{ij} = 1 \quad (7.2)$$

A partir de $\{X(t), t \in \mathbb{R}^+\}$ on construit le processus à temps discret $n \in \mathbb{N}$:

$$\hat{X}(n) = X(S(n)) \quad (7.3)$$

qui est donc la séquence des différentes valeurs prises par $X(t)$. C'est clairement une chaîne de Markov à temps discret dont les probabilités de transition $i \rightarrow j$ sont les $\hat{q}_{ij} = P(\hat{X}(n) = j | \hat{X}(n-1) = i)$

7.1.4 Équations de Kolmogorov

Dans le cas continu, les équations de Chapman-Kolmogorov s'écrivent

$$p_{ij}(t_1 + t_2) = \sum_{k \in \mathcal{S}} p_{ik}(t_1) p_{kj}(t_2) \quad (7.4)$$

Nous définissons aussi les probabilités d'état :

$$\begin{aligned} \pi_i(t) &= P(X(t) = i) \\ \pi_i(t+s) &= \sum_{k \in \mathcal{S}} \underbrace{P(X(t+s) = i | X(t) = k)}_{p_{ki}(s)} \underbrace{P(X(t) = k)}_{\pi_k(t)} = \sum_{k \in \mathcal{S}} \pi_k(t) p_{ki}(s) \end{aligned}$$

Transition

$$Q = [q_{ij}]_{i,j \in \mathcal{S}_X} \quad (7.5)$$

Pour une chaîne à temps discret, si elle est irréductible ($i \leftrightarrow j \forall i, j \in \mathcal{S}_X$), apériodique et récurrente positive, alors

$$\exists! \pi^* \text{ telle que } \pi = \pi P \text{ et } \lim_{n \rightarrow \infty} \pi(n) = \pi^*$$

Pour le temps continu, il suffit de vérifier si la chaîne induite est irréductible et récurrente positive.

Module 8: Files d'attente

A Preuves

A.1 Chapitre 1

A.1.1 Fonction d'une variable aléatoire

Comme nous avons $Y = g(x)$, alors on peut prendre un y , une valeur particulière de Y , et de ceci en découlent x_1, x_2, \dots, x_m les m racines respectives de l'équation $y = g(x_i)$ avec $1 \leq i \leq m$.

Nous voulons poser A l'événement

$$A = \{\zeta | y < Y(\zeta) \leq y + \Delta y\}$$

Nous choisissons une espace de plus d'une valeur, car comme démontré plus haut la probabilité $P(Y = y)$ est nulle. Nous savons par les propriétés établies plus tôt que la probabilité que cet événement se produise est

$$P(A) = P(y \leq Y \leq y + \Delta y) = f_Y(y)|\Delta y| \quad (\text{A.1.1})$$

si $\Delta y \rightarrow 0$. De plus, comme $Y(\zeta) = g(X(\zeta))$ et que $g(x)$ est continue, nous pouvons réécrire A comme

$$\begin{aligned} A &= \{\zeta | y < g(X(\zeta)) \leq y + \Delta y\} \\ &= \{\zeta | x_1 < X(\zeta) \leq x_1 + \Delta x_1\} \cup \\ &\quad \{\zeta | x_2 < X(\zeta) \leq x_2 + \Delta x_2\} \cup \dots \cup \\ &\quad \{\zeta | x_m < X(\zeta) \leq x_m + \Delta x_m\} \end{aligned}$$

Ce qui nous permet d'écrire :

$$P(A) = f_X(x_1)|\Delta x_1| + f_X(x_2)|\Delta x_2| + \dots \quad (\text{A.1.2})$$

En se rappelant que $g'(x_i) = \lim_{\Delta x_i \rightarrow 0} \frac{\Delta y}{\Delta x_i}$ et en égalant (A.1.1) et (A.1.2), nous trouvons

$$f_Y(y) = \sum_i \frac{f_X(x_i)}{|g'(x_i)|} \quad (\text{A.1.3})$$

A.1.2 Existence de la fonction caractéristique

$$\Phi_X(\omega) \leq |\Phi_X(\omega)| = \left| \int_{-\infty}^{\infty} e^{j\omega x} f_X(x) dx \right| \leq \int_{-\infty}^{\infty} |e^{j\omega x} f_X(x)| dx = \int_{-\infty}^{\infty} f_X(x) dx = 1$$

A.1.3 Moment d'ordre k avec la fonction caractéristique

Nous dérivons $\Phi_X(\omega)$ par rapport à ω et l'évaluons en $\omega = 0$ (maximum de la fonction caractéristique) :

$$\left. \frac{d\Phi_X(\omega)}{d\omega} \right|_{\omega=0} = \left. \int_{-\infty}^{\infty} jx e^{j\omega x} f_X(x) dx \right|_{\omega=0} = j \int_{-\infty}^{\infty} x f_X(x) dx = jE[X] \quad (\text{A.1.4})$$

Ce qui donne

$$E[X] = \frac{1}{j} \left. \frac{d\Phi_X(\omega)}{d\omega} \right|_{\omega=0} \quad (\text{A.1.5})$$

Et en répétant l'opération, nous voyons rapidement que

$$E[X^k] = \frac{1}{j^k} \left. \frac{d^k \Phi_X(\omega)}{d\omega^k} \right|_{\omega=0} \quad (\text{A.1.6})$$

A.1.4 Inégalité de Markov

Lorsque la v.a. X ne prend pas de valeurs négatives, et ne limitant $a > 0$, on observe :

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^a x f_X(x) dx + \int_a^{\infty} x f_X(x) dx \\ &\geq \int_a^{\infty} x f_X(x) dx \geq \int_a^{\infty} a f_X(x) dx = a \int_a^{\infty} f_X(x) dx = a P(X \geq a) \end{aligned} \quad (\text{A.1.7})$$

Ce qui implique que e

$$P(X \geq a) \leq \frac{E[X]}{a} \quad (\text{A.1.8})$$

Toujours en conservant les conditions $X \geq 0$, $a > 0$.

A.1.5 Inégalité de Chebychev

En considérant la v.a. $D = (X - \mu_X)^2$, dont la moyenne est $E[D] = E[(X - \mu_X)^2] = \sigma_X^2$ (c'est à dire la variance de X). Posons $b = \sqrt{a}$ dans l'inégalité de Markov (voir A.1.4) :

$$P(D \geq b^2) \leq \frac{\sigma_X^2}{b^2}$$

Or, comme $P(D \geq b^2) = P((X - \mu_X)^2 \geq b^2) = P(|X - \mu_X| \geq b)$ nous avons

$$P(|X - \mu_X| \geq b) \leq \frac{\sigma_X^2}{b^2} \quad (\text{A.1.9})$$

A.1.6 Limite d'une v.a. binomiale vers Poisson

Partons avec $k = 0$; comme $\binom{n}{0} = 1$, on a

$$p_0 = (1-p)^n \stackrel{\mu=np}{=} (1-\mu/n)^n \xrightarrow{n \rightarrow \infty} e^{-\mu}$$

Pour $k > 0$:

$$\frac{p_{k+1}}{p_k} = \frac{\binom{n}{k+1} p^{k+1} (1-p)^{n-k-1}}{\binom{n}{k} p^k (1-p)^{n-k}} = \dots = \frac{\mu(1-k/n)}{(k+1)(1-\mu/n)} \xrightarrow{n \rightarrow \infty} \frac{\mu}{k+1}$$

Ce qui donne la limite

$$p_{k+1} \rightarrow \frac{\mu}{k+1} p_k \rightarrow \frac{\mu}{k+1} \frac{\mu}{k} p_{k-1} \rightarrow \dots \rightarrow \frac{\mu^{k+1}}{(k+1)!} e^{-\mu}$$

A.2 Chapitre 2

A.2.1 La somme de deux v.a. indépendantes est le produit de convolution

definitely
todo

A.3 Chapitre 3

A.4 Chapitre 4

A.5 Chapitre 5

A.5.1 Les deux définitions du processus de Poisson homogène sont équivalentes