M1 E3A - Voie André Ampère

Module 451

Correction de TD

Version du 25 novembre 2023

Un cours de:
Cécile Durieu

Rédigé et amélioré par:
Pierre-Antoine Comby
(basé sur le travail de ?)

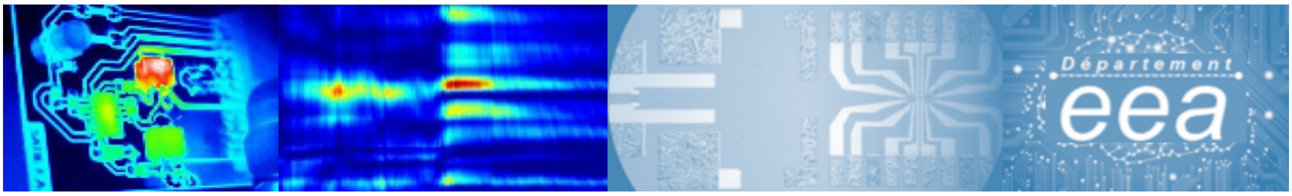


Table des matières

TD1 Variable aléatoire scalaires

Optimisation d'un laminoir

1. Y désigne la longueur de barres perdue, qui est de X (la totalité de la barre) si la barre est trop courte, ou de X-l si la barre est trop longue.

$$Y = \begin{cases} X & \text{si } X < l \\ X - l & \text{si } X > l \end{cases}$$

2.

$$\begin{aligned} m_Y &= E_Y[Y] = E_X[Y] \\ &= \int_0^l x f_X(x)dx + \int_l^\infty (x-l) f_X(x)dx \\ &= \int_0^\infty x f_X(x)dx - \int_l^\infty l f_X(x)dx \\ m_Y &= m_X - l \int_l^\infty f_X(x)dx \end{aligned}$$

3. On suppose que X est une variable aléatoire gaussienne, alors

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$$

- m_X représente la moyenne de X, c'est-à-dire la valeur sur laquelle la gaussienne est centrée
- σ_X représente la dispersion des valeurs autour de m_X. En effet, à $m_X \pm \sigma_X$, on a $f_X(x) = 0.6\frac{1}{\sigma_X\sqrt{2\pi}}$.

$$\begin{aligned} \int_{m_X-\sigma_X}^{m_X+\sigma_X} f_X(x)dx &= 0.68 \\ \int_{m_X-2\sigma_X}^{m_X+2\sigma_X} f_X(x)dx &= 0.95 \\ \int_{m_X-3\sigma_X}^{m_X+3\sigma_X} f_X(x)dx &= 0.99 \end{aligned}$$

Le modèle est valable si la probabilité d'avoir des événements "non-physiques" reste faible, c'est-à-dire si $m_X - 3\sigma_X > 0$ de sorte à ce que $P(X < 0) < 0.5\%$.

4. Sans contrainte sur l'intervalle de variation de m_X, une condition nécessaire à l'obtention

d'un minimum m_X^{opt} est $\frac{dm_Y}{dm_X} = 0$.

$$\frac{dm_Y}{dm_X} = 1 - l \int_l^{+\infty} \frac{\partial f_X(x, m_X)}{\partial m_X} dx$$

$$\text{Or, on a ici une gaussienne, donc } \frac{\partial f_X}{\partial m_X} = -\frac{\partial f_X}{\partial x}$$

$$\begin{aligned} \frac{dm_Y}{dm_X} &= 1 + l \int_l^{\infty} \frac{df_X(x, m_X)}{dx} dx \\ &= 1 + l(f_X(+\infty) - f_X(l)) = 1 - l f_X(l) \end{aligned}$$

$$\text{Ainsi, } \frac{dm_Y}{dm_X} = 0 \quad \Leftrightarrow \quad f_X(l) = \frac{1}{l}$$

$$\begin{aligned} f_X(l) = \frac{1}{l} &\Leftrightarrow 1 = \frac{l}{\sqrt{2\pi}\sigma_X} e^{-\frac{1}{2}(\frac{x-m_X}{\sigma_X})^2} \\ &\Leftrightarrow m_X = l \pm \sigma_X \sqrt{-2\ln \frac{l}{\sqrt{2\pi}\sigma_X}} \text{ avec } l > \sqrt{2\pi}\sigma_X \end{aligned}$$

Pour qu'il s'agisse d'un minimum, il faut que $\frac{d^2 m_Y}{d_X^m} > 0$.

$$\begin{aligned} \frac{d^2 m_Y}{d_X^m} &= -l \frac{df_X(l)}{dm_X} \\ &= \frac{-l}{\sqrt{2\pi}\sigma_X^3}(l - m_X) e^{-\frac{1}{2}(\frac{x-m_X}{\sigma_X})^2} \end{aligned}$$

$$\text{et on obtient } \frac{d^2 m_Y}{d_X^m}|_{m_X+} > 0 \text{ et } \frac{d^2 m_Y}{d_X^m}|_{m_X-} \qquad < 0$$

Le minimum de m_Y est donc bien atteint pour $m_{Xopt} = l + \sigma_X \sqrt{-2\ln \frac{l}{\sqrt{2\pi}\sigma_X}}$

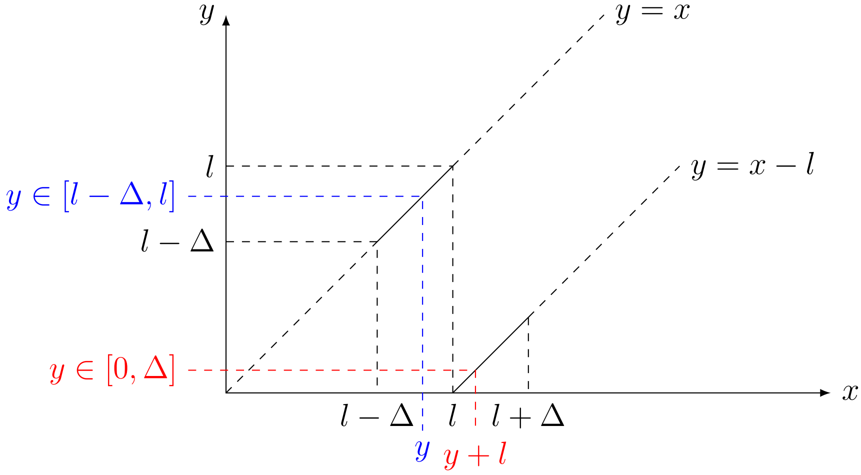
5. Application numérique :

$$m_X = 1 + 0,01\sqrt{2\ln \frac{1}{\sqrt{2\pi}0,01}} = 1,027m$$

Remarque : on obtient $l \approx l + 3\sigma_X$, donc le modèle est valable, en accord avec la question 4.

Changement de variable aléatoire

1. Traçons l'évolution de la VA Y en fonction de X :



Avec l'aide du dessin, on voit immédiatement que :

- si $y \in [0, \Delta[$, alors $F_Y(y) = P[0 \leq Y < y] = P[l \leq x < y + l] = F_X(y + l) - F_X(l)$
- si $y \in [l-\Delta, \Delta]$, alors $F_Y(y) = P[0 \leq Y < y] = F_X(l+\Delta) - F_X(l) + F_X(y) - F_X(l-\Delta)$
- sinon, $F_Y(y) = \text{cste}$.

Ainsi, on en déduit la densité de probabilité en dérivant par rapport à y :

$$f_Y(y) = \begin{cases} f_X(y) & \text{si } y \in [l - \Delta, \Delta] \\ f_X(y + l) & \text{si } y \in [0, \Delta[\\ 0 & \text{sinon} \end{cases}$$

2. On a le changement de variable $g : X \Rightarrow Y = g(X)$ où :

$$g(X) = \begin{cases} X + l & \text{si } Y = X - l \in [0, \Delta] \\ X & \text{si } Y = X \in [l - \Delta, \Delta] \end{cases}$$

On utilise la formule générale du changement de variables :

$$f_Y(y) = F_X(x)|\frac{dx}{dy}||_{x, g(x)=y}$$

On retrouve ainsi le résultat précédent.

3. On considère que la variable X est uniformément répartie sur l'intervalle $[l - \Delta, l + \Delta]$. Ainsi, $f_X(x) = C$ (constante) sur cet intervalle et comme on a :

$$\int_{-\infty}^{+\infty} f_X(x)dx = \int_{l-\Delta}^{l+\Delta} f_X(x)dx = 2\Delta C = 1$$

alors

$$f_X(x) = \begin{cases} \frac{1}{2\Delta} & \text{si } x \in [l - \Delta, l + \Delta] \\ 0 & \text{sinon} \end{cases}$$

d'où

$$f_Y(y) = \begin{cases} \frac{1}{2\Delta} & \text{si } y \in [0, \Delta] \cup [l - \Delta, l] \\ 0 & \text{sinon} \end{cases}$$

Propriété importante : Si une densité de probabilité a une symétrie en x_0 et que sa moyenne existe, sa moyenne est sur l'axe de symétrie, donc vaut x_0.
Preuve :

$$\begin{aligned}\int_{-\infty}^{+\infty} x f_X(x)dx &= \int_{-\infty}^{+\infty}(y+x_0)f_X(y+x_0)dy \\ &= x_0\int_{-\infty}^{+\infty} f_X(y+x_0)dy + \int_{-\infty}^{+\infty} y f_X(y+x_0)dy \\ &= x_0\times 1+0 = x_0\end{aligned}$$

Ainsi, on obtient $m_Y = \frac{l}{2}$

4. On considère le changement de variable : $\Phi \to Y = \cos(\Phi)$.
Comme $\cos(\Phi)\in[-1,1]$, on a pour $y<-1, F_Y(y)=0$ et pour $y\geq 1, F_Y(y)=1$.
Pour $y\in[-1,1]$,

$$\begin{aligned}F_Y(y) &= P(\cos(\Phi)<y) \\ &= P(-pi\leq\Phi<-\arccos(y)) + P(\arccos(y)\leq\Phi<\pi) \\ &= F_\Phi(-\arccos(y)) - F_\Phi(\arccos(y)) + F_\Phi(\pi) - F_\Phi(-\pi) \\ &= F_\Phi(-\arccos(y)) - F_\Phi(\arccos(y)) + 1 \\ f_Y(y) &= \frac{1}{\sqrt{1-y^2}}f_\Phi(-\arccos(y)) + \frac{1}{\sqrt{1-y^2}}f_\Phi(\arccos(y))\end{aligned}$$

5. On utilise le résultat général sur les changements de VA avec $\phi = \arccos(y)$ si $y>0$ et $\phi = -\arccos(y)$ si $y<0$

$$\begin{aligned}f_Y(y) &= f_\Phi(\phi)|\frac{d\phi}{dy}||_{\phi,g(\phi)=y} \\ &= \begin{cases}0 & \text{si } y\notin[-1,1] \\ f_\Phi(\phi)|\frac{d\phi}{dy}||_{\phi=\arccos(y)} + f_\Phi(\phi)|\frac{d\phi}{dy}||_{\phi=-\arccos(y)} & \text{si } y\in[-1,1]\end{cases} \\ &= \begin{cases}0 & \text{si } y\notin[-1,1] \\ \frac{1}{\sqrt{1-y^2}}f_\Phi(-\arccos(y)) + \frac{1}{\sqrt{1-y^2}}f_\Phi(\arccos(y)) & \text{si } y\in[-1,1]\end{cases}\end{aligned}$$

6. Si la VAΦ est uniformément répartie sur $[-\pi,\pi]$, alors on a

$$f_\Phi(\phi) = \begin{cases}\frac{1}{2\pi} & \text{si } \phi\in[-\pi,\pi] \\ 0 & \text{sinon}\end{cases}$$

$$f_Y(y) = \begin{cases}\frac{1}{\pi\sqrt{1-y^2}} & \text{si } y\in[-1,1] \\ 0 & \text{sinon}\end{cases}$$

TD2 Étude d'un couple de VA

(X, Y) est un couple de variables aléatoires uniformément réparti sur un disque D centré en 0 et de rayon R_{max}.

1. Donner la densité de probabilité conjointe des VA X et Y.
   Les variables aléatoires X, Y sont uniformément réparties sur le disque donc $\forall(x, y) \in D$, $f_{XY}(x, y) = A$ constante.
   Or, $\int\int_{\mathbb{R}^2} f_{XY}(x, y)dxdy = 1$ donc $\int\int_D f_{XY}(x, y)dxdy = 1$ d'où $\pi R_{max}^2 A = 1$

$$f_{XY}(x, y) = \begin{cases} \frac{1}{\pi R_{max}^2} & \text{si} \quad (x, y) \in D \\ 0 & \text{sinon} \end{cases}$$

2. Discuter de l'indépendance de X et Y.
   - Est-ce qu'une information sur X implique une information sur Y ?
     Pour la réalisation x_1 de X, on voit que
     $$y_1 \in [y_1^-, y_1^+] = [-\sqrt{R_m^2 - x_1^2}, \sqrt{R_m^2 - x_1^2}]$$
     Ainsi, X et Y ne peuvent pas être indépendantes.
   - On sait que : si X et Y sont indépendantes, alors la ddp est séparable.
     $$f_{XY}(x, y) = \begin{cases} \frac{1}{\pi R_{max}^2} & \text{si} \quad x^2 + y^2 \leq R_{max}^2 \\ 0 & \text{sinon} \end{cases}$$
     On ne peut pas l'écrire sous la forme (fonction de x) x (fonction de y) à cause de la condition $x^2 + y^2 \leq R_{max}^2$ donc X et Y ne sont pas indépendantes.
   - On sait que : si X et Y sont indépendantes, alors $f_{XY}(x, y) = f_X(x)f_Y(y)$. Exhibons un contre-exemple.
     Prenons un point donné par
     $$|x| \leq R_{max}, |y| \leq R_{max}, \text{ tel que } x^2 + y^2 > R_{max}^2$$
     (Dans le carré mais pas dans le cercle). Ainsi, $f_{XY}(x, y) = 0$ alors que $f_X(x) \neq 0$ et $f_Y(y) \neq 0$, donc X et Y ne sont pas indépendantes.
   - On sait que : si X et Y sont indépendantes, alors $F_{XY}(x, y) = F_X(x)F_Y(y)$.
     $$F_{XY}(-\frac{R_{max}}{\sqrt{2}}, \frac{R_{max}}{\sqrt{2}}) = P(X < -\frac{R_{max}}{\sqrt{2}}, Y < \frac{R_{max}}{\sqrt{2}}) = 0$$
     alors que $F_X(x) \neq 0$ et $F_Y(y) \neq 0$, donc X et Y ne sont pas indépendantes.

   Valeur moyenne de x : on peut voir que par symétrie par rapport à l'axe des ordonnées, $m_X = 0$. (Ne pas oublier de vérifier que l'intégrale existe.)

$$\begin{aligned} m_X = E_X[X] &= \int_{\mathbb{R}} x f_X(x)dx = \int_{\mathbb{R}} x(\int_{\mathbb{R}} f_{XY}(x, y)dy)dx \\ &= \int\int_{\mathbb{R}^2} x f_{XY}(x, y)dxdy \\ &= \int\int_D \frac{x}{\pi R_{max}^2} dxdy = 0 \end{aligned}$$

Coefficient de corrélation : $\rho_{XY} = \frac{E[(X-m_X)(Y-m_Y)]}{\sigma_X \sigma_Y}$
Le coefficient de corrélation détermine le degré de linéarité entre X et Y. Ici, il n'y a pas de direction privilégiée, donc $\rho_{XY} = 0$.
Par le calcul, $\rho_{XY} = \frac{E[XY]}{\sigma_X \sigma_Y}$ car les moyennes sont nulles.

$$\begin{aligned}
E[XY] &= \int\int_{\mathbb{R}^2} xyf_{XY}(x,y)dxdy \\
&= \int\int_{Q_{++}} xyf_{XY}(x,y)dxdy + \int\int_{Q_{+-}} xyf_{XY}(x,y)dxdy \\
&\quad + \int\int_{Q_{-+}} xyf_{XY}(x,y)dxdy + \int\int_{Q_{--}} xyf_{XY}(x,y)dxdy \\
&= 0
\end{aligned}$$

3. $f_X(x) = \int_{\mathbb{R}} f_{XY}(x,y)dy, \forall x \in \mathbb{R}$

   - $$\forall x/|x| \leq R_{max}, f_X(x) = \int_{-\sqrt{R_{max}^2-x^2}}^{\sqrt{R_{max}^2-x^2}} \frac{1}{\pi R_{max}^2} dy = \frac{2\sqrt{R_{max}^2 - x^2}}{\pi R_{max}^2}$$
     $$f_X(x) = \begin{cases} \frac{2\sqrt{R_{max}^2-x^2}}{\pi R_{max}^2} & \text{si} \quad |x| < R_{max} \\ 0 & \text{sinon} \end{cases}$$
   - $$\begin{aligned} P[x \leq X < x + \Delta x] &= f_X(x)\Delta x \\ &= P[(X,Y) \in D_x] \\ &= \frac{AireD_x}{\pi R_{max}^2} \end{aligned}$$

   On calcule de même $f_Y(y)$ et on remarque que $f_{XY}(x,y) \neq f_X(x)f_Y(y), \forall (x,y) \in D$.
   X et Y ne sont pas indépendantes.

4. On pose $X = R\cos(\Phi)$ et $Y = R\sin(\Phi)$ avec $R \geq 0$ et $Pphi \in]-\pi, \pi[$.
   La répartition est isotrope : $f_{R\Phi}(r,\phi)$ ne dépend pas de ϕ si $|\phi| < \pi$. Ainsi,
   $$f_{R\Phi}(r,\phi) = \begin{cases} 0 & \text{si} \quad r \notin [0, R_{max}] \text{ ou } \phi \notin]-\pi, \pi[\\ g_R(r) & \text{sinon} \end{cases}$$
   La fonction est séparable donc R et Φ sont indépendantes.
   Expliciter $f_\Phi(\phi)$.
   - On voit que Φ suit une loi uniforme sur $]-\pi, \pi[$.
   - $$\begin{aligned} P[\phi \leq \Phi < \phi + \Delta\phi] &= f_\Phi(\phi)\Delta\phi \\ &= P[(R,\Phi)] \in D_\phi = P[(X,Y) \in D_\phi] \\ &= \frac{\frac{\Delta\phi}{2\pi}\pi R_{max}^2}{\pi R_{max}^2} = \frac{\Delta\phi}{2\pi} \end{aligned}$$

Expliciter $f_R(r)$. Attention, la loi n'est pas uniforme ! On applique le même raisonnement que ci-dessus avec un domaine D_r en couronne.

$$\begin{cases} f_R(r) = \frac{2r}{R_{max}^2} & \text{si} \quad r \notin [0, R_{max}] \\ 0 & \text{sinon} \end{cases}$$

5. En utilisant les formules de changement de variables, on exprime $f_{R\Phi}(r, \phi)$

$$f_{R\Phi}(r,\phi) = f_{XY}(x,y)|J| \text{ avec } |J| = \begin{pmatrix} \cos(\phi) & r\sin(\phi) \\ \sin(\phi) & r\cos(\phi) \end{pmatrix}$$

$$f_{R\Phi}(r,\phi) = \begin{cases} f_{XY}(r\cos(\phi), r\sin(\phi))|r| & \text{si} \quad r \in [0, R_{max}[\text{ et } \phi \in]-\pi, \pi] \\ 0 & \text{sinon} \end{cases}$$

$$\text{Or, } f_{XY}(x,y) = \begin{cases} \frac{1}{\pi R_{max}^2} & \text{si} \quad x^2 + y^2 \leq R_{max}^2 \\ 0 & \text{sinon} \end{cases}$$

$$\text{Donc } f_{R\Phi}(r,\phi) = \begin{cases} \frac{r}{\pi R_{max}^2} & \text{si} \quad r \in [0, R_{max}[\text{ et } \phi \in]-\pi, \pi] \\ 0 & \text{sinon} \end{cases}$$

On peut écrire $f_{R\Phi}(r,\phi) = g_R(r)g_\Phi(\phi)$, $f_{R\Phi}(r,\phi)$ est séparable donc R et Φ sont indépendantes.

Si $r \in [0, R_{max}[$,

$$\begin{aligned} f_R(r) &= \int_{\mathbb{R}} f_{R\Phi}(r,\phi)d\phi \\ &= \frac{r}{\pi R_{max}^2} \int_{-\pi}^{\pi} d\phi \\ &= \frac{2r}{R_{max}^2} \end{aligned}$$

Si $\phi \in]-\pi, \pi]$,

$$\begin{aligned} f_\Phi(\phi) &= \int_{\mathbb{R}} f_{R\Phi}(r,\phi)dr \\ &= \int_0^{R_{maxint}} \frac{r}{\pi R_{max}^2} dr \\ &= \frac{1}{2\pi} \end{aligned}$$

6.

$$\begin{aligned} m_R &= E[R] \\ &= \int_{\mathbb{R}} r f_R(r) dr \\ &= \int_0^{R_{max}} \frac{2r^2}{R_{max}} dr \\ &= \frac{2}{3} R_{max} \end{aligned}$$

TD3 Somme de VA et TCL

1. À l'extérieur du carré, $f_{XY}(x,y) = 0$. À l'intérieur, le couple (X,Y) est uniformément réparti donc :

$$f_{XY}(x,y) = \begin{cases} \frac{1}{a^2} & \text{si} \quad x \in [0,a[, y \in [0,a[\\ 0 & \text{sinon} \end{cases}$$

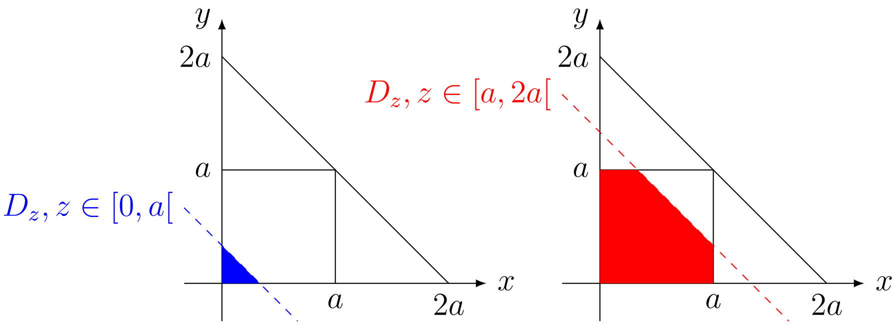
   $f_{XY}(x,y) = g_X(x)g_Y(x)$, f_{XY} est séparable donc X et Y sont indépendantes.

2. On considère la VA $Z = X + Y$. Comme $X \in [0,a[$ et $Y \in [0,a[$, $Z \in [0,2a[$
   Calculons la fonction de répartition de la VA Z.

$$F_Z(z) = \begin{cases} 1 & \text{si} \quad z > 2a \\ ? & \text{si} \quad z \in [0,2a] \\ 0 & \text{si} \quad z < 0 \end{cases}$$

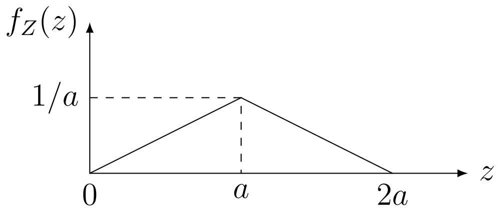
   Si $z \in [0,2a[$, l'expression de la fonction de répartition n'est pas immédiate :

$$\begin{aligned} F_Z(z) &= P[Z < z] \\ &= P[(X,Y) \in D_z] \\ &= \begin{cases} \frac{z^2/2}{a^2} & \text{si} \quad z \in [0,a] \\ \frac{a^2 - \frac{(2a-z)^2}{2}}{a^2} & \text{si} \quad z \in [a,2a] \end{cases} \end{aligned}$$



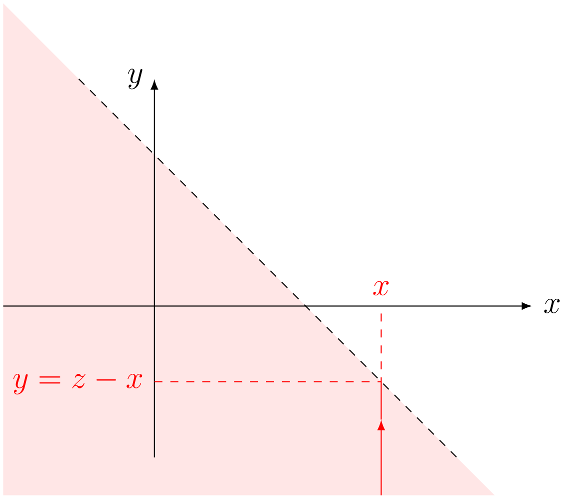
   On peut donc résumer les résultats comme suit :

$$F_Z(z) = \begin{cases} 0 & \text{si} \quad z < 0 \\ \frac{z^2/2}{a^2} & \text{si} \quad z \in [0,a] \\ \frac{a^2 - \frac{(2a-z)^2}{2}}{a^2} & \text{si} \quad z \in [a,2a] \\ 1 & \text{si} \quad z > 2a \end{cases} \qquad f_Z(z) = \begin{cases} 0 & \text{si} \quad z < 0 \text{ ou } z > 2a \\ \frac{z}{a^2} & \text{si} \quad z \in [0,a] \\ \frac{2a-z}{a^2} & \text{si} \quad z \in [a,2a] \end{cases}$$



3. On commence par expliciter $F_Z(z)$ en fonction de $f_{XY}(x,y)$:

$$\begin{aligned}
F_Z(z) &= P[Z < z] = \int_{-\infty}^{z} f_Z(w)dw \\
&= P[(X,Y) \in D_z] \\
&= \int\int_{D_z} f_{XY}(x,y)dxdy \\
F_Z(z) &= \int_{-\infty}^{+\infty} (\int_{-\infty}^{z-x} f_{XY}(x,y)dy)dx
\end{aligned}$$



On en déduit $f_Z(z)$

$$\begin{aligned}
f_Z(z) &= \frac{dF_Z(z)}{dz} = \frac{d}{dz}\int_{-\infty}^{+\infty} (\int_{-\infty}^{z-x} f_{XY}(x,y)dy)dx \\
&= \int_{-\infty}^{+\infty} \frac{\partial}{\partial z}(\int_{-\infty}^{z-x} f_{XY}(x,y)dy)dx \\
f_Z(z) &= \int_{-\infty}^{+\infty} f_{XY}(x, z-x)dx
\end{aligned}$$

4. Les VA X et Y indépendantes donc la ddp $f_{XY}(x,y)$ est séparable :

$$\begin{aligned}
f_Z(z) &= \int_{-\infty}^{+\infty} f_{XY}(x, z-x)dx \\
&= \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx \\
f_Z(z) &= (f_X * f_Y)(z)
\end{aligned}$$

5. Par définition de la fonction caractéristique de la VA Z :

$$\begin{aligned}
\phi_Z(u) &= E[e^{juZ}] \\
&= \int_{\mathbb{R}} f_Z(z)e^{juz}dz
\end{aligned}$$

Ainsi, on peut réécrire

$$\begin{aligned}\phi_Z(u) &= TF[f_Z(z)]_{f=-\frac{u}{2\pi}} \\ &= E[e^{ju(X+Y)}] = E[e^{juX}e^{juY}]\end{aligned}$$

Et par indépendance de X et Y,

$$\begin{aligned}\phi_Z(u) &= \phi_X(u)\phi_Y(u) \\ f_Z(z) &= TF^{-1}[\phi_Z(u)](z) \\ &= (TF^{-1}[\phi_X(u)] * TF^{-1}[\phi_Y(u)])(z) \\ f_Z(z) &= (f_X * f_Y)(z)\end{aligned}$$

6. $X_1, ...X_n$ indépendantes dans leur ensemble

$$\begin{array}{rcl} Y_{12} = & X_1 + X_2 & \rightarrow f_{Y_{12}}(y) = (f_{X_1} * f_{X_2})(y) \\ Y_{123} = & X_1 + X_2 + X_3 = Y_{12} + X_3 & \rightarrow f_{Y_{123}}(y) = (f_{Y_{12}} * f_{X_3})(y) = (f_{X_1} * f_{X_2} * f_{X_3})(y) \end{array}$$

Par récurrence, on montre alors que pour $Y = X_1 + ... + X_n$,

$$f_Y(y) = (f_{X_1} * ... * f_{X_n})(y)$$

On montre que pour $X_n, n = 1, ..., N$ VA réelles et scalaires indépendantes et identiquement distribuées, centrées et d'écart-type σ,

$$Z_N = \frac{\sum_{n=1}^N X_n}{\sqrt{N}} \text{ tend vers une VA gaussienne quand N tend vers } +\infty$$

7. Par linéarité de l'espérance, et comme les variables X_N sont centrées ($E[X_N] = 0$),

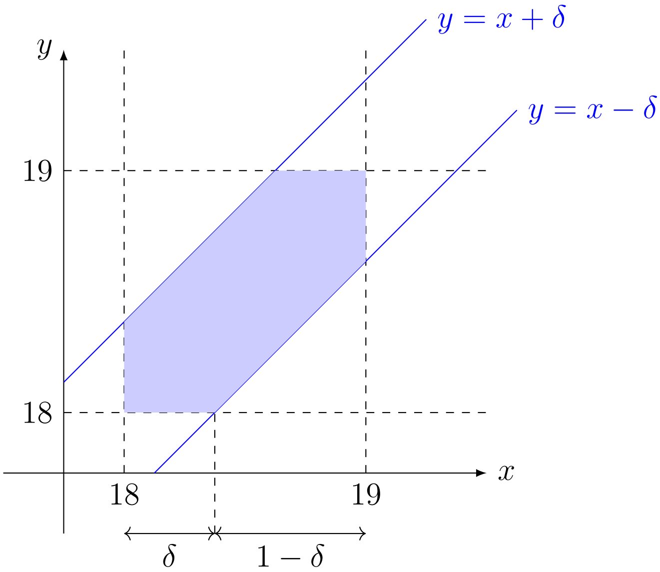
$$E[Z_N] = E[\frac{\sum_{n=1}^N X_n}{\sqrt{N}}] = \frac{\sum_{n=1}^N E[X_n]}{\sqrt{N}} = 0$$

De plus,

$$\begin{aligned}\sigma_Z^2 &= E[(Z_N - m_{Z_N})^2] = E[Z_N^2] \\ &= \frac{1}{N}E[(\sum_{n=1}^N X_n)^2] \\ &= \frac{1}{N}E[\sum_{n=1}^N X_n^2 + \sum_{i\neq j} X_i X_j] \\ &= \frac{1}{N}(\sum_{n=1}^N E[X_n^2] + \sum_{i\neq j} E[X_i X_j]) \\ &= \frac{1}{N}\sum_{n=1}^N \sigma^2 \\ &= \sigma^2\end{aligned}$$

8. Deux personnes se donnent rendez-vous entre 18h et 19h. On associe aux deux instants d'arrivées deux VA X et Y indépendantes, de ddp uniforme sur l'intervalle [18,19].
On introduit la VA $\Delta = |Y - X|$. Calculons sa fonction de répartition.

$$\begin{aligned} F_\Delta(\delta) &= P[\Delta \leq \delta] \\ &= P[|Y - X| \leq \delta] \\ &= P[Y - X \leq \delta \quad \text{et} \quad X - Y \leq \delta] \\ &= P[Y \leq X + \delta \quad \text{et} \quad Y \geq X - \delta] \end{aligned}$$



Ainsi,

$$F_\Delta(\delta) = \begin{cases} 0 & \text{si} \quad \delta < 0 \\ 1 - (1 - \delta)^2 & \text{si} \quad 0 \leq \delta < 1 \\ 1 & \text{si} \quad \delta \geq 1 \end{cases}$$

Donc

$$f_\Delta(\delta) = \begin{cases} 2 - 2\delta & \text{si} \quad 0 \leq \delta < 1 \\ 0 & \text{sinon} \end{cases}$$

TD4 Lois marginale, loi conditionnelles et estimation

On considère une variable aléatoire scalaire et réelle Y de densité de probabilité :

$$f_Y(y) = \frac{1}{X} e^{-\frac{y}{X}} u(x)$$

où $u(y)$ est la fonction échelon d'Heaviside et X un paramètre réel, inconnu, positif et supposé certain dans un premier temps.

1. Calcul de la valeur moyenne et de l'écart type de la VA Y

$$\begin{aligned}
m_Y = E[Y] &= \int_{\mathbb{R}} y f_Y(y) dy \\
&= \int_0^{\infty} y \frac{1}{X} e^{-\frac{y}{x}} dy \\
&= [y(-e^{-\frac{y}{X}})]_0^{\infty} - \int_0^{\infty} (-e^{-\frac{y}{X}}) dy \\
&= \int_0^{\infty} (-e^{-\frac{y}{X}}) dy \\
&= [-X e^{-\frac{y}{X}}]_0^{\infty} \\
m_Y &= X
\end{aligned}$$

De plus,

$$\begin{aligned}
\sigma_Y^2 &= E[(Y - m_Y)^2] \\
&= E[Y^2] - m_Y^2 \\
E[Y^2] &= \int_0^{\infty} y^2 \frac{1}{X} e^{-\frac{y}{X}} dy \\
&= [y^2(-e^{-\frac{y}{X}})]_0^{\infty} + 2X \int_0^{\infty} y \frac{1}{X} e^{-\frac{y}{X}} dy \\
&= 2X m_Y = 2X^2 \\
\sigma_Y^2 &= 2X^2 - X^2 = X^2 \\
\sigma_Y &= X
\end{aligned}$$

$$\boxed{m_Y = \sigma_Y = X}$$

2. On considère N VA $Y_n, n = 1..N$ indépendantes et identiquement distribuées. On note $(y_1, ...y_N)$ les réalisations de $(Y_1, ...Y_N)$.

   Grâce au résultat précédent $m_Y = X$, on peut estimer

$$\hat{x} = \frac{\sum_{n=1}^N y_n}{N}$$

$$\hat{X} = \frac{\sum_{n=1}^N Y_n}{N}$$

   De plus, l'estimateur est non biaisé car

$$E[\hat{X}] = \frac{\sum_{n=1}^N E[Y_n]}{X} = X$$

3. On souhaite exprimer la ddp conjointe des VA $Y_1, ...Y_N$

$$\begin{aligned}
f_{\mathbf{Y}}(\mathbf{y}) &= f_{Y_1,...Y_N}(y_1, ...y_N) \\
&= \prod_{n=1}^{N} f_{Y_n}(y_n) \text{ par indépendance de } Y_1, ...Y_N \\
&= \frac{1}{X^N} \exp\left(-\frac{\sum_{n=1}^N Y_n}{X}\right) u(y_1)...u(y_N)
\end{aligned}$$

4. On utilise l'estimateur du maximum de vraisemblance

$$\hat{x}_{MV} = \arg\max_X f_{\mathbf{Y}}(\mathbf{y})$$

$\hat{x}_{MV}$ est la valeur de x qui rend les valeurs $y_1, ...y_N$ les plus probables.
Condition nécessaire (non suffisante) :

$$\begin{aligned}
\frac{df_{\mathbf{Y}}(\mathbf{y})}{dX}|_{X=\hat{x}_{MV}} = 0 \Leftrightarrow ... \\
&\Leftrightarrow (-\frac{N}{X} + \frac{\sum_{n=1}^{N} y_n}{X^2}) f_{\mathbf{Y}}(\mathbf{y})|_{X=\hat{x}_{MV}} = 0 \\
&\Rightarrow \frac{-N}{\hat{x}_{MV}} + \frac{\sum_{n=1}^{N} y_n}{\hat{x}_{MV}^2} = 0 \\
&\Rightarrow \hat{x}_{MV} = \frac{\sum_{n=1}^{N} y_n}{N} \\
&\Rightarrow \hat{X}_{MV} = \frac{\sum_{n=1}^{N} Y_n}{N}
\end{aligned}$$

Vérifier que c'est un maximum :

$$\begin{aligned}
\frac{d^2 f_{\mathbf{Y}}(\mathbf{y})}{dX^2} &> \text{ ou } < 0? \\
\frac{df_{\mathbf{Y}}(\mathbf{y})}{dX} &> 0 \text{ pour } x \to 0^+
\end{aligned}$$

Calculons la moyenne et l'écart type de $\hat{X}_{MV}$

$$\begin{aligned}
E[\hat{X}_{MV}] &= ... = X \\
\sigma_{MV}^2 &= E[(\hat{X}_{MV} - X)^2] \\
&= E[(\frac{\sum_{n=1}^{N} Y_n}{N} - \frac{NX}{N})^2] \\
&= E[(\frac{\sum_{n=1}^{N}(Y_n - X)}{N})^2] \\
&= \frac{1}{N^2}(\sum_{n=1}^{N} E[(Y_n - X)^2] + \sum_{i \neq j} E[(Y_i - X)(Y_j - X)] \\
&= \frac{NX^2}{N^2} \\
\sigma_{MV} &= \frac{X}{\sqrt{N}}
\end{aligned}$$

5. On a beaucoup plus d'*a priori* (d'informations sur X sans faire l'expérience) avec $\alpha = 1$ qu'avec $\alpha = 10$.
La courbe pour $\alpha = 10$ est beaucoup plus étalée :

$$\sigma_{X,\alpha=1} < \sigma_{X,\alpha=10}$$

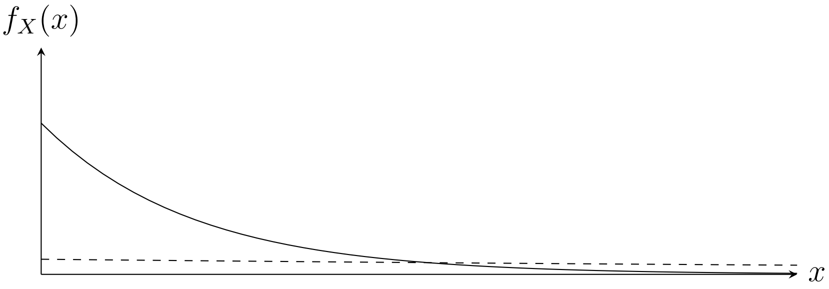


FIGURE 1 – Tracé de $f_X(x)$ pour $\alpha = 1$ (–) et $\alpha = 10$ (- -)

6. $\hat{x}_{MAP} = \arg\max_X f_{X/\mathbf{Y}=\mathbf{y}}(x)$
   Or,

$$\begin{aligned} f_{X/\mathbf{Y}=\mathbf{y}}(x) &= f_{\mathbf{Y},X=x}(\mathbf{y})\frac{f_X(x)}{f_{\mathbf{Y}}(\mathbf{y})} \\ &= \prod_{n=1}^{N} f_{Y_i,X=x(y_i)}\frac{f_X(x)}{f_{\mathbf{Y}}(\mathbf{y})} \\ &= g(x)u(y_1)...u(y_N)\frac{1}{f_{\mathbf{Y}}(\mathbf{y})} \end{aligned}$$

   Ainsi,

$$\hat{x}_{MAP} = \arg\max_X f_{X/\mathbf{Y}=\mathbf{y}}(x) = \arg\max g(x)$$

   Condition nécessaire :

$$\begin{aligned} \frac{dg(x)}{dx}|_{x=\hat{X}_{MAP}} = 0 \Leftrightarrow \hat{X}_{MAP} &= \frac{-N\alpha + \sqrt{\alpha^2N^2 + 4\alpha\sum_{n=1}^{N} y_n}}{2} \\ \Leftrightarrow \hat{X}_{MAP} &= \frac{-N\alpha + \alpha N\sqrt{1 + \frac{4}{\alpha N^2}\sum_{n=1}^{N} y_n}}{2} \end{aligned}$$

7. Si $\alpha \to +\infty$,

$$\begin{aligned} \hat{X}_{MAP} &\approx \frac{-N\alpha + \alpha N(1 + \frac{1}{2}\frac{4}{\alpha N^2}\sum_{n=1}^{N} y_n}{)}2 \\ \hat{X}_{MAP} &\approx \frac{\sum_{n=1}^{N} Y_n}{N} = \hat{X}_{MV} \end{aligned}$$

   On n'a pas d'a priori sur X, on n'a que les observations.

   Si $\alpha \to 0$, m_X et $\sigma_X \to 0$: $\hat{X}_{MAP} \to 0$.
   L'a priori est fort.

TD5 Signaux aléatoire

Signal sinusoïdal à phase équirépartie

On considère le signal aléatoire

$$X_t = x(t) = E_0 \sin(2\pi f_o t + \Phi)$$

ϕ est une variable aléatoire uniformément répartie sur $[0, 2\pi[$
E_0, f_0 sont des grandeurs déterministes strictement positives.

1. À t donné, $f_{X,t}(x,t) = f_{Xt}(x) = f_X(x,t)$
   Méthode : changement de variable

$$\Phi \to X_t = x(t) = g(\Phi) = E_0 \sin(2\pi f_0 t + \Phi)$$

$$f_\Phi(\phi) = \begin{cases} \frac{1}{2\pi} & \text{si} \quad \phi \in [0, 2\pi[\\ 0 & \text{sinon} \end{cases}$$

$X_t = x(t) \in [-E_0, E_0]$ donc $f_{Xt}(x) = 0$, pour tout x tel que $|x| > E_0$.
Théorème de changement de variables aléatoires.
Soit x tel que $|x| < E_0$

$$f_{Xt}(x) = f_X(x,t) = f_\Phi(\phi)|\frac{d\phi}{dx}||_{\phi, g(\phi)=x}$$

Pour tout $x \in [-E_0, E_0]$ (sauf les points où la dérivée s'annule, ensemble de mesure nulle), il y a deux points d'intersection $\phi_i \in [0, 2\pi[$

$$f_X(x,t) = \sum_{i=1}^{2} f_\Phi(\phi_i) \frac{1}{|\frac{dx}{d\phi}|}|_{\phi_i, g(\phi_i)=x}$$

$$\frac{dx}{d\phi} = E_0 \cos(2\pi f_0 t + \phi) \text{ donc } |\frac{dx}{d\phi}|_{\phi=\phi_1} = |\frac{dx}{d\phi}|_{\phi=\phi_2}$$

$$\frac{dx}{d\phi}|_{\phi=\phi_i} = \sqrt{E_0^2 \cos^2(2\pi f_0 t + \phi_i)} = \sqrt{E_0^2(1 - \sin^2(2\pi f_0 t + \phi_i))} = \sqrt{E_0^2 - x^2}$$

Ainsi, on a

$$f_{X_t}(x) = f_X(x,t) = \begin{cases} 0 \quad \text{si} \quad |x| \geq E_0 \\ \frac{1}{\pi\sqrt{E_0^2 - x^2}} \quad \text{sinon} \end{cases}$$

Remarque : $f_X(x,t)$ est finie en $x = \pm E_0$ car Φ VA continue et g fonction continue $\to X_t$ est une VA continue.
Pour conclure quant à la stationnarité à l'ordre un, on regarde si $E[x(t)]$ dépend du temps
Or, $E[x(t)] = \int_{\mathbb{R}} x_t f_X(x_t, t) dx_t$ et $f_X(x_t, t)$ ne dépend pas de t, donc la VA $x(t)$ est stationnaire à l'ordre un.
Autre méthode : fonction de répartition
$F_X(x, t = F_{X_t}(x) = P[X_t < x]$

2.

$$\begin{aligned} E[x(t)] &= \int_{\mathbb{R}} x f_X(x,t)dx = ... = 0 \\ \text{ou } &= E[E_0 \sin(2\pi f_0 t + \phi)] \\ &= \int_{\mathbb{R}} E_0 \sin(2\pi f_0 t + \phi) f_\phi(\phi) d\phi \\ &= E_0 \int_0^{2\pi} \frac{1}{2\pi} \sin(2\pi f_0 t + \phi) d\phi = 0 \end{aligned}$$

La moyenne statistique ne dépend pas de l'origine des temps : stationnarité du moment d'ordre 1.

$$\overline{x(t)} = \frac{1}{T_0} \int_{[T_0]} E_0 \sin(2\pi f_0 t + \phi) dt = 0$$

La moyenne temporelle ne dépend pas de ϕ (de la réalisation) : ergodicité du moment d'ordre 1.

3.

$$\begin{aligned} E[x(t)x(t-\tau)] &= \gamma_{xx}(t, t-\tau) \\ &= E[E_0^2 \sin(2\pi f_0 t + \phi) \sin(2\pi f_0 (t-\tau) + \phi)] \\ &= \frac{E_0^2}{2} E[\cos(2\pi f_0 \tau) - \cos(4\pi f_0 t - 2\pi f_0 \tau + 2\phi)] \\ &= \frac{E_0^2}{2} \cos(2\pi f_0 \tau) \end{aligned}$$

$E[x(t)x(t-\tau)]$ ne dépend pas de l'origine des temps, donc on a $E[x(t)x(t-\tau)] = \gamma_{xx}(\tau)$: stationnarité du moment d'ordre 2

$$\begin{aligned} \overline{x(t)x(t-\tau)} &= \frac{E_0^2}{2} (\overline{\cos(2\pi f_0 \tau)} - \overline{\cos(4\pi f_0 t - 2\pi f_0 \tau + 2\phi)}) \\ &= \frac{E_0^2}{2} \cos(2\pi f_0 \tau) \end{aligned}$$

$\overline{x(t)x(t-\tau)}$ ne dépend pas de ϕ (de la réalisation) : ergodicité du moment d'ordre 2

On a donc ergodicité et stationnarité, à l'ordre 1 et 2.
Remarque : on a donc égalité des moments d'ordre 1 et 2 :

$$\gamma_{xx}(\tau) = \overline{x(t)x(t-\tau)}$$

$\gamma_{xx}(0) = P_x = \frac{E_0^2}{2} < \infty$ et stationnarité du moment d'ordre 1 et 2 $\longrightarrow$ stationnaire au sens large

4. Pour calculer une DSP d'un signal stationnaire au sens large : $\Gamma_{xx}(f) = TF[\gamma_{xx}(\tau)]$

$$\begin{aligned} \Gamma_{xx}(f) &= TF[\gamma_{xx}(\tau)] \\ &= TF[\frac{E_0^2}{2} \cos(2\pi f_0 \tau)] \\ &= \frac{1}{2} \frac{E_0^2}{2} (\delta(f_0 - f) + \delta(f_0 + f)) \end{aligned}$$

Propriétés de la fonction de corrélation

On considère $x(t)$ un SA scalaire et stationnaire.

1.

$$\begin{aligned} m_{x(t)} &= m_x \text{ par stationnarité} \\ &= E[x(t)] \\ \gamma_{xx}(\tau) &= E[x(t)x^*(t-\tau)] \\ \Gamma_{xx}(f) &= TF[\gamma_{xx}(\tau)](f) \end{aligned}$$

2.

$$\begin{aligned} P_x &= E[|x(t)|^2] \\ &= E[x(t)x^*(t)] \\ &= \gamma_{xx}(0) \\ \gamma_{xx}(\tau) &= TF^{-1}[\Gamma_{xx}(f)](\tau) \\ &= \int_{\mathbb{R}} \Gamma_{xx}(f)e^{j2\pi f\tau}df \\ P_x &= \int_{\mathbb{R}} \Gamma_{xx}(f)df \end{aligned}$$

3.

$$\begin{aligned} \gamma_{xx}(-\tau) &= E[x(t)x^*(t+\tau)] \\ &= E[x(t)x^*(t+\tau)]^{**} \\ &= E[x^*(t)x(t+\tau)]^* \\ &= E[x(t+\tau)x^*(t)]^* \\ &= \gamma_{xx}(\tau)^* \end{aligned}$$

Ainsi,

$$\begin{aligned} \Gamma^*_{xx}(f) &= (\int_{\mathbb{R}} \gamma_{xx}(\tau)e^{-j2\pi f\tau}d\tau)^* \\ &= \int_{\mathbb{R}} \gamma^*_{xx}(\tau)e^{j2\pi f\tau}d\tau \\ &= \int_{\mathbb{R}} \gamma_{xx}(-\tau)e^{j2\pi f\tau}d\tau \\ &= \int_{\mathbb{R}} \gamma_{xx}(\tau')e^{-j2\pi f\tau'}d\tau' \\ &= \Gamma_{xx}(f) \end{aligned}$$

Donc $\Gamma_{xx}(f) \in \mathbb{R}$

4. On suppose que $x(t) \in \mathbb{R}$ Montrons que $\Gamma_{xx}(\tau)$ est paire :

$$\begin{aligned}\Gamma_{xx}(-f) &= \int_{\mathbb{R}} \gamma_{xx}(\tau)e^{-j2\pi(-f)\tau}d\tau \\ &= \int_{\mathbb{R}} \gamma_{xx}(-\tau)e^{-j2\pi f\tau}d\tau \text{en posant } \tau = -\tau' \\ &= \int_{\mathbb{R}} \gamma^*_{xx}(\tau)e^{-j2\pi f\tau}d\tau \\ &= \int_{\mathbb{R}} \gamma_{xx}(\tau)e^{-j2\pi f\tau}d\tau \text{ car } x \text{ est réel} \\ &= \Gamma_{xx}(f)\end{aligned}$$

Ainsi, $\Gamma_{xx}(f)$ est bien paire.

5. Montrons que $\gamma_{xy}(-\tau) = \gamma^*_{yx}(-\tau)$

$$\begin{aligned}&\forall \tau \in \mathbb{R} \\ \gamma_{xy}(-\tau) &= E[x(t)y^*(t+\tau)]*^* \\ &= E[y(t+\tau)x^*(t)]^* \\ &= \gamma_{xy}(\tau)^*\end{aligned}$$

Remarque : on retrouve la formule du 3. si y(t) = x(t).

6.
   - $\gamma_{xx}(0)$ est le maximum de l'autocorrélation.
   - donc sa dérivée en 0 est nulle.
   - et sa dérivée seconde est négative pour assurer la concavité (car maximum).

   Remarque : FL(filtre linéaire) = SL(système linéaire) + temps invariant.
   si x(t) est stationnaire alors x'(t) aussi et : $m_{x'} = E[x'(t)] = E[\frac{dx(t)}{dt}] = \frac{d}{dt}E[x(t)]$ car l'espérance ne dépent pas du temps donc l'interversion est possible.
   $\gamma'_{xx}(\tau) = \frac{d}{d\tau}E[x(t)x^*(t-\tau)] = E[x(t)\frac{\partial}{\partial\tau}x^*(t-\tau)] = -\gamma_{xx}(\tau)$

7.

$$\begin{aligned}\gamma_{xx}(\tau) &= \int_{\mathbb{R}} \Gamma_{xx}(f)e^{+j2\pi f\tau}df \\ \gamma'_{xx}(\tau) &= \frac{d}{d\tau}\gamma_{xx}(\tau) = \int_{\mathbb{R}} (j2\pi f)\Gamma_{xx}(\tau)e^{+j2\pi f\tau}df \\ \gamma''_{xx}(\tau) &= \int_{\mathbb{R}} (j2\pi f)^2\Gamma_{xx}(\tau)e^{+j2\pi f\tau}df \\ |\gamma_{xx}(\tau)| &\leq \int_{\mathbb{R}} |\Gamma_{xx}(f)|df = \int_{\mathbb{R}} \Gamma_{xx}(f)df = \gamma_{xx}(0) \\ |\gamma'_{xx}(\tau)| &\leq \int_{\mathbb{R}} 2\pi|f|\Gamma_{xx}(f)df < +\infty \text{ si } \Gamma_{xx}(f) \text{ décroit plus vite que } \frac{1}{f^2} \text{ en } \pm\infty\end{aligned}$$

et ainsi de suite pour des ordres supérieurs

$$\gamma'_{xx}(0) = 0$$

car $f\Gamma_{xx}(f)$ est impaire donc l'intégrale est nulle sur $\mathbb{R}$. Ou alors, γ' est réelle et égal à j fois un réel, donc est nulle.

8. $s(t) = (h * e)(t) = \int_{\mathbb{R}} h(\theta)e(t-\theta)d\theta$ avec h la réponse impulsionnelle.

$$\begin{aligned}
m_s &= E[s(t)] \\
&= E[\int_{\mathbb{R}} h(\theta)e(t-\theta)d\theta] \\
&= \int_{\mathbb{R}} h(\theta)E[e(t-\theta)]d\theta \\
&= m_e \int_{\mathbb{R}} d\theta \\
&= H(0)m_e \\
H(f) &= \int_{\mathbb{R}} h(t)e^{-j2\pi ft}dt \\
H(0) &= \int_{\mathbb{R}} h(t)dt
\end{aligned}$$

9. $\Gamma_{ss}(f)$ en fonction de $H(f)$, $\Gamma_{ee}(f)$. Formules des interférences :

$$\begin{aligned}
\Gamma_{ss}(f) &= H(f)H^*(f)\Gamma_{ee}(f) \\
&= |H(f)|^2\Gamma_{ee}(f)df \\
P_s = \int_{\mathbb{R}} \Gamma_{ss}(f)df &= \int_{\mathbb{R}} |H(f|^2\Gamma_{ee}(f)df \geq 0
\end{aligned}$$

On suppose par l'absurde qu'il existe $f_0 \in \mathbb{R}$ tel que $\Gamma_{ee}(f_0) < 0$
$\Rightarrow \exists (f_1, f_2) \in \mathbb{R}^2 \quad$ avec $\quad f_2 > f_0 > f_1$ tel que

$$\forall f \in]f_1, f_2[, \Gamma_{ee}(f) < 0$$

On utilise un filtre passe-bande idéal de gain unitaire et :

$$P_s = \int_{\mathbb{R}} \Gamma_{ss}(f)df = \int_{f_1}^{f_2} |H(f)|^2\Gamma_{ee}(f)df < 0$$

Impossible donc $\forall f \in \mathbb{R}, \Gamma_{ee}(f) \geq 0$.

10. On considère deux signaux stationnaires dans leur ensemble. La formule des interférences donne :

$$\Gamma_{s_1s_2}(f) = H_1(f)H_2^*(f)\Gamma_{e_1e_2}(f)$$

Montrons que :

$$\gamma_{xx'}(\tau) = -\gamma'_{xx}(\tau)$$

Avec $H_1(f) = 1$ et $H_2(f) = j2\pi f$ (dérivateur de x), on a :

$$\Gamma xx'(f) = -j2\pi f \Gamma xx(f)$$

Par transformée de Fourier inverse il vient :

$$\gamma_{xx'}(\tau) = -\gamma'_{xx}(\tau)$$

De même, avec $H_1(f) = H_2(f) = j2\pi f$, on a :

$$\Gamma x'x'(f) = -(j2\pi f)^2 \Gamma xx(f)$$

avec la transformée inverse de Fourier il vient :

$$\gamma_{x'x'}(\tau) = -\gamma''_{xx}(\tau)$$

TD6 Détection

On envoie un signal x(t) binaire de valeur a de probabilité p et de valeur -a de probabilité 1-p. On reçoit $y(t) = x(t) + b(t)$ avec $x(t)$ la partie utile du signal et $b(t)$ le bruit.

A chaque instant t, $x(t)$ et $b(t)$ sont des VA réelles notées X_t et B_t, et indépendantes. Le bruit B_t suit une loi gaussienne $N(0, \sigma^2)$.

On observe à l'instant t_0, $Y_0 = X_0 + B_0$ que l'on compare au seuil S. La détection, c'est le cas particulier de l'estimation mais avec un nombre discret de valeurs possibles.

1. Avec p $=$ 0.5, les valeurs $+a$ et $-a$ interviennent avec la même probabilité. Le bruit est centré. Le problème est symétrique. Il n'y a pas de raison de privilégier les valeurs strictement positive, ou négative. On pose donc $S = 0$.

2. B est une VA gaussienne centrée et d'écart type σ :

$$f_B(b) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{b^2}{2\sigma^2})$$

3. X est certain et $Y = X + B$ avec $B : N(0, \sigma^2)$, donc on peut intuiter que Y suit la loi gaussienne : $N(X, \sigma^2)$.
   Sinon : Théorème de changement de variable : $f_Y(y) = f_B(b)|\frac{db}{dy}|_{btqy=X+b} = f_B(y - X)$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(y-x)^2}{2\sigma^2})$$

4.

$$\begin{aligned}
F_Y(S) = Pr[Y < S] &= \int_{-\infty}^{S} f_Y(y)dy \\
&= \int_{-\infty}^{S} \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(y-x)^2}{2\sigma^2})dy \\
&= \int_{-\infty}^{\frac{S-X}{\sigma}} \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{1}{2}t^2)dt \\
&= P(\frac{S-X}{\sigma})
\end{aligned}$$

   On en déduit donc

$$\begin{aligned}
F_Y(X + 3\sigma) &= P(3) = 0.9987 \\
F_Y(X - 3\sigma) &= 1 - P(3) = 0.0013
\end{aligned}$$

5. On cherche $\frac{a^2}{\sigma^2}$

$$P_X = E[X^2] = pa^2 + (1-p)(-a)^2 = a^2$$
$$P_B = E[B^2] = \int_{\mathbb{R}} \frac{b^2}{\sqrt{2\pi}\sigma} exp(-\frac{1}{2}\frac{b^2}{\sigma^2})db = \sigma^2$$

Ainsi, $\frac{a^2}{\sigma^2}$ est égal au rapport signal sur bruit (RSB).

6. Comme X et B sont indépendants, si on fixe $X = a$, on se ramène au cas précédent avec X certain, donc

$$f_{Y/X=a}(y) = f_B(y-a)$$

Ainsi,

$$\begin{aligned} Pr[Y < S/X = a] &= \int_{-\infty}^{S} f_{Y/X=a}(y)dy \\ &= P(\frac{S-a}{\sigma} \end{aligned}$$

7. On commet une erreur si
   - on envoie a (probabilité p) et qu'on reconstruit $-a$
   - ou si on envoie $-a$ (probabilité $1-p$ et qu'on reconstruit a

$$\begin{aligned} P_\epsilon &= Pr[(Y < S \quad \text{et} \quad X = a) \quad \text{ou} \quad (Y > S \quad \text{et} \quad X = -a)] \\ &= Pr[(Y < S \quad \text{et} \quad X = a)] + Pr[(Y > S \quad \text{et} \quad X = -a)] \\ &= Pr[(Y < S/X = a)]Pr[X = a] + Pr[(Y > S/X = -a)]Pr[X = -a] \\ P_\epsilon &= P(\frac{S-a}{\sigma})p + (1 - P(\frac{S+a}{\sigma}))(1-p) \end{aligned}$$

Remarque :
Si $S \to +\infty$ (i.e. on reconstruit toujours $-a$), alors on a $P_\epsilon \to p$ (i.e. la probabilité d'erreur correspond à la celle d'envoyer un a).
Si $S \to -\infty$ (i.e. on reconstruit toujours a), alors on a $P_\epsilon \to 1-p$ (i.e. la probabilité d'erreur correspond à la celle d'envoyer un $-a$).

8. Condition nécessaire pour avoir un optimum (attention à vérifier aux bornes) :

$$\frac{dP_\epsilon}{dS}|_{S=S_{opt}} = 0$$

$$\begin{aligned} \frac{dP_\epsilon}{dS}|_{S=S_{opt}} = 0 &\leftrightarrow \frac{p}{\sigma}P(\frac{S-a}{\sigma}) - \frac{1-p}{\sigma}P(\frac{S+a}{\sigma}) = 0 \\ &\leftrightarrow pe^{-\frac{1}{2}(\frac{S-a}{\sigma})^2} - (1-p)e^{-\frac{1}{2}(\frac{S+a}{\sigma})^2} = 0 \\ &\leftrightarrow S_{opt} = \frac{\sigma^2}{2a}\ln(\frac{1-p}{p}) \end{aligned}$$

Pour montrer qu'il s'agit d'un minimum, on peut calculer $\frac{d^2P_\epsilon}{dS^2}|_{S=S_{opt}}$ et vérifier que c'est positif, ou vérifier que $\frac{dP_\epsilon}{dS}$ change de signe en S_{opt}.

9. Lorsque p tend vers 1, S_{opt} tend vers $-\infty$. En effet, si on envoie toujours un a, pour reconstruire uniquement a, il faut toujours être au dessus du seuil.

   Lorsque p tend vers 0, alors S_{opt} tend vers $+\infty$.

10. Lorsque $p = \frac{1}{2}$, $S_{opt} = 0$.

$$\begin{aligned} P_\epsilon &= \frac{1}{2}P(-\frac{a}{\sigma}) + (1 - P(\frac{a}{\sigma}))\frac{1}{2} \\ &= \frac{1}{2}(1 + P(-\frac{a}{\sigma}) - P(\frac{a}{\sigma})) \\ P_\epsilon &= P(-\frac{a}{\sigma}) \end{aligned}$$

   Lorsque a/σ "grand" (bon RSB), alors $P_\epsilon = P(-\frac{a}{\sigma}) = 1-P(\frac{a}{\sigma}) \to 0$. Si le bruit est faible, l'erreur aussi.

   Lorsque a/σ "petit", alors $P_\epsilon \to \frac{1}{2}$. Si le bruit est élevée, on a autant de chance d'avoir la bonne valeur que de se tromper.

   $a/\sigma = 3 : P_\epsilon = 1 - P(3) = 0.0013$

11. Pour diminuer la probabilité d'erreur, on peut par exemple réaliser deux mesures au lieu d'une sur chaque intervalle de temps : cela permet de "moyenner" l'effet du bruit. En effet, pour deux VA Y_1 et Y_2 décrites par $N(0, \sigma^2)$, on a pour la VA $Y = \frac{Y_1+Y_2}{2}$ un écart type de $\sigma' = \frac{\sigma}{\sqrt{2}}$.

TD7 Prédiction

TD7.1 Introduction

- Grandeur à estimer : VA $\theta = x(t+\Delta t)$, $\Delta t > 0$.
- Information a priori : $x(t)$ SA réel, scalaire, centré ($\forall t \in \mathbb{R}, E[x(t)] = 0$), stationnaire ($E[x(t)] = m_x(t) = m_x$).
- Observations / mesures : dans la partie II, $Y = x(t)$ et dans la partie III, $\mathbf{Y} = \begin{bmatrix} x(t) \\ x'(t) \end{bmatrix}$
- Choix de l'estimateur : estimateur linéaire : $\hat{\theta} = \mathbf{HY}$
  - Partie II : $\hat{\theta} = \hat{x}(t+\Delta t) = Hx(t)$
  - Partie III : $\hat{\theta} = \hat{x}(t+\Delta t) = \mathbf{H} \begin{bmatrix} x(t) \\ x'(t) \end{bmatrix} = [a \quad b] \begin{bmatrix} x(t) \\ x'(t) \end{bmatrix} = ax(t) + bx'(t)$
- Calcul des caractéristiques statistiques de l'estimateur
  - $\tilde{\theta} = \hat{x}(t+\Delta t) - x(t+\Delta t)$
  - $E[\tilde{\theta}] =$ biais(moyen)
  - $E[\tilde{\theta}^2] = P_{\tilde{\theta}} = E[(\hat{\theta} - \theta)^2] =$ erreur quadratique moyenne (puissance de l'erreur)

Objectif : Minimiser $P_{\tilde{\theta}}$

- Variations lentes : $x(t+\Delta t) \approx x(t)$.
  $\hat{x}(t+\Delta t) = x(t) = 1.x(t)$. L'erreur d'estimation vient de celle de $x(t+\Delta t) \approx x(t)$.
  $\hat{x}(t+\Delta t) = 1.x't) + \Delta t x'(t)$
- Fortement corrélé : $\gamma_{xx}(\Delta t) \approx \gamma_{xx}(0)$.
  On obtient les mêmes expressions que précédemment pour $\hat{x}(t+\Delta t)$.
- Faiblement corrélé : la fonction d'autocorrélation est "plus étroite" : $\gamma_{xx}(\Delta t) \approx 0$
  $\hat{x}(t+\Delta t) = 0$ i.e. $a = 0, b = 0$
- Signal sinusoïdal
  $\hat{x}(t+\Delta t) = \frac{x_1(t+\Delta t)+x_2(t+\Delta t)}{2}$ si on a seulement accès à $x(t)$
  $\hat{x}(t+\Delta t) = x(t+\Delta t)$ si on a accès à $x(t)$ et $x'(t)$.

TD7.2 Estimateur à partir de $x(t)$

On utilise l'estimateur suivant :

$$\hat{x}(t+\Delta t) = a.x(t)$$

1. Calculons l'erreur moyenne :

$$\begin{aligned} E[\tilde{x}(t+\Delta t)] &= E[\hat{x}(t+\Delta t) - x(t+\Delta t)] \\ &= E[a.x(t) - x(t+\Delta t)] \\ &= aE[x(t)] - E[x(t+\Delta t)] \\ &= 0 \text{ car le signal est centré} \end{aligned}$$

L'estimateur est non biaisé car la moyenne de l'estimateur est égale à la moyenne du signal.

2. Calculons l'erreur quadratique :

$$\begin{aligned} P_{\tilde{\theta}} &= E[\tilde{x}(t+\Delta t)^2] \\ &= E[(ax(t) - x(t+\Delta t))^2] \\ P_{\tilde{\theta}}(a) &= a^2\gamma_{xx}(0) - 2a\gamma_{xx}(\Delta t) + \gamma_{xx}(0) \end{aligned}$$

   $P_{\tilde{\theta}}(a)$ est une parabole et $\gamma_{xx}(a) > 0$ donc on a la CNS de maximum :

$$\frac{dP_{\tilde{\theta}}(a)}{da}|_{a_{opt}} = 0 \Leftrightarrow a_{opt} = \frac{\gamma_{xx}(\Delta t)}{\gamma_{xx}(0)} \in [-1, 1]$$

   On en déduit l'erreur quadratique minimale :

$$P_{min} = P_{\tilde{\theta}}(a_{opt}) = \gamma_{xx}(0) - \frac{\gamma_{xx}^2(\Delta t)}{\gamma_{xx}(0)} = \gamma_{xx}(0)(1 - (\frac{\gamma_{xx}(\Delta t)}{\gamma_{xx}(0)})^2)$$

   Calculons l'erreur moyenne de $\tilde{x}(t+\Delta t)x(t)$:

$$\begin{aligned} E[\tilde{x}(t+\Delta t)x(t)] &= E[(ax(t) - x(t+\Delta t))x(t)] \\ &= a\gamma_{xx}(0) - \gamma_{xx}(\Delta t) \end{aligned}$$

   Pour $a = a_{opt}$,

$$E[\tilde{x}(t+\Delta t)x(t)] = 0 \text{ (principe d'orthogonalité)}$$

   On peut réécrire ce résultat :

$$E[\tilde{x}(t+\Delta t)x(t)] = \gamma_{x\tilde{x}}(\Delta t) = 0$$

   Autrement dit, il ne reste plus d'information commune entre $\tilde{x}(t+\Delta)$ et $x(t)$. On a extrait ce qu'on pouvait. Si on ne l'avait pas fait ($\gamma_{x\tilde{x}}(\Delta t) \neq 0$), on pourrait trouver un meilleur estimateur.

3. Dans le cas du bruit blanc $\gamma_{xx}(\Delta t) = 0$ donc $a_{opt} = 0$.
   Dans le cas du faiblement corrélé, $\gamma_{xx}(\Delta t) = 0$.
   Fortement corrélé : $\gamma_{xx}(\Delta t) \approx \gamma_{xx}(0)$ donc $a_{opt} \approx 1$

TD7.3 Estimateur à partir de $x(t)$ et $x'(t)$

On considère l'estimateur :

$$\hat{x}(t+\Delta t) = ax(t) + bx'(t)$$

Hypothèses :

- $\tau \rightarrow \gamma_{xx}(\tau)$ est dérivable 2 fois

- $\gamma'_{xx}(0) = 0$
- $\gamma''_{xx}(0) < 0$

1. Biais de l'estimateur ?

$$\begin{aligned} E[\tilde{x}(t+\Delta t)] &= E[(\hat{x}(t+\Delta t) - x(t-\Delta t)] \\ &= E[ax(t) + bx'(t) - x(t+\Delta t)] \\ &= aE[x(t)] + bE[x'(t)] - E[x(t+\Delta t)] \\ &= 0 \end{aligned}$$

L'estimateur est non biaisé, et e $\forall a, b \in \mathbb{R}^2$.

2. Erreur quadratique moyenne de l'estimateur ?

$$\begin{aligned} P_{\tilde{\theta}} &= E[\tilde{x}(t+\Delta)^2] \\ &= E[(ax(t) + bx'(t) - x(t+\Delta t))^2] \\ &= a^2E[x(t)^2] + b^2E[x'(t)^2] + E[x(t+\Delta t)^2] + 2abE[x(t)x'(t)] \\ &\quad - 2aE[x(t)x(t+\Delta t) - 2bE[x'(t)x(t+\Delta t)] \end{aligned}$$

D'après les résultats démontrés au TD précédent (via formule des interférences) :

$$\begin{aligned} P_{\tilde{\theta}} &= a^2\gamma_{xx}(0) - b^2\gamma''_{xx}(0) + \gamma_{xx}(0) - 2ab\gamma'_{xx}(0) - 2a\gamma_{xx}(\Delta t) + 2b\gamma'_{xx}(\Delta t) \\ &= (1+a^2)\gamma_{xx}(0) - b^2\gamma''_{xx}(0) - 2a\gamma_{xx}(\Delta t) + 2b\gamma'_{xx}(\Delta t) \end{aligned}$$

Ceci définit sans conteste un fantastique paraboloïde tourné ver le haut ! En effet, les coefficients de vant a^2 et b^2 ont le bon goût d'être positifs (car $gxx(0) = P_x > 0$ et $\gamma''_{xx}(0) < 0$ (puissance maximum en 0)).
Tout ça pour ne pas minimiser la belle fonction à deux variables, car on a maintenant une CNS de minimum de l'erreur quadratique :

$$\frac{\partial P_{\tilde{\theta}}}{\partial a}|_{a=a_{opt}} = 0 \qquad \text{et} \qquad \frac{\partial P_{\tilde{\theta}}}{\partial b}|_{b=b_{opt}}$$

On en déduit donc :

$$a_{opt} = \frac{\gamma_{xx}(\Delta t)}{\gamma_{xx}(0)} \qquad \text{et} \qquad b_{opt} = \frac{\gamma'_{xx}(\Delta t)}{\gamma''_{xx}(0)}$$

puis

$$P_{\tilde{\theta}}(a_{opt}, b_{opt}) = \gamma_{xx}(0) - \frac{\gamma^2_{xx}(\Delta t)}{\gamma_{xx}(0)} + \frac{\gamma'^2_{xx}(\Delta t)}{\gamma''_{xx}(0)}$$

On compare les 2 estimateurs :

$$P_{min,2} = \gamma_{xx}(0)[1 - (\frac{\gamma_{xx}(\Delta t)}{\gamma_{xx}(0)})^2] + \frac{\gamma'^2_{xx}(\Delta)}{\gamma''_{xx}(0)} = P_{min,1} + \frac{\gamma'^2_{xx}(\Delta)}{\gamma''_{xx}(0)}$$

Or, $\frac{\gamma'^2_{xx}(\Delta)}{\gamma''_{xx}(0)} < 0$ donc $P_{min,2} < P_{min,1}$

$$\begin{aligned}
E[\tilde{x}(t+\Delta t)x(t)] &= E[(ax(t)+bx'(t)-x(t+\Delta t))x(t) \\
&= a\gamma_{xx}(0) - b\gamma'_{xx}(0) - \gamma_{xx}(\Delta t) \\
&= a\gamma_{xx}(0) - \gamma_{xx}(\Delta t) \\
&= 0 \quad \text{avec} \quad a = a_{opt}
\end{aligned}$$

$$\begin{aligned}
E[\tilde{x}(t+\Delta t)x(t)] &= E[(ax(t)+bx'(t)-x(t+\Delta t))x'(t)] \\
&= a\gamma'_{xx}(0) - b\gamma''_{xx}(0) + \gamma'_{xx}(\Delta t) \\
&= -b\gamma''_{xx}(0) + \gamma'_{xx}(\Delta t) \\
&= 0 \quad \text{avec} \quad b = b_{opt}
\end{aligned}$$

On aurait pu utiliser ce résultat (principe d'orthogonalité) pour trouver les valeurs de a_{opt} et b_{opt}.

Résumé : dans le cadre d'un estimateur linéaire :

$$\hat{\theta} = \hat{x}(t+\Delta t) = \mathbf{H}\begin{bmatrix} x(t) \\ x'(t) \end{bmatrix} = [a \quad b]\begin{bmatrix} x(t) \\ x'(t) \end{bmatrix} = ax(t) + bx'(t)$$

- 1ère méthode : exprimer $P_{\tilde{\theta}} = E[\tilde{\theta}^2]$, chercher le $\mathbf{H}_{opt} = [a_{opt} \quad b_{opt}]$ tel que $P_{\tilde{\theta}}$ est minimale. On en déduit $\hat{\theta} = \mathbf{H}_{opt}\mathbf{Y}$.
- 2ème méthode : Principe d'orthogonalité, revient à chercher $\mathbf{H}$ tel que $E[\tilde{\theta}\mathbf{Y}^T] = 0$.

$$E[\tilde{\theta}\mathbf{Y}^T] = 0 \Leftrightarrow \text{Chercher } P_{\tilde{\theta}} min + \text{ estimateur lin.}$$

Remarque : Innovation = l'erreur $\tilde{x}(t+\Delta t)$ dans le cas où l'estimateur minimise $P_{\tilde{\theta}}$

TD7.4 Comparaison

1. - Les deux estimateurs sont non biaisés.
   - $P_{min,2} \leq P_{min,1}$: le 2ème est cool !

2. On suppose Δt "petit". Au début du TD, on avait alors intuité que

$$\hat{x}(t+\Delta t) = 1.x't) + \Delta t x'(t)$$

soit $a_{opt} = 1$ et $b_{opt} = \Delta t$.

$$a_{opt} = \frac{\gamma_{xx}(\Delta t)}{\gamma_{xx}(0)} \approx \frac{\gamma_{xx}(0) + \Delta t\gamma'_{xx}(0)}{\gamma_{xx}(0)} = 1$$

$$b_{opt} = \frac{\gamma'_{xx}(\Delta t)}{\gamma''_{xx}(0)} \approx \frac{\gamma'_{xx}(0) + \Delta t\gamma''_{xx}(0)}{\gamma''_{xx}(0)} = \Delta t$$

WIRKLICH WUNDERBAR !

TD7.5 Application

On s'intéresse maintenant au signal :

$$x(t) = E_0 \sin(2\pi f_0 t + \phi)$$

où Φ est une VA uniformément répartie sur $[0, 2\pi[$

On a montré dans un TD précédent que ce signal est stationnaire et ergodique (à l'ordre 2).

1. $E[x(t)] = 0$ car Φ est une VA uniforme. Par ergodicité et stationnarité au 1er ordre, $\overline{x(t)} = E[x(t)] = 0$.

   On calcule la fonction d'autocorrélation et comme on l'a déjà vu :

   $$\gamma_{xx}(\tau) = \frac{E_0^2}{2} \cos(2\pi f_0 t)$$

2. $\gamma_{xx}(\tau) = \frac{E_0^2}{2} \cos(2\pi f_0 t)$ a sensiblement l'air périodique, d'amplitude $\frac{E_0^2}{2}$ et de fréquence f_0.

3. 1er estimateur : $\hat{x}_1(t + \Delta t) = a_{opt} x(t)$.

   Or, $a_{opt} = \frac{\gamma_{xx}(\Delta t)}{\gamma_{xx}(0)} = \cos(2\pi f_0 \Delta t)$, donc

   $$\hat{x}_1(t + \Delta t) = E_0 \cos(2\pi f_0 \Delta t) x(t) = E_0 \cos(2\pi f_0 \Delta t) \sin(2\pi f_0 t + \phi)$$

4. 2ème estimateur : $\hat{x}_2(t + \Delta t) = a_{opt} x(t) + b_{opt} x'(t)$.

   Or, $a_{opt} = \frac{\gamma_{xx}(\Delta t)}{\gamma_{xx}(0)} = \cos(2\pi f_0 \Delta t)$ et $b_{opt} = \frac{\gamma'_{xx}(\Delta t)}{\gamma''_{xx}(0)} = \frac{\sin(2\pi f_0 \Delta t)}{2\pi f_0}$, donc

   $$\hat{x}_2(t + \Delta t) = E_0 \cos(2\pi f_0 \Delta t) \sin(2\pi f_0 t + \phi) + E_0 \frac{\sin(2\pi f_0 \Delta t)}{2\pi f_0} (2\pi f_0 \cos(2\pi f_0 t + \phi))$$

   $$\hat{x}_2(t + \Delta t) = E_0 \sin(2\pi f_0 (t + \Delta t) + \phi) = x(t + \Delta t)$$

TD8 Estimation de la vitesse d'un véhicule

On cherche à estimer V (paramètre constant).
On relève la position du véhicule le long du rail à des instants $t_n = nT$.
On considère $Y_n = nTV + B_n$, avec B_n $N(0, \sigma_B^2)$.
À $t_n = t_0 = 0$, le mobile se trouve en 0.

TD9 Expériences

1. On trace la droite qui passe au mieux par tous les points et l'origine, on trouve une pente d'environ 1 m/s.

   L'hypothèse "bruit blanc" a l'air de marcher mais ne pas faire de conclusion rapide (on n'a que 10 mesures).

2. La "meilleure droite" ne passe pas par l'origine. Les causes possibles sont : un bruit non centré, ou une position non nulle à $t_0 = 0$. On a l'impression que le bruit est corrélé, mais on ne peut pas tirer de conclusion.

3. Pour obtenir la ddp de $\hat{V}$:
   - Méthode basée sur l'expérience : chaque jeu d'observation donne $\hat{v}_i$ et on trace l'histogramme (voir TP d'initiation à Matlab).
   - Méthode de changement de variable : ddp de B_n, puis ddp de Y_n et enfin (passage difficile) ddp de $\hat{V}$.

   Remarque : $\hat{V}\ N(m_{\hat{v}}, \sigma^2_{\hat{V}})$
   - 1er estimateur (non biaisé) : $m_{\hat{v}} = 1m/s$ et $\sigma_{\hat{V}} = 0,08m/s$
   - 2ème estimateur (non biaisé) : $m_{\hat{v}} = 1m/s$ et $\sigma_{\hat{V}} = 0,04m/s$, meilleur estimateur car meilleur écart-type.

TD9.1 Estimateur empirique

Dans cette partie, $Y_n = nTV + B_n$ avec bruit faible, donc $V = \frac{y_n}{nT}$.

1. On mesure y_n.

$$\hat{V}_{emp} = \frac{Y_n}{nT}$$

$$E[\hat{V}_{emp}] = \frac{E[Y_n]}{nT} = \frac{E[nTV + B_n]}{nT} = V + \frac{E[B_n]}{nT} = V$$

   donc l'estimateur est non biaisé.

$$\sigma^2_{\hat{V}_{emp}} = E[(\hat{V}_{emp} - m_{\hat{V}_{emp}})^2] = E[(\frac{Y_n}{nT} - V)^2] = E[(\frac{nTV + B_n}{nT} - V)^2] = \frac{1}{(nT)^2}E[B_n^2]$$

$$\sigma_{\hat{V}_{emp}} = \frac{\sigma_B}{nT}$$

   Ainsi, pour minimiser $\sigma_{\hat{V}_{emp}}$, on prend $n = N$ (la plus grande mesure).

2. On dispose de N mesures $y_1, ...y_N$

$$\hat{V}_{emp} = \frac{\sum_n \frac{Y_n}{nT}}{n}$$

L'estimateur n'est pas biaisé car :

$$E[\hat{V}_{emp}] = \frac{\sum_n \frac{E[Y_n]}{nT}}{N} = V$$

Écart-type de l'estimateur :

$$\begin{aligned}
\sigma^2_{\hat{V}_{emp}} &= E[(\hat{V}_{emp} - m_{\hat{V}_{emp}})^2] \\
&= E[(\hat{V}_{emp} - V)^2] \\
&= E[(\frac{\sum_n V + \frac{B_n}{nT}}{N} - \frac{NV}{N})^2] \\
&= \frac{1}{N^2} E[(\sum_n \frac{B_n}{nT})^2] \\
&= \frac{1}{(NT)^2} E[(\sum_n \frac{B_n}{n})^2] \\
&= \frac{1}{(NT)^2} (\sum_n \frac{E[B_n^2]}{n^2} + \sum_{n \neq m} \frac{E[B_n B_m]}{nm})
\end{aligned}$$

Le bruit est blanc, donc les $E[B_n B_m] = 0$

$$\begin{aligned}
\sigma^2_{\hat{V}_{emp}} &= \frac{\sigma_B^2}{(NT)^2} \sum_n \frac{1}{n^2} \\
\sigma_{\hat{V}_{emp}} &= \frac{\sigma_B}{nT} \sqrt{S_1}
\end{aligned}$$

Or, $\frac{\sigma_B}{nT}\sqrt{S_1} > \frac{\sigma_B}{nT}$. Cela signifie que notre estimateur avec 1 mesure est "meilleur" que celui avec n mesures. On est triste d'avoir considéré les premières mesures qui sont très sensibles, comme nous, mais au bruit.

TD9.2 Préliminaires

V grandeur certaine mais inconnue, $Y_n = nTV + B_n$ et $B_n = N(0, \sigma_B^2)$.

1. Moyenne de Y_n : $m_n = E[Y_n] = nTV$.
   Écart-type de Y_n : $\sigma_n^2 = E[(Y_n - m_n)^2] = E[B_n^2]$ donc $\sigma_n = \sigma_B$
   Coefficient de corrélation :

$$\rho_{mn} = \frac{E[(Y_n - m_n)(Y_m - m_m)]}{\sigma_n \sigma_m} = \frac{E[B_n B_m]}{\sigma_B^2} = \begin{cases} 1 & \text{si} \quad m = n \\ 0 & \text{si} \quad m \neq n \end{cases} = \delta_{n-m}$$

2. Par changement de variables aléatoires si ça t'amuse,

$$f_{Y_n}(y_n) = f_{B_n}(y_n - nTV) = frac1\sqrt{2\pi\sigma_b^2} e^{-frac(y_n - nTV)^2 2\sigma_b^2}$$

   Le caractère gaussien se conserve par transformation linéaire, i.e. toute combinaison linéaire de VA suivant une loi gaussienne suit aussi une loi gaussienne. Attention, ne pas sommer les ddp.

Les Y_i étant indépendants (car les B_i sont indépendants car blancs) :

$$f_{Y_1,\ldots Y_n}(y_1,\ldots y_n) = \prod_{i=1}^{N} f_{Y_i}(y_i) = \frac{1}{(2\pi)^{N/2}\sigma_B^N}\exp(-\frac{1}{2}\frac{\sum_n (y_n - nTV)^2}{\sigma_B^2})$$

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{N/2}\sigma_B^N}\exp(-\frac{1}{2}\frac{(\mathbf{y}-E[\mathbf{y}])^T(\mathbf{y}-E[\mathbf{y}])}{\sigma_B^2})$$

Rappel : la décorrélation n'implique pas l'indépendance, il faut le caractère "gaussiens dans leur ensemble".

TD9.3 Estimateur des moindres carrés

1. L'estimateur au sens des moindres carrés de V est :

$$\hat{v}_{MC} = \arg_V \min \sum_{n=1}^{N} (y_n - nTV)^2$$

En posant $J_{MC} : V \to \sum_{n=1}^{N}(y_n - nTV)^2$, J_{MC} est une parabole (concavité tournée vers le haut). On a alors une condition nécessaire et suffisante à la minimisation :

$$\begin{aligned}
\frac{dJ_{MC}}{dV}|_{V=\hat{V}_{MC}} = 0 &\Leftrightarrow \sum_{n=1}^{N} 2(-nT)(y_n - nTV) = 0\\
&\Leftrightarrow \sum_{n=1}^{N} ny_n - \hat{v}_{MC}T\sum_{n=1}^{N} n^2 = 0\\
&\Leftrightarrow \hat{v}_{MC} = \frac{\sum_n ny_n}{T\sum_{n=1}^{N} n^2}\\
&\Leftrightarrow \hat{V}_{MC} = \frac{\sum_n nY_n}{TS_2}
\end{aligned}$$

2. Calculons la moyenne de l'estimateur.

$$m_{MC} = E[\hat{v}_{MC}] = \frac{\sum_n nE[Y_n]}{TS_2} = \frac{\sum_n n(nTV)}{TS_2} = V$$

L'estimateur non biaisé.

On s'intéresse à son écart-type.

$$\begin{aligned}
\sigma_{MC}^2 &= E[(\hat{V}_{MC} - m_{MC})^2]\\
&= E[(\hat{V}_{MC} - V)^2]\\
&= E[(\frac{\sum_n nY_n}{TS_2} - V)^2]\\
&= E[(\frac{\sum_n n^2(TV + nB_n) - TV\sum_n n^2}{TS_2})^2]\\
&= E[(\frac{\sum_n nB_n}{TS_2})^2]
\end{aligned}$$

Comme les B_n sont décorrélés, les doubles produits sont tous nuls

$$\sigma_{MC}^2 = \frac{\sum_n n^2 E[B_n^2]}{T^2 S_2^2}$$
$$\sigma_{MC} = \frac{\sigma_B}{T\sqrt{S_2}}$$

Comme $S_2 = \sum_n n^2 > N^2$, on a $\sqrt{S_2} > N$ donc $\sigma_{MC} < \sigma_{emp} = \frac{\sigma_B}{NT}$.
Notre estimateur est meilleur que l'estimateur empirique.

TD9.4 Estimateur du maximum de vraisemblance

1. L'estimateur du maximum de vraisemblance de V est donné par

$$\hat{v}_{MV} = \arg_V \max f_{\mathbf{Y}}(\mathbf{y})$$

Remarque : dans le cas où V est incertain, $\arg_v \max f_{\mathbf{Y}/V=v}(\mathbf{y})$.

Y suit une loi gaussienne donc :

$$\hat{v}_{MV} = \arg_V \max f_{\mathbf{Y}}(\mathbf{y}) = \arg_V \min \sum_n (y_n - nTV)^2 = \arg\min J_{MC}(V)$$

2. Identique à la partie précédente car $\hat{V}_{MV} = \hat{V}_{MC}$.

TD9.5 Estimateur du maximum a posteriori

- V suit une loi gaussienne $N(V_0, \sigma_V^2)$
- les VA B_n et V sont indépendantes 2 à 2

On a 2 types d'informations :

- celle qui vient des mesures y_n
- celle qui vient de l'a priori V

1. Dans le cas où $V = v$ (v est certain), on ne change pas pour autant le comportement de $B_1, ...B_n$, donc de $Y_1...Y_n$:

$$f_{\mathbf{Y}/V=v}(\mathbf{y}) = \frac{1}{(2\pi)^{N/2}\sigma_B^N} \exp(-\frac{1}{2}\frac{\sum_n (y_n - nTv)^2}{\sigma_B^2})$$

2. On utilise la règle de Bayes :

$$f_{V/\mathbf{Y}=\mathbf{y}}(v) = \frac{f_{\mathbf{Y}/V=v}(\mathbf{y}) f_V(v)}{f_{\mathbf{Y}}(\mathbf{y})}$$

$f_{\mathbf{Y}}(\mathbf{y})$ ne dépend pas de v car on peut la calculer selon $f_{\mathbf{Y}}(\mathbf{y}) = \int_{\mathbb{R}} f_{\mathbf{Y},V}(\mathbf{y}, v) dv$.

3. L'estimateur du maximum a posteriori de V est donné par :

$$\hat{v}_{MAP} = \arg_v \max f_{V/\mathbf{Y}=\mathbf{y}}(v)$$

Or

$$f_{V/\mathbf{Y}=\mathbf{y}}(v) = cste \times \exp(-\frac{1}{2}(\frac{\sum_n (y_n - nvT)^2}{\sigma_B^2} + \frac{(v - V_0)^2}{\sigma_V^2}))$$

On pose $J_{MAP} = \frac{\sum_n (y_n - nvT)^2}{\sigma_B^2} + \frac{(v-V_0)^2}{\sigma_V^2}$ et on a alors $\hat{v}_{MAP} = \arg_v \min J_{MAP}(v)$.

CNS de maximisation :

$$\frac{dJ_{MAp}}{dV}|_{V=\hat{v}_{MAP}} = 0 \Leftrightarrow -2T\frac{\sum_n (y_n - nvT)n}{\sigma_B^2} + 2\frac{(v-V_0)}{\sigma_V^2} = 0$$

$$\Leftrightarrow \hat{v}_{MAP} = \frac{\frac{T\sum_n ny_n}{\sigma_B^2} + \frac{V_0}{\sigma_V^2}}{\frac{T^2\sum_n n^2}{\sigma_B^2} + \frac{1}{\sigma_V^2}}$$

$$\Leftrightarrow \hat{v}_{MAP} = \frac{\frac{\hat{v}_{MV}}{\sigma_{MV}^2} + \frac{V_0}{\sigma_V^2}}{\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2}}$$

Donc

$$\hat{V}_{MAP} = \frac{\frac{\hat{V}_{MV}}{\sigma_{MV}^2} + \frac{V_0}{\sigma_V^2}}{\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2}}$$

C'est un barycentre entre les mesures représentées par $\hat{v}_{MV}$ et l'a priori V_0.

- Si σ_V "petit", alors $\hat{V}_{MAP} \approx V_0$: beaucoup d'a priori donc on n'a pas exploité les mesures.
- Si σ_V "grand", alors $\hat{V}_{MAP} \approx \hat{V}_{MV}$: l'a priori est tellement pourri qu'on n'en tient pas compte.

4. On forme l'erreur d'estimation

$$\tilde{V}_{MAP} = \hat{V}_{MAP} - V$$

- Biais ?

$$E[\tilde{V}_{MAP}] = E[\hat{V}_{MAP}] - E[V] = 0$$

- Variance de l'erreur d'estimation : puissance de l'erreur dans le cas non biaisé.

$$\begin{aligned}
\sigma_{MAP}^2 &= E[(\tilde{V}_{MAP} - E[\hat{V}_{MAP}])^2] \\
&= E[(\hat{V}_{MAP} - V)^2] \\
&= E[(\frac{\frac{\hat{V}_{MV}-V}{\sigma_{MV}^2} + \frac{V_0-V}{\sigma_V^2}}{\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2}})^2] \\
&= \frac{1}{(\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^2} E[(\frac{\hat{V}_{MV}-V}{\sigma_{MV}^2} + \frac{V_0-V}{\sigma_V^2})^2] \\
&= \frac{1}{(\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^2} (\frac{E[(\hat{V}_{MV}-V)^2]}{\sigma_{MV}^4} + \frac{E[(V_0-V)^2}{\sigma_V^4}) \\
&= \frac{1}{(\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^2} (\frac{\sigma_{MV}^2}{\sigma_{MV}^4} + \frac{\sigma_V^2}{\sigma_V^4}) \\
\sigma_{MAP}^2 &= (\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^{-1}
\end{aligned}$$

On a donc $\sigma_{emp} > \sigma_{MV} > \sigma_{MAP}$.