Notions d'estimation

Supposons que l'on veuille connaître le poids 'p' ou la taille 'T' que devrait avoir un enfant de 4 ans. Pour ce faire, on prendra un échantillon d'enfants de 4 que l'on pèsera et dont on mesurera la taille. Une moyenne sur ces mesures nous fournira une estimation $\hat{p}$ du poids une estimation $\hat{T}$ de la taille. Cet échantillon devra bien sûr être le plus grand possible et le plus hétérogène possible (plusieurs ethnies). L'estimation de la durée moyenne d'une communication téléphonique est très utile aux opérateurs de téléphonie pour ajuster les prix et offres en conséquence ou pour la gestion du trafic. Autre exemple : les intentions de vote. Dans chacun des cas, un échantillon est considéré pour représenter la population.

1. Estimateur

Dans de nombreuses situations, on ne dispose pas directement de mesures sur la variable d'intérêt mais que d'une observation liée à notre variable inconnue. Le but des techniques d'estimation est d'utiliser les observations pour extraire de l'information sur la grandeur d'intérêt. Ceci doit fait de la meilleure façon qui soit et implique plusieurs choix dont la relation mathématique entre l'observation et la variable à estimer.

Dans les problèmes d'estimation, on a, en général, affaire à deux catégories de variables:

$\underline{X}$: Un vecteur de variables inconnues que l'on cherche à déterminer

$\underline{Y}$: Un vecteur de variables mesurées (observation ou échantillon) liées à x.

Déterminer **X** à partir de **Y** (et donc en fonction de **Y**) consiste à traiter de manière déterministe les mesures **Y** pour obtenir une grandeur $\hat{X}$, proche de **X** , soit : $\hat{X} = fct(Y)$.

On peut distinguer les deux cas suivants :

-X et Y sont aléatoire (estimateurs de Bayes).

- Si X est déterministe et Y aléatoire (estimateurs de Fisher)

Un estimateur est donc une valeur $\hat{X}$ calculée sur un échantillon Y tiré au hasard (dit aussi observations), ce qui fait de $\hat{X}$ une variable aléatoire possédant une moyenne et une variance.

__Exemple__

Considérons que l'on veuille estimer la valeur d'une résistance R (voir Exo 1 du TP n°5). On procède comme suit : On réalise N expériences indépendantes où l'on mesure le courant I(k) passant par la résistance et la tension U(k) à ses bornes. Ces mesures de courant et de tension sont toutes bruitées par un bruit dont on ignore la densité de probabilité.

Observations : Y = (I(1), U(1), I(2), U(2), . . ., I(N), U(N))

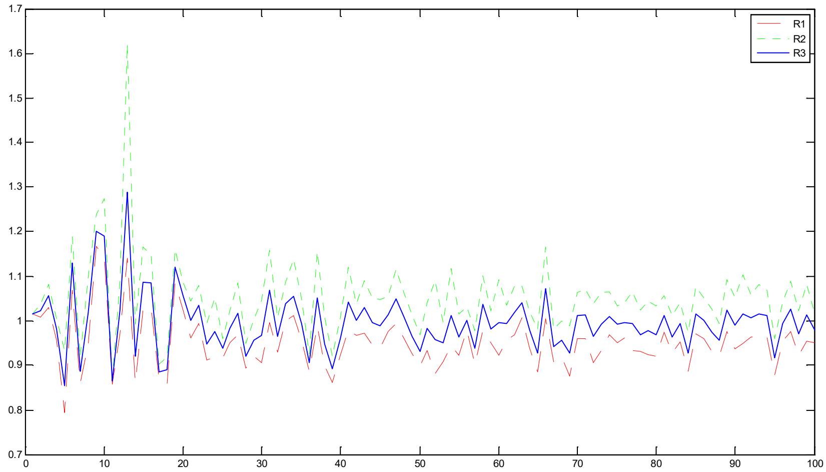
Variable à estimer : X= R

Bien que X et Y ne soient pas supposés avoir une densité de probabilité, on peut construire différents estimateur de R en fonctions des observations U(k) et I(k)

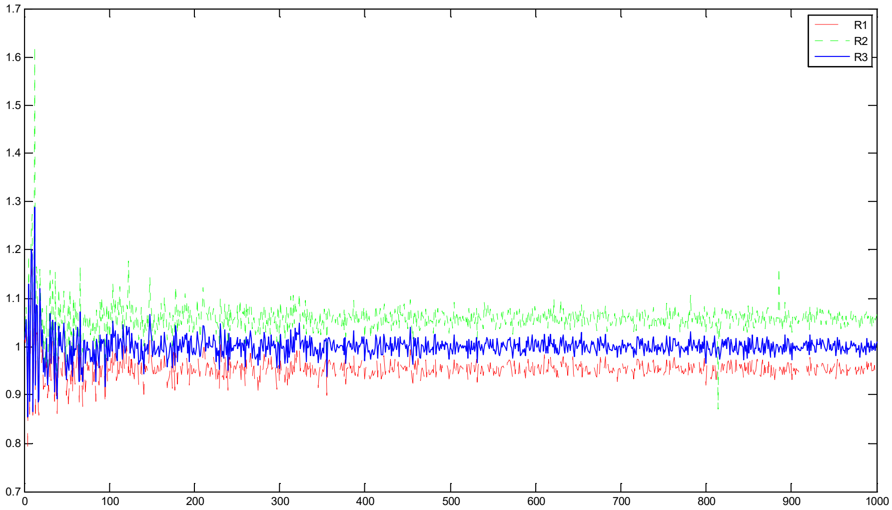
Estimateur 1 : $\hat{R}_1 = \dfrac{\sum_{k=1}^{N} U(k)I(k)}{\sum_{k=1}^{N} I^2(k)}$

Estimateur 2 : $\hat{R}_2 = \dfrac{1}{N}\sum_{k=1}^{N}\dfrac{U(k)}{I(k)}$

Estimateur 3 : $\hat{R}_3 = \dfrac{\sum_{k=1}^{N} U(k)}{\sum_{k=1}^{N} I(k)}$



Ces figures montrent l'évolution de l'estimée d'une résistance R=1 en fonction du nombre de mesures pour U=I=1. Ci-contre une estimation pour N=100 et ci-bas pour N=1000



On remarque que le premier estimateur donne des valeurs comprises entre 0.9 et 1. Le deuxième fournit une estimation proche de 1 et 1.2 Ils sont donc tous deux fortement biaisés. Le troisième estimateur est quant à lui non biaisé.

Ainsi, pour pouvoir déterminer d'une façon constructive une règle d'estimation, il faut définir un critère qui évalue la qualité des résultats, et définir l'estimée comme l'application de Y en X qui optimise ce critère (minimiser une erreur ou une distance, dite aussi fonction de coût). Deux mesures de qualité de l'estimation sont largement utilisées, il s'agit du biais et de la variance.

2. Propriétés des estimateurs

Le biais est la moyenne de l'écart et la variance est la puissance de l'écart (mesure les fluctuations de l'estimateur autour de la valeur souhaité).

L'estimateur $\hat{x}$ a pour biais $b = E\{\hat{x} - x\}$ et pour variance: $\sigma^2 = E\{(\hat{x} - E\{\hat{x}\})^2\} = E\{\hat{x}^2\} - E\{\hat{x}\}^2$

- Si x est déterministe, alors le biais est $b = E\{\hat{x}\}$ -x
- Si x est aléatoire, alors le biais vaut $b = E\{\hat{x}\}$ -μ_x

❑ Le biais indique la valeur moyenne de l'erreur d'estimation, trois cas sont possibles:

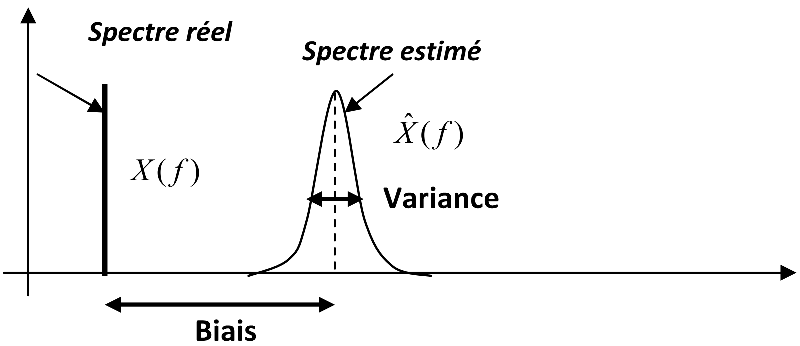
• $E[\hat{x}\}$ =x ou μ_x pour toutes les valeurs possibles du paramètre. On dit alors que l'estimée est non-biaisée ;

• $E[\hat{x}\}$ =x+b ou μ_x+b. Si b est indépendant de x, dans ce cas l'estimateur a un biais constant et connu, qui peut toujours être éliminé ;

• $E[\hat{x}\}$ =x+ b(x) , c'est-à-dire, on a un biais qui dépend de x(qui est inconnu).

❑ La variance doit être aussi petite que possible, de façon à que l'estimée soit concentrée autour de la vraie valeur du paramètre.

❑ Plus b et σ^2 sont faibles, meilleur est l'estimateur (estimateur non biaisée à variance minimale). Malheureusement, la diminution de l'un provoque l'augmentation de l'autre. Dans ces conditions, un estimateur biaisé pourrait être préférable, si le biais est faible, à un estimateur non biaisé mais de très grande variance. C'est le compromis biais-variance , on peut démontrer que $b^2=E\{\hat{x}-x\}^2 - \sigma^2$ où $E\{\hat{x}-x\}^2$ est l'erreur quadratique moyenne.



❑ On dit qu'un estimateur est *consistant* s'il tend vers la vraie valeur du paramètre quand le nombre d'observations tend vers infini (comportement asymptotique) :

$$\lim_{N\to+\infty} b_N = \lim_{N\to+\infty} \sigma_N = 0$$

❑ Si l'estimateur est non biaisé, nous avons une borne *inférieure de la variance* (borne de Cramer Rao). Si la variance atteint la borne de CR, on dit que l'estimateur est *efficace.* Autrement dit, un estimateur non-biaisé est appelé efficace si sa variance dans est plus petite que celle de n'importe quel autre estimateur non-biaisé.

<u>Remarque:</u> Dans ce qui suit, on parlera de variables x_i iid (indépendantes et identiquement distribuées). L'exemple classique est celui du lancé de dé : Les variables aléatoires représentent chaque résultat des lancers (0 pour face et 1 pour pile) suivent toutes la même loi de Bernoulli. Bien que les lancers soient successifs, le résultats d'un lancé donné ne dépend pas des résultats précédents et n'aura aucune influence sur les prochains lancés (pas de lien de dépendance)

Ce sont des variables aléatoires qui suivent toutes la même loi de probabilité (donc même moyenne et même variance) et sont indépendantes.

On en déduit les propriétés suivantes:

- $E\{x_i\}=m \quad \forall i$

- $E\{(x_i-m)^2\}=\sigma^2$

- $E\{(x_i-m)(x_k-m)\}=0$ pour $k \neq i$ (indépendants et centrés)

Exemple d'application

Supposons que l'on observe N échantillons indépendants $Y_1, Y_2, \ldots\ldots, Y_n$ d'une variables aléatoire Y de moyenne $m= E\{Y_i\}$ et de variance $\sigma^2=E\{Y_i^2\}-m^2$.

1. On désire estimer X=m (déterministe) en supposant σ^2 connue. Pour cela, on étudie les deux estimateurs suivants :

$$X_1=\frac{1}{n}\sum_{i=1}^{n}Y_i \qquad\qquad X_2=\frac{2}{n(n+1)}\sum_{i=1}^{n}i.Y_i$$

- Biais de $X_1= E\{\hat{X}_1\}-X=0$ - Variance de $X_1=\sigma^2/n$
- Biais de $X_2= E\{\hat{X}_2\}-X=0$ - Variance de $X_2=2\sigma^2(2n+1)/3n(n+1)$

2. On désire estimer X= σ^2 en supposant m connue. Pour cela, on étudie les deux estimateurs suivants :

$$X_1=\frac{1}{n}\sum_{i=1}^{n}(Y_i-m)^2 \text{ et } X_1=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\mu_Y)^2$$

- Biais de $X_1= E\{\hat{X}_1\}-X=0$ - Biais de $X_2= E\{\hat{X}_2\}-X=(n-1).\sigma^2/n$

3. Estimateurs des moyennes statistiques

Les quantités $E[x\}$,$E[x\}^2$, ,$Sxx(f)$,... sont impossibles à calculer sur un ordinateur car elles nécessiteraient un nombre de points infini. Sur calculateur, on ne dispose que d'une séquence discrète et finie de *N* points. En réalité, on calcule des estimées de ces grandeurs en supposant généralement le signal stationnaire et ergodique. On remplace alors le calcul des moyennes statistiques par des moyennes temporelles.

Estimation de Moyenne

Soit N échantillons (x_0, x_1, …..,x_{N-1}) indépendants et identiquement distribuées (même loi avec même paramètre) d'un signal aléatoire stationnaire. On définit l'estimée $\hat{m}$ de E{x}

$$\hat{m}=\frac{1}{N}\sum_{k=0}^{N-1}x_k$$

On remarque donc qu'on remplace la moyenne statistique par la moyenne temporelle sur la séquence finie. Cet estimateur est non biaisé. En effet, du fait que le signal soit iid , $\forall k; E\,[x_k\} = m$, alors :

$$E\{\hat{m}\}=\frac{1}{N}\sum_{k=0}^{N-1}E\{x_k\}=\frac{1}{N}.N.m=m \quad\Rightarrow\quad b\{\hat{m}\}=0$$

Pour la variance d'estimation, on obtient :

$$\sigma_{\hat{m}}^2 = \mathrm{var}\{\hat{m}\} = E\left\{(\hat{m} - E\{\hat{m}\})^2\right\} = E\left\{\hat{m}^2\right\} - E\{\hat{m}\}^2$$
$$= E\left\{\left(\frac{1}{N}\sum_{k=0}^{N-1} x_k - m\right)^2\right\} = E\left\{\left(\frac{1}{N}\sum_{k=0}^{N-1} x_k - \frac{1}{N}\sum_{k=0}^{N-1} m\right)^2\right\} = \frac{1}{N^2} E\left\{\left(\sum_{k=0}^{N-1}(x_k - m)\right)^2\right\}$$

Sachant que les x_k sont indépendants et que $E\{x_k-m\}=0$, alors : $\sigma_{\hat{m}}^2 = \frac{1}{N^2}\sum_{i=0}^{N-1}(x_k - m)^2 = \sigma^2 / N$

Cet estimateur est consistant, puisque pour $N\rightarrow\infty$, biais et variance tendent vers 0.

Estimateurs de variance

Selon que l'on connaisse ou non la moyenne, on utilisera l'un des deux estimateurs suivants :

$$\hat{\sigma}_1{}^2 = \frac{1}{N}\sum_{k=0}^{N-1}(x_k - m)^2 \qquad\qquad \hat{\sigma}_2{}^2 = \frac{1}{N-1}\sum_{k=0}^{N-1}(x_k - \hat{m})^2$$

Calculons le biais des deux estimateurs

- Pour le premier, on a : $E\left\{\hat{\sigma}_1{}^2\right\} = \frac{1}{N}\sum_{k=0}^{N-1} E\left\{(x_k - m)^2\right\} = \sigma^2 \qquad \Rightarrow \qquad b_{\hat{\sigma}_1^2} = 0$

- Pour le deuxième, on a : $E\left\{\hat{\sigma}_2{}^2\right\} = \frac{1}{(N-1)}\sum_{k=0}^{N-1} E\left\{(x_k - \hat{m})^2\right\} = \frac{1}{(N-1)}\sum_{k=0}^{N-1} E\left\{(x_k - m + m - \hat{m})^2\right\}$

$$E\left\{\hat{\sigma}_2{}^2\right\} = \frac{1}{(N-1)}\sum_{k=0}^{N-1} E\left\{\left(x_k - m + m - \frac{1}{N}\sum_{k=0}^{N-1} x_j\right)^2\right\} = \frac{1}{(N-1)}\sum_{k=0}^{N-1} E\left\{\left((x_k - m) - \frac{1}{N}\sum_{k=0}^{N-1}(x_j - m)\right)^2\right\}$$

Sachant que $E\{(x_k - m)\} = 0$, on a :

$$E\left\{\hat{\sigma}_2{}^2\right\} = \frac{1}{(N-1)}\sum_{k=0}^{N-1}\left(E\left\{(x_k - m)^2\right\} + \left(\frac{1}{N}\sum_{k=0}^{N-1}(x_j - m)\right)^2 - \frac{2}{N}E\left\{(x_k - m)^2\right\}\right)$$
$$= \frac{1}{(N-1)}\sum_{i=0}^{N-1}\left(\sigma^2 + \frac{1}{N^2}\sum_{k=0}^{N-1} E\left\{(x_k - m)^2\right\} - \frac{2\sigma^2}{N}\right) = \frac{1}{(N-1)}\sum_{i=0}^{N-1}\left(.\sigma^2 + \frac{\sigma^2}{N} - \frac{2\sigma^2}{N}\right) = \frac{N\sigma^2}{(N-1)}\left(\frac{N-1}{N}\right) = \sigma^2$$

Les deux estimateurs ne sont pas biaisés puisque on obtient la vraie valeur à chaque fois.

Estimateurs de la corrélation

Dans le cas discret, la fonction d'autocorrélation d'un signal aléatoire x; supposé ergodique, est définie par : $R_{xx}(k) = E\{x(n)x^*(n-k)\} = E\{x(n)x^*(n-k)\} = \lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{n=+N} x(n)x^*(n-k)$

A partir d'un échantillon (x_0, x_1, …..,x_{N-1}) indépendants et identiquement distribuées, plusieurs estimateurs sont alors possible:

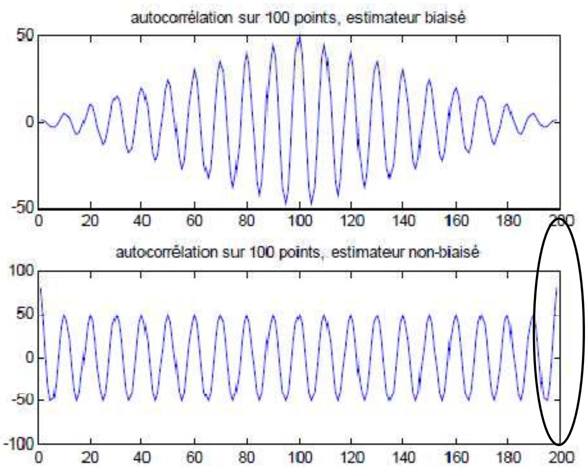
$$\hat{R}_{xx}(k) = \frac{1}{N}\sum_{n=k}^{N-1} x(n)x^*(n-k) \qquad \hat{R}_{xx}(k) = \frac{1}{N-k}\sum_{n=k}^{N-1} x(n)x^*(n-k)$$

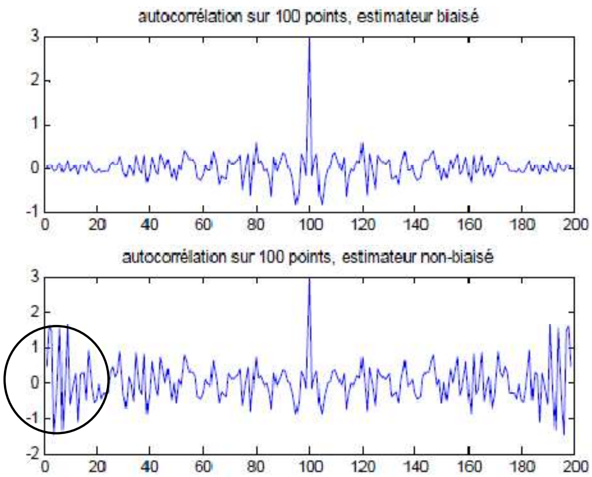
Le premier estimateur est biaisé car on ne tient pas compte du nombre d'échantillons disponibles qui varie avec le pas k. Le deuxième en tient compte, il est non biaisé. En effet :

$$E\{\hat{R}_{xx}(k)\} = \frac{1}{N-k}\sum_{n=k}^{N-1} E\{x(n)x^*(n-k)\} = \frac{1}{N-k}\sum_{n=k}^{N-1} R_{xx}(k) = R_{xx}(k)$$

On peut démontrer que la variance de cet estimateur tend vers 0 quand N tend vers l'infini. Cet estimateur est donc consistant.

On observe par contre que lorsque k approche du nombre d'échantillons N, la variance de l'estimateur non biaisé devient excessive (voir exemple pour un bruit blanc et une sinusoïde). Alors que celle de l'estimateur biaisé varie beaucoup moins. C'est une des raisons pour lesquelles cet estimateur est souvent utilisé par la suite, malgré son biais.





<u>Estimateurs de la densité spectrale</u>

a. Méthode du périodogramme

Supposons que l'on dispose d'une séquence de N points du signal aléatoire : x [x_0,x_1,...,x_{N-1}}.

Soit donc la séquence y_k obtenue à partir de la séquence x_k pondérée par la fenêtre f_k : $y_k = x_k . f_k$

La transformée de Fourier calculée sur la séquence finie sera donc convoluée avec le spectre de la fenêtre rectangulaire c'est-à-dire un sinus cardinal. Rappelons que les propriétés spectrales du sinus cardinal ne sont pas bien adaptées à l'analyse spectrale du signal (lobes secondaires importants). Pour y remédier, on emploie des fenêtres de pondérations (Hamming, Hanning, Kaiser,etc.)

La méthode du périodogramme à prendre la TF d'une réalisation :

$$\hat{S}_{xx}(f) = \frac{1}{N}|Y(f)|^2 \qquad avec \qquad Y(f) = \sum_{k=0}^{N-1} y_k e^{-2\pi jfk}$$

D'où
$$\hat{S}_{xx}(f)=\frac{1}{N}\sum_{l=0}^{N-1}y_l e^{-2\pi jfl}\sum_{m=0}^{N-1}y_m e^{2\pi jfm}\;\frac{1}{N}\sum_{l=0}^{N-1}\sum_{m=0}^{N-1}y_l y_m\; e^{-2\pi jf(l-m)}$$
$$=\frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=k}^{N-1}y_l y_{l-k}\; e^{-2\pi jfk}=\sum_{k=0}^{N-1}\hat{R}_{yy}(k)e^{-2\pi jfk}=TF\left(\hat{R}_{yy}(k)\right)$$

Le calcul de l'espérance de $\hat{S}_{xx}(f)$ donne :

$$E\left\{\hat{S}_{xx}(f)\right\}=E\left\{\sum_{k=0}^{N-1}\hat{R}_{yy}(k)e^{-2\pi jfk}\right\}=E\left\{\frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=k}^{N-1}y_l y_{l-k}\; e^{-2\pi jfk}\right\}$$

$$E\left\{\hat{S}_{xx}(f)\right\}==\frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=k}^{N-1}E\left\{y_l y_{l-k}\right\}e^{-2\pi jfk}=\sum_{k=0}^{N-1}\frac{N-k}{N}R_{yy}(k)\,e^{-2\pi jfk}=S_{yy}(f)*N\sin c(fN)^2$$

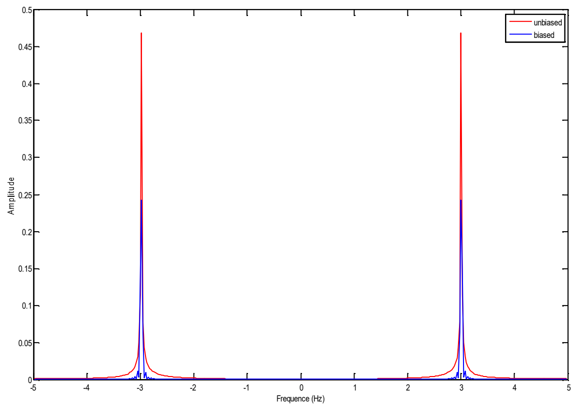
Le périodogramme est un estimateur biaisé. Le périodogramme est donc en moyenne la convolution du véritable spectre avec la transformée de Fourier de la fenêtre triangulaire. Néanmoins, lorsque N $\rightarrow\infty$, le biais devient nul. La variance est pratiquement indépendante de N et proportionnelle au spectre :≃ $Sx(f)^2$

Afin de diminuer la variance de cet estimateur, on peut utiliser un périodogramme moyenné. Cela consiste à séparer le signal en K tranches (de longueur N/K), à calculer le périodogramme sur chaque tranche et à faire la moyenne.

Du fait des K moyennages, la variance est presque divisée par K : néanmoins, les tranches étant plus courtes, la résolution diminue. En pratique, un taux de recouvrement de 40% donne de bons résultats. Au delà, l'indépendance entre les tranches n'est plus respectée.

b. La méthode du corrélogramme consiste à calculer d'abord l'estimée de *Rxx*(*k*) de la fonction d'autocorrélation puis à prendre pour estimée de la densité spectrale la TF de cette estimée.

$$\hat{S}_{xx}(f)=TF\left[\hat{R}_{xx}(k)\right]$$



Les corrélogrammes présentent chacun un extrema pour la fréquence normalisée 3 qui est bien la fréquence de la sinusoïde étudiée. Il ressort également de ces courbes que le corrélogramme non-biaisé permet de mieux faire ressortir la fréquence de la sinusoïde mais au dépend d'une plus grande variance sur les bords.

c) Dans cette méthode, on recherche un modèle AR, MA ou ARMA pour la séquence x_k. Dans le cas d'un modèle auto-régressif $x_k = a_1x_{k-1}+....+a_rx_{k-r}+u_k$

$$\hat{S}_{xx}(f)=\left|\frac{1}{1+\sum_{k=1}^{r}a_k e^{-i2\pi fk}}\right|^2 u_0$$

TD n°5 : Notions d'estimation

1. Les éléments d'une population possèdent un caractère X qui suit une loi de probabilité dont la densité est donnée par :

$$p_\theta(x) = \begin{cases} \frac{kx^3}{\theta^4} & si\ \ 0 < x < \theta \\ 0 & sinon \end{cases}$$

Où k =cste et θ>0 est le paramètre inconnu.

- Déterminer la constante k.
- Montrer que l'estimateur $\frac{1}{N}\sum_{i=1}^{N} X_i$ de θ est biaisé et en déduire un non biaisé.

2. Soit $(X_1, X_2, \ldots\ldots X_N)$ avec N>10 un échantillon aléatoire issu d'une population suivant une loi de Bernoulli de paramètre p. p(x)=p.δ(x-1)+(1-p).δ(x) . Considérons les trois estimateurs du paramètre p :

$$\hat{p}_1 = \left(\sum_{i=1}^{N} X_i - 1\right)/N \qquad \hat{p}_2 = \frac{1}{N/2}\sum_{i=1}^{N/2} X_{2i} \qquad \hat{p}_3 = \frac{1}{10}\sum_{i=1}^{10} X_i$$

- Comparer les trois estimateurs (consistance)

3. Soit $(X_1, X_2, \ldots\ldots X_N)$ une population de moyenne μ et de variance σ^2. On considère les deux estimateurs de la moyenne :

$$\mu_1 = \frac{1}{N}\sum_{i=1}^{N} X_i \qquad \mu_2 = \sum_{i=1}^{N} a_i X_i / \sum_{i=1}^{N} a_i \qquad a_i \geq 1$$

- Montrer que ces deux estimateurs sont non biaisés.
- Lequel sera le plus favorable à un étudiant ?

4. Soient $(X_1, X_2, \ldots\ldots X_N)$, N variables aléatoires indépendantes suivant des lois de poisson de paramètre λ. Soient les estimateurs suivants du paramètre λ suivant :

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^{N} x_i}{N} \qquad \hat{\lambda}_2 = \frac{x_1 + x_N}{2}$$

- Que signifient le biais et la variance pour un estimateur ?
- Calculer biais et variance des 2 estimateurs.
- Lequel est le plus efficace et pourquoi ?

<u>Solutions</u>

1. k=4 $\qquad$ b=(4θ/5)-θ $\qquad \theta = \frac{5}{4N}\sum_{i=1}^{N} x_i$

2. p(x)=pδ(x-1)+(1-p)δ(x) $\qquad \mu_x = p \qquad \sigma_x^2 = p(1-p)$

- p_1 b=1/N σ^2= p(1-p)/N (=0 N→∞) consistant
- p_2 b=0 σ^2=2.p(1-p)/N (=0 N→∞) consistant
- p_3 b=0 σ^2=p(1-p)/10 (≠0 N→∞) non consistant

p_1 meilleur estimateur

3. μ_1 voir cours $\qquad \mu_2$: b=0 ,σ^2= $\sum_{i=1}^{n} a_i^2 \sigma^2 / \left(\sum_{i=1}^{n} a_i\right)^2$

4. $b_{\lambda 1}=0$ $\quad b_{\lambda 2}=0$ $\quad \sigma_{\lambda 1}^2 = \sigma^2/N$ $\quad \sigma_{\lambda 2}^2 = \sigma^2/2$ Le premier (variance plus petite pour N>2)

Exercices supplémentaires

1. Soit X une variable aléatoire qui suit une loi Binomiale B(16; p), où p est un paramètre inconnu. On se propose d'estimer p a l'aide d'un échantillon de taille n de X : (x1; .....;xn). On se propose d'étudier les estimateurs suivants :

$$T_1 = \frac{X_1 + ... + X_n}{n} \;;\; T_2 = \frac{X_1 + ... + X_n}{16n} \;;\; T_3 = \frac{1}{16}\left[X_1 + \frac{X_2 + ... + X_n}{n-1}\right]$$

- Dire lesquels sont sans biais, convergents.
- Lequel doit-on employer de préférence?

2. Soit D la variable aléatoire représentant la durée d'une communication téléphonique. On suppose que la loi de D est la loi Uniforme U sur [0,θ], pour θ > 0. On veut estimer θ, pour cela, on observe un n échantillon : soit les durées D_i de N communications. Etudier l'estimateur suivant : $\theta = \frac{2}{N}\sum_{i=1}^{N} D_i$

- Calculer son biais, sa variance, est-il consistant?

3. Soient $s_1^2, s_2^2, \ldots, s_k^2$ les k variances-échantillons obtenues à partir de k échantillons aléatoires simples indépendants de tailles respectives $n_1, n_2, \ldots, n_k$.

1. $U = \frac{n_1 s_1^2 + n_2 s_2^2 + \ldots + n_k s_k^2}{n_1 + n_2 + \ldots + n_k}$ est-il un estimateur non biaisé de σ^2 ?
2. Si ce n'est pas le cas, comment le biais peut-il être enlevé ?

4. Soient les variables aléatoires $X_i (i = 1; 2)$ i.i.d. de moyenne μ et de variance σ^2. Lequel des deux estimateurs de μ non biaisés suivants conseilleriez-vous :

$$M_1 = \frac{X_1 + X_2}{2} \text{ ou } M_2 = \frac{aX_1 + bX_2}{a + b},$$

où $a, b \in \mathbb{R}$?

TP n°5 : Notions d'estimation

But : Aborder les notions d'estimation ainsi que les propriété des estimateurs (biais, variance) à travers un exemple puis étudier les estimateurs de la moyenne, de la variance, de la corrélation et de la DSP.

Exercice 1: On reprend les 3 estimateurs de la valeur d'une résistance R donnés comme suit :

$$\hat{R}_1 = \frac{\sum_{k=1}^{N} U(k)I(k)}{\sum_{k=1}^{N} I^2(k)} \qquad \hat{R}_2 = \frac{1}{N}\sum_{k=1}^{N}\frac{U(k)}{I(k)} \qquad \hat{R}_3 = \frac{\sum_{k=1}^{N} U(k)}{\sum_{k=1}^{N} I(k)}$$

Le programme suivant permet de simuler des courants et tensions bruitées et d'étudier les 3 estimateurs.

```
clc;clear all; close all;
randn('state', sum(100*clock));N=1000; it=1:N;R=1;
for k=2:N
    U=1+0.22*randn(1,k); I=1+0.22*randn(1,k);
R1(k)=sum(U.*I)/sum(I.^2);     R2(k)=sum(U./I)/k;     R3(k)=sum(U)/sum(I);
end
figure; plot(it(1:100),R1(1:100),'-.r',it(1:100),R2(1:100),':g',it(1:100),R3(1:100),'-b');
legend('R1','R2','R3');
figure; plot(it,R1,'-.r',it,R2,':g',it,R3,'-b');legend('R1','R2','R3');
for k=2:N
biais_R1(k)=mean(R1(1:k))-R;biais_R2(k)=mean(R2(1:k))-R;biais_R3(k)=mean(R3(1:k))-R;
end
figure ;subplot(311); plot(biais_R1); grid; legend('Biais de R1');
subplot(312); plot(biais_R2); grid; legend('Biais de R2');
subplot(313); plot(biais_R3);grid; legend('Biais de R3');
for k=2:N
var_R1(k)=var(R1(1:k));var_R2(k)=var(R2(1:k));var_R3(k)=var(R3(1:k));
end
figure ;subplot(311); plot(var_R1); legend('Var de R1');
subplot(312); plot(var_R2); legend('Var de R2');subplot(313); plot(var_R3);legend('Var de
R3');
```

1. Expliquer toutes les boucles.
2. Etudier le biais et la variance de chaque estimateur. Conclure.
3. Lesquels sont consistants? Lequel est le plus efficace?

Exercice 2: Utiliser le programme ci-dessus pour construire les estimateurs de la moyenne et de la variance

$$\hat{m} = \frac{1}{N}\sum_{k=0}^{N-1} x_k \qquad \hat{\sigma}_1^{\,2} = \frac{1}{N}\sum_{k=0}^{N-1}(x_k - m)^2 \qquad \hat{\sigma}_2^{\,2} = \frac{1}{N-1}\sum_{k=0}^{N-1}(x_k - \hat{m})^2$$

1. Prendre m=1 et σ^2=2. Etudier le biais, la variance et la consistance de l'estimateur de la moyenne
2. Quelle est la différence entre les deux estimateurs de la variance ?
3. Comparer les 2 estimateurs de la variance sur les 100ères valeurs. Quel est le biais de chacun d'eux ?
4. Que remarque-t-on lorsque N augmente ? Les 3 estimateurs sont-ils efficaces?

Exercice 3 :Les deux estimateurs classiques de la fonction d'autocorrélation sont donnés par :

$$\hat{R}_{kk}(k) = \hat{R}_{kk}(-k) = \frac{1}{N}\sum_{i=1}^{N-k} x(n)x(n+k) \qquad et \qquad \hat{R}_{kk}(k) = \hat{R}_{kk}(-k) = \frac{1}{N-k}\sum_{i=1}^{N-k} x(n)x(n+k)$$

```
clc; clear all; close all;
randn('state', sum(100*clock));
N = 200; f0=1; Fe=5*f0; Te=1/Fe; t=(1:N)*Te; m=0; v=2;tt=(-N+1:N-1)*Te;ff=Fe/2*(-N+1:N-1)/N;
bb=m+sqrt(v)*randn(1,N);Rbb1=xcorr(bb,'unbiased');Rbb2=xcorr(bb,'biased');
ss=cos(2*pi*f0*t); Rss1=xcorr(ss,'unbiased');Rss2=xcorr(ss,'biased');
figure ; subplot(2,1,1); plot(tt,Rbb1); subplot(2,1,2) ; plot(tt,Rbb2);
figure ; subplot(2,1,1); plot(tt,Rss1); subplot(2,1,2) ; plot(tt,Rss2);
figure;plot(ff,abs(fftshift(fft(Rss1)/N)),'r',ff,abs(fftshift(fft(Rss2)/N)),'b');
```

```
legend('unbiased','biased'); xlabel('Frequence (Hz)'); ylabel('Amplitude');
```

1. Des 2 estimations classiques de l'autocorrélation pour le bruit, laquelle paraît la plus satisfaisante ?
2. Même question pour l'autocorrélation de la sinusoide
3. Observer la DSP dans chaque et commenter en justifiant les différences (amplitude, lobes secondaires, etc.)

Exercice 4 : Il existe différentes méthodes d'estimation de la densité spectrale de puissance. Les méthodes les plus couramment utilisées sont :

- Périodogramme simple: $\hat{S}_{xx}(f) = \frac{1}{N}|X_T(f)|^2$
- Corrélogramme: TF de la corrélation $\hat{S}_{xx}(f) = TF\left[\left\{\hat{R}_{xx}(k)\right\}\right]$
- Périodogramme moyenné**:** Une manière de réduire la variance de l'estimateur est de subdivisé l'intervalle de définition du signal x(n) en un certain nombre de sous-intervalles. On calcule alors la moyenne des péridogrames sur chacun des sous-intervalles (Méthode de Welch).
- Périodogramme moyenné avec recouvrement (overlapping): Cette fois-ci, on considère des sous intervalles de mêmes longueurs que précédemment mais avec des recouvrements d'une demi-longueur.

Le programme suivant permet d'estimer la DSP par les deux première méthodes

```
clc; close all; clear all;
N = 1024;    Te = 1/N;  Fe=N;    t = (0:N-1)*Te;  f1 = 100; f2 = 150; A1 = 2; A2=1.2; A3=4;
x = A1*cos(2*pi*f1*t) + A2*cos(2*pi*f2*t)+A3*randn(1,N); figure;plot(t,x);

% periodogramme
Sx1=abs(fft(x).^2)/N; ff1= Fe*(0:N/2-1)/N;

% spectre de l'autocorrélation
Rx=xcorr(x,'biased'); Sx2=fft(Rx); N2=length(Rx);
ff2=Fe*(0:N2/2-1)/N2;

% Méthode de welch
T=256; SXXF=[];k=1;
for ii=1:T:N-T
    xt=x(ii+1:ii+T); Sx3(k,:)=abs(fft(xt,N)).^2/T;   k=k+1;
end
Moy_Sx3=mean(Sx3);

% Périodogramme moyenné avec recouvrement
T=256; Sx4=[];k=1;
for ii=1:T/2:N-T
    xt=x(ii+1:ii+T); Sx4(k,:)=abs(fft(xt,N)).^2/T;   k=k+1;
end
Moy_Sx4=mean(Sx4);

%affichage
figure; plot(ff1,Sx1(1:N/2),'b',ff2,abs(Sx2(1:N2/2)),'r');
legend('Périodogramme','Correlogramme'); xlabel('Frequence (Hz)'); ylabel('Amplitude');
AXIS([min(ff2) max(ff2) min(Sx1) max(Sx1)]);
figure; plot(ff1,Sx1(1:N/2),'b',ff1,Moy_Sx3(1:N/2),'r');
legend('Périodogramme','Périodogramme moyenné'); xlabel('Frequence (Hz)');
ylabel('Amplitude');
figure; plot(ff1,Sx1(1:N/2),'b',ff1,Moy_Sx4(1:N/2),'r');
legend('Périodogramme','Périodogramme moyenné avec recouvrement'); xlabel('Frequence (Hz)');
ylabel('Amplitude');
figure; plot(ff1,Moy_Sx3(1:N/2),'b',ff1,Moy_Sx4(1:N/2),'r');
legend('Périodogramme moyenné','Périodogramme moyenné avec recouvrement'); xlabel('Frequence
(Hz)'); ylabel('Amplitude');
```

1. Comparer les 4 techniques en faisant varier A1, A2 et A3 et commenter.
2. Rapprocher les fréquences et observer.
3. Commenter les quatre techniques d'estimation de la DSP en faisant varier N.
4. Quelle est l'avantage et l'inconvénient du moyennage?