

**Statistique des Processus  
2009-2010**

**Fiche de TD/TP : Statistique des processus**

Estimation

**Ex.1.**

Soient les observations  $\{x_t, t = 1, \dots, 100\}$  d'un processus  $X$ , à partir desquelles on obtient les valeurs  $\bar{X} = 46.93$  et  $\hat{\gamma}(0) = 1382.2$ ,  $\hat{\gamma}(1) = 1114.4$ ,  $\hat{\gamma}(2) = 591.73$ ,  $\hat{\gamma}(3) = 96, 216$ .

On centre alors le processus  $X$  ; le nouveau processus centré est noté  $Y$ .

1. Trouver les estimateurs de Yule-Walker de  $\phi_1$ ,  $\phi_2$  et  $\sigma^2$  du modèle  $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$ , avec  $\varepsilon$  bruit blanc de variance  $\sigma^2$ .

En supposant que les observations sont réellement la réalisation d'un processus  $AR(2)$ , donner les intervalles de confiance à 95 % pour  $\phi_1$  et  $\phi_2$ .

Application SAS sur la série E921.

2. Même question en considérant les EMV (sur la série E921).

**Ex.2.**

Considérons  $X$  processus  $AR(2)$  satisfaisant

$$X_t = \phi X_{t-1} + \phi^2 X_{t-2} + \varepsilon_t, \quad \text{avec } (\varepsilon_t) \text{ b.b de variance } \sigma^2.$$

1. Pour quelles valeurs de  $\phi$  le processus est-il causal ?
2. On a calculé, à partir des observations  $x_1, \dots, x_{200}$ ,  $\hat{\gamma}(0) = 6.06$  et  $\hat{\gamma}(1) = 0.687$ . Donner les estimateurs de  $\phi$  et  $\sigma^2$  en résolvant les équations de Yule-Walker.  
Remarque : si vous trouvez plus d'une solution, choisissez celle qui est causale.

**Ex.3.**

On dispose de 200 observations d'une série  $X$ , de moyenne  $\mu$ , pour laquelle la sortie SAS nous donne :

$$\begin{aligned} \text{sample mean : } & \bar{X}_{200} = 3.82 \\ \text{sample variance : } & \hat{\gamma}(0) = 1.15 \\ \text{sample ACF : } & \hat{\rho}(1) = .427 \\ & \hat{\rho}(2) = .475 \\ & \hat{\rho}(3) = .169 \end{aligned}$$

1. Peut-on supposer à partir de ces résultats que  $(X_t - \mu)$  est un bruit blanc ?
2. Supposons que l'on modélise  $(X_t - \mu)_t$  par un  $AR(2)$ , donné par

$$X_t - \mu - \phi_1(X_{t-1} - \mu) - \phi_2(X_{t-2} - \mu) = \varepsilon_t,$$

avec  $(\varepsilon_t)$  bruit blanc fort, de variance  $\sigma^2$ .  
Donner les estimateurs de  $\mu$ ,  $\phi_1$ ,  $\phi_2$  et  $\sigma^2$ .

3. Peut-on conclure que  $\mu = 0$  ?
4. Construire des intervalles de confiance à 95 % pour  $\phi_1$  et  $\phi_2$ .
5. En supposant que les données sont simulées à partir d'un  $AR(2)$ , donner des estimateurs des autocorrélations partielles empiriques pour  $h \geq 1$ .

## Propriétés asymptotiques des EMV

### Ex.4.

Soit  $\beta := {}^t(\Phi_1, \dots, \Phi_p, \theta_1, \dots, \theta_q)$  paramètres d'un processus  $ARMA(p, q)$  ( $p \geq 1, q \geq 1$ ), et  $\hat{\beta}$  son EMV.

Alors pour  $n$  assez grand, on a  $\hat{\beta} \sim AN(\beta, n^{-1}V(\beta))$ , où

$$V(\beta) = \sigma^2 \begin{pmatrix} E[\mathbf{U}_t {}^t\mathbf{U}_t] & E[\mathbf{U}_t {}^t\mathbf{V}_t] \\ E[\mathbf{V}_t {}^t\mathbf{U}_t] & E[\mathbf{V}_t {}^t\mathbf{V}_t] \end{pmatrix}^{-1},$$

avec  $\mathbf{U}_t = {}^t(U_t, \dots, U_{t+1-p})$ ,  $\mathbf{V}_t = {}^t(V_t, \dots, V_{t+1-q})$ , ( $U_t$ ) et ( $V_t$ ) étant des processus autorégressifs définis par  $\Phi(B)\mathbf{U}_t = \varepsilon_t$  et  $\Theta(B)\mathbf{V}_t = \varepsilon_t$ .

1. Montrer que dans le cas d'un processus  $AR(p)$ ,  $V(\beta)$  est la matrice de covariance asymptotique obtenue pour les estimateurs de Yule-Walker. Donner  $V(\beta)$  explicitement pour  $p = 1$  et  $p = 2$ .
2. Quel résultat obtient-on pour un processus  $MA(q)$  ?

On en déduit des régions de confiance au niveau  $1 - \alpha$  pour les paramètres  $(\Phi_1, \dots, \Phi_p, \theta_1, \dots, \theta_q)$  d'un  $ARMA(p, q)$  causal et inversible, à savoir

$$\{\beta \in \mathbb{R}^{p+q} : {}^t(\beta - \hat{\beta})V^{-1}(\beta - \hat{\beta}) \leq n^{-1}\chi_{p+q}^2(1 - \alpha)\},$$

et en désignant le  $j$ ème élément diagonal de  $V(\hat{\beta})$  par  $v_{jj}(\hat{\beta})$ , on a comme région de confiance au niveau  $1 - \alpha$  pour  $\beta_j : \{\beta \in \mathbb{R} : |\beta - \hat{\beta}_j| \leq n^{-1/2}q_{1-\alpha/2}(v_{jj}(\hat{\beta}))^{1/2}\}$ , avec  $q_{1-\alpha/2}$  quantile de la loi normale centrée réduite d'ordre  $1 - \alpha/2$ .

## Prévision

### Ex.5.

1. On considère le processus  $X = (X_t, t \in \mathbb{Z})$  défini par  $X_t = \varepsilon_t - \alpha\varepsilon_{t-1}$ , avec  $\varepsilon$  bruit blanc de variance  $\sigma^2$  et  $|\alpha| < 1$ .

a. Montrer que la prédiction linéaire optimale de  $X_{t+1}$  à partir du passé  $\{X_j, -\infty < j \leq t\}$  est

$$\hat{X}_{t+1} = - \sum_{j=1}^{\infty} \alpha^j X_{t+1-j}.$$

b. Quelle est la variance de l'erreur de prédiction ?

c. Application SAS : on choisit  $\alpha = 0.7$ ; simuler 2000 valeurs du modèle associé, puis comparer la prévision fournie par SAS avec celle que donne le calcul théorique.

2. On considère cette fois un processus  $AR(p)$ ,  $Y = (Y_t, t \in \mathbb{Z})$ , de polynôme autorégressif  $\Phi(z) =$

$$1 + \sum_{i=1}^p \Phi_i z^i.$$

Utiliser les équations de projection pour vérifier que le prédicteur linéaire optimal de  $Y_{t+1}$  connaissant tout le passé  $\{Y_j, -\infty < j \leq t\}$  est

$$\hat{Y}_{t+1} = - \sum_{j=1}^p \Phi_j Y_{t+1-j}.$$

**Ex.6.**

Soit  $X = (X_t, t \in \mathbb{Z})$  processus  $ARIMA(p, d, q)$  centré défini par

$$(1 - B)^d \Phi(B)X_t = \Theta(B)\varepsilon_t,$$

avec  $(\varepsilon_t, t \in \mathbb{Z})$  bruit blanc de variance  $\sigma^2$  et

$Z_{-1} = (X_{-1}, \dots, X_{-p-d}, \varepsilon_{-1}, \dots, \varepsilon_{-q})$  non corrélé avec  $(\varepsilon_k, k \geq 0)$ .

On suppose  $\Phi$  et  $\Theta$  de racines de modules strictement supérieurs à 1.

On suppose que l'on observe  $(Z_{-1}, X_0, \dots, X_T)$  et l'on veut prédire  $X_{T+h}, h > 0$ .

On note alors  $\hat{X}_t^T = P_{[Z_{-1}, X_0, \dots, X_T]}(X_t)$  et  $\hat{\varepsilon}_t^T = P_{[Z_{-1}, X_0, \dots, X_T]}(\varepsilon_t)$ .

1. a. Montrer que  $\hat{X}_t^T = X_t, 0 \leq t \leq T$  et en déduire que  $\hat{\varepsilon}_t^T = \varepsilon_t, 0 \leq t \leq T$ .  
 b. Montrer que  $\forall t > T, \hat{\varepsilon}_t^T = 0$  (i.e.  $\forall h > 0, \hat{\varepsilon}_{T+h}^T = 0$ ).  
 (Ce résultat sera utilisé dans toute la suite du problème).  
 c. En déduire que  $\forall t > T + q, (1 - B)^d \Phi(B)\hat{X}_t^T = 0$ .
2. a. Calculer la fonction de prévision  $h \rightarrow \hat{X}_{T+h}^T$ , avec  $h > 0$ , associée respectivement aux modèles définis par :  
 i)  $(1 - B)X_t = \varepsilon_t$   
 ii)  $(1 - B)X_t = (1 - \beta B)\varepsilon_t, \beta \neq 0, 1$ .  
 b. Pour chacun des deux modèles (i) et (ii), exprimer l'erreur de prédiction  $e_h = X_{T+h} - \hat{X}_{T+h}^T$ ,  $h > 0$ , puis en calculer la variance  $var(e_h)$ . Quelle est la limite de cette variance quand  $h \rightarrow \infty$ ? comment l'interpréter?
3. On considère le cas particulier où  $X$  est un processus  $MA(q)$ . Calculer les prédictions  $\hat{X}_{T+h}^T$  pour  $h > q$ .

**Ex.7.**

Soit  $X = (X_t, t \in \mathbb{Z})$  le processus défini par

$$(1 - \frac{1}{4}B)X_t = 3 + (1 - 0.5B)\varepsilon_t, \quad t \in \mathbb{Z},$$

avec  $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$  b.b. de var  $\sigma^2$ . On suppose que l'on observe  $\varepsilon_{-1}, X_{-1}, X_0, \dots, X_T$ .

Calculer  $\varepsilon_t$  en fonction de  $\varepsilon_{-1}, X_{-1}, X_0, \dots, X_T$  et en déduire la prédiction de  $X_{T+1}$ . Calculer complètement la fonction de prévision et la suite des erreurs de prédiction.

**Ex.8.**

Modéliser les série E911, E921, E923, E951, E1042, E1062, E1331 et E1332 à l'aide de la procédure ARIMA de SAS.

**Ex.9.**

Modéliser la série AIRPASS à l'aide de la procédure ARIMA de SAS.

Identifier les modèles possibles.

Choisir un modèle en le justifiant via les critères adéquats.

Prévoir la série sur 1 an.