*Université Paris I, Panthéon - Sorbonne*

**Première Année Master M.A.E.F.** 2016 − 2017

# Statistiques II

Contrôle continu n°2, avril 2017

*Examen de 1h30. Tout document ou calculatrice est interdit.*

1. **(7 points + 2 points)** On suppose que $M = (M_k)_{k \in \mathbb{Z}}$ est un processus centré stationnaire du second ordre tel qu'il existe $\alpha > 0$ et $\beta \in \mathbf{R}$ vérifiant pour tout $k \in \mathbf{Z}$:

$$\mathbb{E}\big(M_k^2\big) = \alpha, \quad \mathbb{E}\big(M_k M_{k+1}\big) = \beta \quad \text{et} \quad \mathbb{E}\big(M_k M_{k+i}\big) = 0 \quad \text{pour} \quad i \geq 2. \tag{1}$$

   (a) Montrer que l'on a nécessairement $|\beta| \leq \alpha$ **(1pt)**.

   (b) Rappeler l'expression de la densité spectrale de $M$. En déduire que nécessairement $2|\beta| \leq \alpha$ **(1.5pts)**.

   (c) Soit $(\xi_k)_{k \in \mathbf{Z}}$ une suite de v.a.i.i.d. centrées et de variance $\sigma^2 > 0$ et soit $(N_k)_{k \in \mathbf{Z}}$ un processus tel que pour tout $k \in \mathbf{Z}$, $N_k = \xi_k + c\,\xi_{k-1}$ avec $c \in \mathbf{R}$. Quel type de processus est $(N_k)$ **(0.5pts)**? Déterminer l'autocovariance de $(N_k)$ **(1pt)**. En déduire que $(N_k)_{k \in \mathbf{Z}}$ vérifie (1) en précisant $c$ et $\sigma^2$ en fonction de $\alpha$ et $\beta$ **(1pt)**.

   (d) Si $M$ est un processus gaussien vérifiant (1), montrer que $M$ est nécessairement un processus MA(1) et donner son écriture à l'aide d'un bruit blanc gaussien $\varepsilon = (\varepsilon_k)_{k \in \mathbb{Z}}$ dont vous préciserez la variance **(2pts)**.

   (e) **(Question facultative sur 2 points)** Montrer que si $M$ est la somme d'un processus MA(1) et d'un processus ARCH($p$) stationnaire d'ordre 2, indépendant de ce MA(1), alors $M$ vérifie également la relation (1). Y-a-t-il une contradiction avec la propriété précédente?

2. **(7 points)** On considère deux processus stationnaires indépendants $X = (X_k)_{k \in \mathbb{Z}}$ et $Y = (Y_k)_{k \in \mathbb{Z}}$ tels que pour tout $k \in \mathbb{Z}$,

$$X_k = \rho_1 X_{k-1} + u_k, \qquad \text{avec } 0 < |\rho_1| < 1, \tag{2}$$
$$Y_k = \rho_2 Y_{k-1} + v_k, \qquad \text{avec } 0 < |\rho_2| < 1, \tag{3}$$

   où $u = (u_k)_{k \in \mathbb{Z}}$ et $v = (v_k)_{k \in \mathbb{Z}}$ sont deux bruits blancs gaussiens indépendants de variances respectives $\sigma_u^2$ et $\sigma_v^2$. Soit enfin $Z = (Z_k)_{k \in \mathbb{Z}}$ tel que

$$Z_k = X_k + Y_k \quad \text{pour tout } k \in \mathbb{Z}.$$

   (a) Si $\rho_1 = \rho_2$, quel type de processus est $Z$ **(1pt)**?

   (b) Si $\rho_1 \neq \rho_2$, montrer que $Z$ est un processus ARMA$(2,1)$ dont vous donnerez l'écriture (on pourra considérer $Z_k - (\rho_1 + \rho_2) Z_{k-1} + \rho_1 \rho_2 Z_{k-2}$, puis s'aider de la question 1.(d)) **(4pts)**.

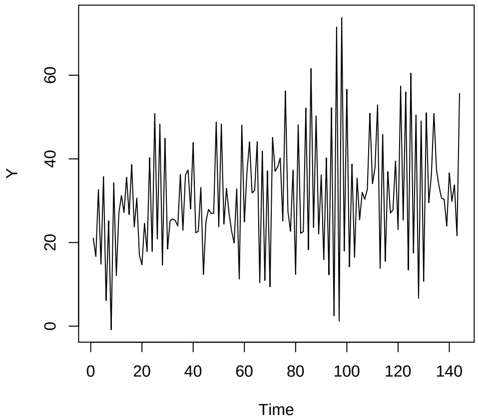
   (c) Quelle relation doivent satisfaire $\rho_1$, $\rho_2$, $\sigma_u^2$ et $\sigma_v^2$ pour que $Z$ soit un AR(2) **(2pts)**?

3. **(9 points)** Voici des simulations effectuées avec le logiciel R.

   (a) On tape d'accord les commandes suivantes:

```
epsilon=3*rnorm(144)
X=0
for (j in c(1:143))
 {X[j+1]=-0.7*X[j]+epsilon[j+1]-2*epsilon[j]}
t=c(1:144)
Y=15+4*log(t)+7*cos(pi*(t-2)/6)+X
ts.plot(Y)
```

Voici le graphe obtenu:




*Questions: Quel est le processus simulé par le vecteur $X$ (préciser tous ses paramètres)* **(0.5pts)***? Qelles sont les tendances et saisonnalités de $Y$* **(1pt)***? Quelle est sa variance (théorique)* **(2pts)***?*

(b) Voici les commandes tapées ensuite:

```
Z1=t
Z2=t^2
Z3=sqrt(t)
Z4=cos(pi*t/6)
Z5=sin(pi*t/6)
reg=lm(Y~Z1+Z2+Z3+Z4+Z5)
summary(reg)
```

Voici les résultats obtenus:

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.6447975 11.1173261   0.957 0.339991
Z1          -0.4077850  0.5735870  -0.711 0.478322
Z2           0.0009389  0.0017558   0.535 0.593669
Z3           5.3872548  5.0063068   1.076 0.283763
Z4           3.4346042  1.5587953   2.203 0.029227 *
Z5           6.0576026  1.5746636   3.847 0.000182 ***

Residual standard error: 13.22 on 138 degrees of freedom
Multiple R-squared: 0.1789,     Adjusted R-squared: 0.1492
```

*Questions: Qu'a-t-on fait par ces commandes* **(0.5pts)***? Que représentent les valeurs* 10.6447975*,* 0.593669 *et* 1.5746636 **(1.5pts)***? Que conclure à la lecture de ces résultats* **(0.5pts)***?*

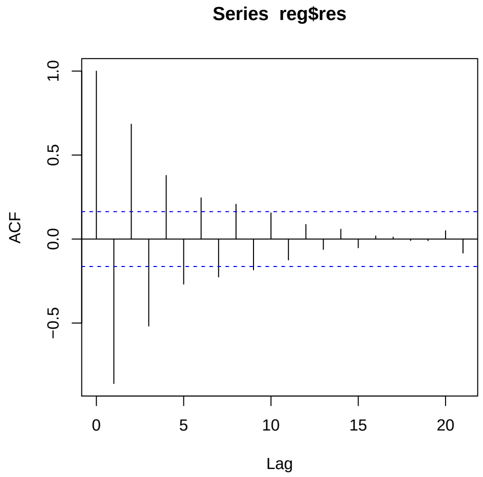
(c) On tape enfin les commandes suivantes:

```
reg=lm(Y~Z1+Z3+Z4+Z5)
summary(reg)
reg=lm(Y~Z3+Z4+Z5)
summary(reg)
acf(reg$res)
```

Voici les résultats obtenus ainsi que le graphe tracé:

```
lm(formula = Y ~ Z1 + Z3 + Z4 + Z5)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  15.2777     6.9492   2.198   0.0296 *
Z1           -0.1099     0.1367  -0.804   0.4225
Z3            2.9408     2.0282   1.450   0.1493
Z4            3.4318     1.5548   2.207   0.0289 *
Z5            5.9595     1.5599   3.820   0.0002 ***
```

```
Residual standard error: 13.19 on 139 degrees of freedom
Multiple R-squared: 0.1772,     Adjusted R-squared: 0.1536

lm(formula = Y ~ Z3 + Z4 + Z5)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  20.1762     3.3436   6.034 1.35e-08 ***
Z3            1.3403     0.3928   3.412 0.000843 ***
Z4            3.4173     1.5527   2.201 0.029384 *
Z5            5.9084     1.5567   3.796 0.000219 ***

Residual standard error: 13.17 on 140 degrees of freedom
Multiple R-squared: 0.1734,     Adjusted R-squared: 0.1557
```




*Questions: Expliquez pourquoi on a répété deux fois la commande lm et quel(s) critère(s) guide(nt) cette répétition* **(0.5pts)***? Est-on satisfait du modèle obtenu* **(0.5pts)***? Pouvait-on s'attendre aux valeurs numériques* 3.4173 *et* 5.9084 **(0.5pts)***? Expliquez ce que représente le graphe* **(0.5pts)***. La valeur numérique décrite par la seconde barre est approximativement de* $-0.85$*. Expliquez ce que l'on pouvait espérer comme valeur théorique* **(1pt)***.*