

ECN 6578A, Économétrie des marchés financiers,  
Hiver 2023

Cours 2

William McCausland

2023-01-15

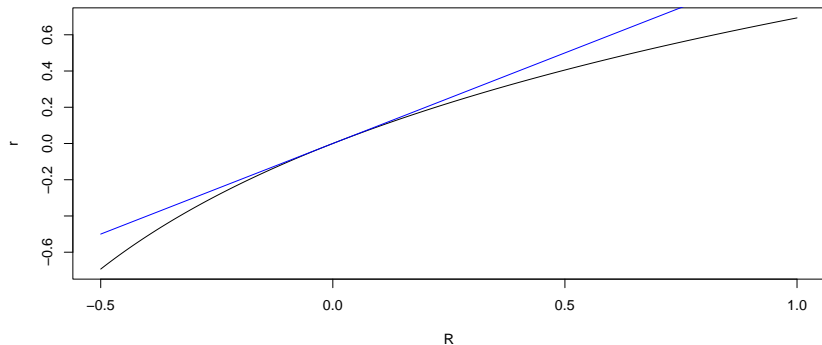
# Plan

1. Rendements de plusieurs périodes et log rendements
2. Asymétrie et Aplatissement
3. Autocovariance et autocorrélation
4. Faits empiriques

## La relation entre le log-rendement $r$ et le rendement $R$

- ▶ Rappelons que  $r_t \equiv \log(1 + R_t)$ .
- ▶ La fonction  $r = \log(1 + R)$  est concave et  $r \leq R$

```
R = seq(-0.5, 1.0, by=0.01)
plot(R, log(1+R), xlab="R", ylab="r", type='l')
lines(R, R, col='blue')
```



## Rendements multipériodes

- ▶ Rendements multipériodes : ( $k$  périodes, net, bruts, log)

$$R_t[k] \equiv \frac{P_t - P_{t-k}}{P_{t-k}}, \quad 1 + R_t[k] \equiv \frac{P_t}{P_{t-k}}, \quad r_t[k] \equiv \log(1 + R_t[k])$$

- ▶ Notez que

$$1 + R_t[k] = \frac{P_{t-k+1}}{P_{t-k}} \cdots \frac{P_{t-1}}{P_{t-2}} \frac{P_t}{P_{t-1}} = \prod_{\tau=t-k+1}^t (1 + R_\tau)$$

$$r_t[k] = (p_{t-k+1} - p_{t-k}) + \cdots + (p_{t-1} - p_{t-2}) + (p_t - p_{t-1}) = \sum_{\tau=t-k+1}^t r_\tau$$

- ▶ Avantage aux log-rendements pour l'analyse intertemporelle

## Intuition pour le rendement continument composé

- ▶ Rendement composé : divisez une période en  $n$  sous-périodes,
  - ▶ soit  $r/n$  le rendement net à chaque sous-période,
  - ▶ soit  $R$  le rendement net pour la période entière.
- ▶ Alors

$$(1 + R) = (1 + r/n)^n.$$

- ▶  $1 + R$  est croissant en  $n$  grace au rendements composés ...
- ▶ ... mais il y a une limite : (rendement continument composé)

$$(1 + R) = \lim_{n \rightarrow \infty} (1 + r/n)^n = e^r.$$

- ▶ En logs (Séries Mercator) :

$$\lim_{n \rightarrow \infty} n \log(1 + r/n) = \lim_{n \rightarrow \infty} n \left[ \frac{r}{n} - \frac{1}{2} \left( \frac{r}{n} \right)^2 + \frac{1}{3} \left( \frac{r}{n} \right)^3 - \dots \right] = r.$$

- ▶ Pour un log-rendement constant  $r$  et une valeur initial  $V_0$  à  $t_0$ ,

$$V_t = V_0 e^{r(t-t_0)}.$$

# Annualisation

- ▶ Mettons que l'unité de temps est l'année.
- ▶ Question : Quelle rendement net annuel constant  $R$  pendant  $k$  ans donne-t-il un rendement net après  $k$  périodes de  $R_t[k]$ ?
- ▶ C'est à dire quelle valeur de  $R$  vérifie  $1 + R_t[k] = (1 + R)^k$ ?
- ▶ L'annualisation d'un rendement multipériode :

$$(1 + R) = (1 + R_t[k])^{1/k}, \quad R = (1 + R_t[k])^{1/k} - 1.$$

- ▶ En logarithmes :

$$r = \frac{1}{k} \log(1 + R_t[k]) = r_t[k]/k.$$

- ▶ Remarquez encore la linéarité des log rendements avec l'aggrégation temporelle.

## Asymétrie et aplatissement

- ▶ Définitions : pour la variable aléatoire  $X \sim (\mu, \sigma^2)$ , l'asymétrie  $S$  et l'aplatissement  $K$  de la population sont

$$S = E[(X - \mu)^3]/\sigma^3, \quad K = E[(X - \mu)^4]/\sigma^4.$$

- ▶ Pour une v.a. gaussienne,  $S = 0$  et  $K = 3$ .
- ▶ Pour l'asymétrie et l'aplatissement de l'échantillon, on utilise les moments analogues de l'échantillon.
- ▶ Asymétrie et aplatissement dans le tableau 1.2 de Tsay.
  - ▶ Asymétrie (skewness) souvent négative—actions et indices.
  - ▶ Aplatissement (kurtosis) toujours plus grand que 3—actions et indices.
  - ▶  $S$  et  $K$  plus sévères pour  $r_t$  que pour  $R_t$ .
  - ▶  $S$  et  $K$  moins sévères (mais pas toujours) pour  $r_t$  mensuels que pour  $r_t$  journaliers. (Comme si un théorème central limite s'applique)

## Aspects non-gaussiens du rendement IBM I

```
# Données IBM, r_t journalier, 1970-2008  
da = read.table('d-ibm3dx7008.txt', header=T)  
r = da$rtn  
mu = mean(r); sigma = sd(r)  
  
S_ch = mean((r-mu)^3)/sigma^3;  
cat(sprintf('Asymétrie, Sch=%f', S_ch))
```

Asymétrie, Sch=0.061399

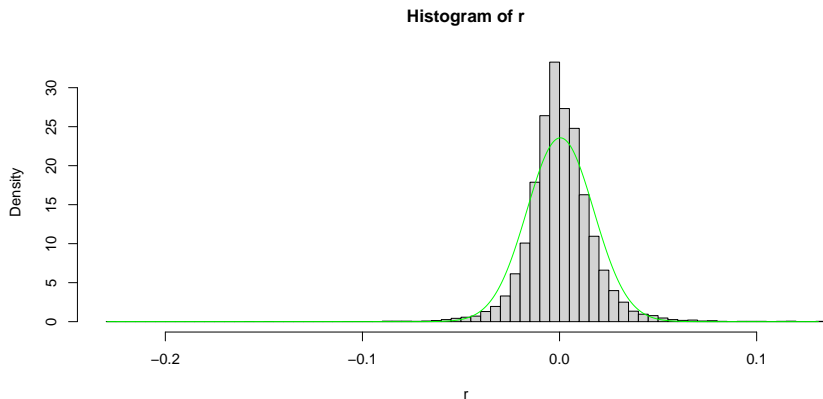
```
K_ch = mean((r-mu)^4)/sigma^4;  
cat(sprintf('Aplatissement, Kch=%f', K_ch))
```

Aplatissement, Kch=12.916359



## Aspects non-gaussiens du rendement IBM II

```
hist(r, breaks=100, freq=FALSE)  
x = seq(min(r), max(r), by=0.0005)  
lines(x, dnorm(x, mu, sigma), col='green')
```



## Tests de normalité basés sur $S$ et $K$

- ▶ L'hypothèse nulle ici :  $H_0 : r_t \sim \text{iid } N(\mu, \sigma^2)$ ,  $t = 1, \dots, T$ .
- ▶ Loi asymptotique de l'assymétrie de l'échantillon sous  $H_0$ :

$$z_S = \frac{\hat{S}}{\sqrt{6/T}} \sim N(0, 1).$$

- ▶ Loi asymptotique de l'aplatissement de l'échantillon sous  $H_0$ :

$$z_K = \frac{\hat{K} - 3}{\sqrt{24/T}} \sim N(0, 1).$$

- ▶ Il s'avère que  $z_S$  et  $z_K$  sont asymptotiquement indépendents.
- ▶ Si un nombre  $\nu$  de v.a.  $N(0, 1)$  sont indépendentes, la somme de leurs carrées à une loi  $\chi^2(\nu)$ .
- ▶ Loi asymptotique de la statistique Jarque-Bera sous  $H_0$ :

$$\text{JB} = \frac{\hat{S}^2}{6/T} + \frac{(\hat{K} - 3)^2}{24/T} \sim \chi^2(2).$$

## Tests statistiques unilatéraux et bilatéraux

- ▶ Un test avec la statistique  $\hat{S}$  est souvent bilatéral—l'hypothèse alternative est  $S \neq 0$ . On peut défendre aussi un test unilatéral avec une hypothèse alternative  $S < 0$ .
- ▶ Un test avec la statistique  $\hat{K}$  peut être unilatéral, auquel cas l'hypothèse alternative devrait être  $K > 3$ , ou bilatéral, auquel cas l'hypothèse alternative est  $K \neq 3$ .
- ▶ Un test avec la statistique JB est toujours unilatéral. On ne rejette pas la normalité si l'aplatissement et l'asymétrie sont plus près de 3 et 0 que d'habitude sous l'hypothèse nulle.

Exemple 1: test de l'hypothèse  $H_0$  contre  $H_1 : S \neq 0$  à un niveau de 5% (rendement IBM)

```
T = length(r)           # Taille de l'échantillon
alpha = 0.05           # Niveau du test
z_S = S_ch/sqrt(6/T)   # Statistique test
c_1 = qnorm(alpha/2)   # Valeurs critiques
c_2 = qnorm(1-alpha/2)
cat(sprintf('Statistique test : %f', z_S))
```

Statistique test : 2.487093

```
cat(sprintf('Région de non rejet : [%f, %f]', c_1, c_2))
```

Région de non rejet : [-1.959964, 1.959964]

Exemple 2: test de l'hypothèse  $H_0$  contre  $H_1 : K > 3$  à un niveau de 1% (rendement IBM)

```
alpha = 0.01           # Niveau du test
z_K = (K_ch-3)/sqrt(24/T) # Statistique test
c = qnorm(1-alpha)     # Valeur critique
cat(sprintf('Statistique test : %f', z_K))
```

Statistique test : 200.841983

```
cat(sprintf('Région de non rejet : [-inf, %f]', c))
```

Région de non rejet : [-inf, 2.326348]

## Attention : tests multiples!

Trois tests asymptotiques de l'hypothèse  $H_0$  de normalité:

1. Test avec la statistique  $z_S$ ,  $H_0$  contre  $H_1 : S \neq 0$ . La région de non-rejet est

$$q_{N,\alpha/2} < z_S < q_{N,1-\alpha/2}.$$

où  $\alpha$  est le niveau du test;  $q_{N,p}$ , la quantile  $p$  d'une loi  $N(0, 1)$ .

2. Test avec la statistique test  $z_K$ ,  $H_0$  contre  $H_1 : K \neq 3$ . La région de non-rejet est

$$q_{N,\alpha/2} < z_K < q_{N,1-\alpha/2}.$$

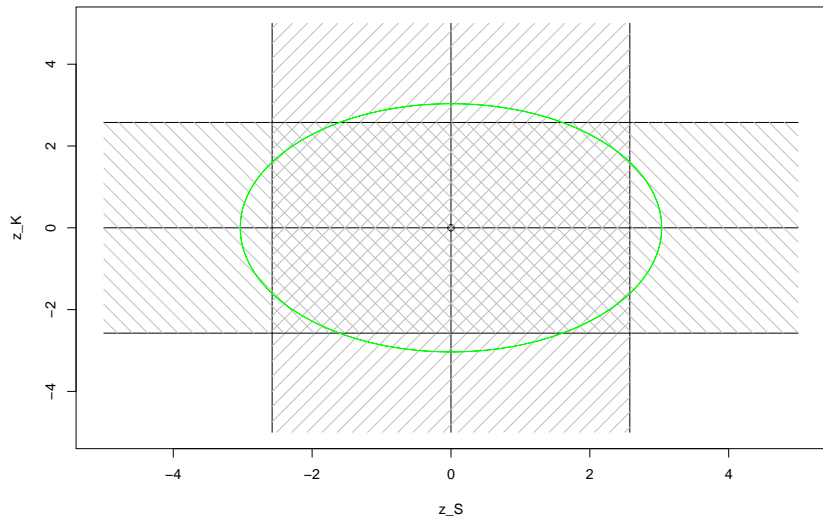
3. Test avec la statistique test JB,  $H_0$  contre  $H_1 : S \neq 0$  ou  $K \neq 3$ . La région de non-rejet est

$$JB = z_S^2 + z_K^2 < q_{\chi^2(2),1-\alpha},$$

où  $q_{\chi^2(2),p}$  est la quantile  $p$  d'une loi  $\chi^2(2)$ .

Si on fait les deux tests 1 et 2, la probabilité d'au moins un rejet sous  $H_0$  excède  $\alpha$ .

# Régions de non-rejet pour les trois tests, $\alpha = 0.01$



## Stationnarité

- ▶ Une série  $r_t$  est stationnaire ssi pour chaque  $k, t_1, \dots, t_k$  et  $\tau$ , les deux distributions suivantes sont identiques :

$$(r_{t_1}, r_{t_2}, \dots, r_{t_k}), \quad (r_{t_1-\tau}, r_{t_2-\tau}, \dots, r_{t_k-\tau}).$$

- ▶ Une série  $r_t$  est covariance-stationnaire ssi  $E[r_t] \equiv \mu$  et  $\gamma_k \equiv \text{Cov}[r_t, r_{t-k}]$  ne dépend pas de  $t$ .
- ▶ Pertinence des deux types de stationnarité
  - ▶ Des hypothèses que le futur ressemble, dans un sens, au passé.
- ▶ Pertinence de covariance-stationnarité
  - ▶ Sauf pour l'existence des moments, elle est moins forte.
  - ▶ Les variances et covariances des fonctions linéaires des v.a. dépendent seulement des variances et covariances de ces v.a.
  - ▶ Les fonctions d'autocorrélation et d'autocorrélation partielle sont bien définies.
  - ▶ Conditions de covariance-stationnarité en termes des coefficients ARMA(p,q)



## Non-corrélation versus indépendance I

- ▶ Tirer un paramètre d'échelle commun  $\omega \sim \chi^2(\nu)$ .
- ▶ Sachant  $\omega$ , tirer  $X_1$  et  $X_2$ , conditionnellement iid et gaussiens :

$$X_1, X_2 | \omega \sim \text{iid } N(0, \nu/\omega).$$

- ▶  $X_1$  et  $X_2$  ont chacune une loi marginale  $t$  de Student:

$$X_i \sim t(\nu), \quad i = 1, 2.$$

- ▶  $X_1$  et  $X_2$  sont non-corrélés :

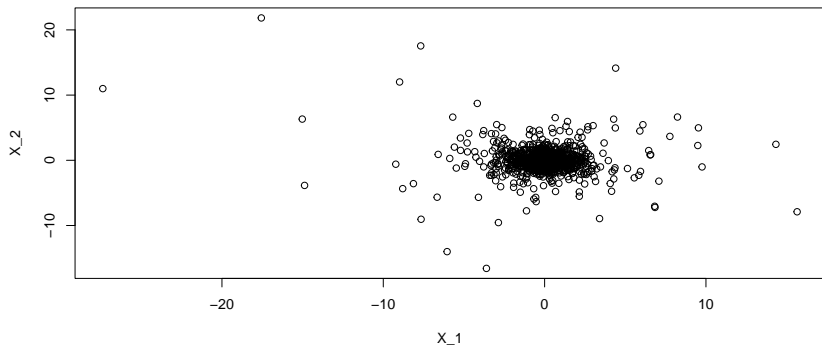
$$\text{Cov}[X_1, X_2] = E[\text{Cov}[X_1, X_2 | \omega]] + \text{Cov}[E[X_1 | \omega], E[X_2 | \omega]] = 0$$

- ▶  $X_1$  et  $X_2$  ne sont pas indépendents :

$$\begin{aligned} \text{Cov}[X_1^2, X_2^2] &= E[\text{Cov}[X_1^2, X_2^2 | \omega]] + \text{Cov}[E[X_1^2 | \omega], E[X_2^2 | \omega]] \\ &= 0 + \text{Var}[\nu/\omega] \neq 0. \end{aligned}$$

## Non-corrélation versus indépendance II

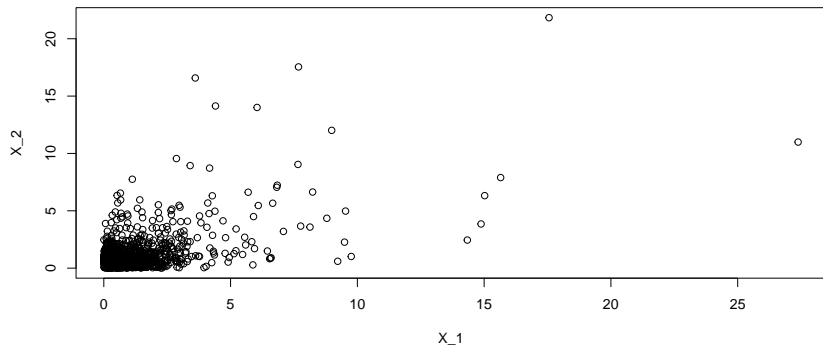
```
nu = 2  
T = 1000; set.seed(12345)  
omega = rchisq(T, nu)  
X1 = rnorm(T, 0, sqrt(nu/omega))  
X2 = rnorm(T, 0, sqrt(nu/omega))  
plot(X1, X2, xlab='X_1', ylab='X_2')
```



## Non-corrélation versus indépendance III

Remarquez la dépendance entre  $|X_1|$  et  $|X_2|$  :

```
plot(abs(X1), abs(X2), xlab='X_1', ylab='X_2')
```



# Fonctions d'autocorrélation, d'autocorrélations partielles

Fonction d'autocorrélation:

$$\rho_{\tau} = \text{CORR}[r_t, r_{t-\tau}] = \text{Cov}[r_t, r_{t-\tau}] / \text{Var}[r_t].$$

Fonction d'autocorrélation partielle:

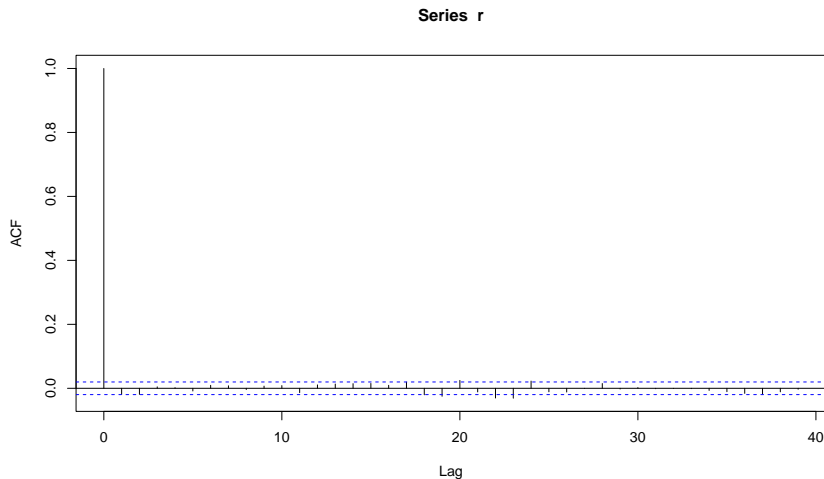
$$\phi_{\tau} = \text{CORR}[r_t, r_{t-\tau} | r_{t-1}, \dots, r_{t-\tau+1}].$$

Quand  $r_t$  est covariance stationnaire,  $\phi_{\tau}$  est le coefficient  $\phi_{\tau\tau}$  de  $r_{t-\tau}$  dans la régression

$$r_t = \phi_{1\tau} r_{t-1} + \dots + \phi_{\tau\tau} r_{t-\tau} + \epsilon_t, \quad E[\epsilon_t | r_{t-1}, \dots, r_{t-\tau}] = 0.$$

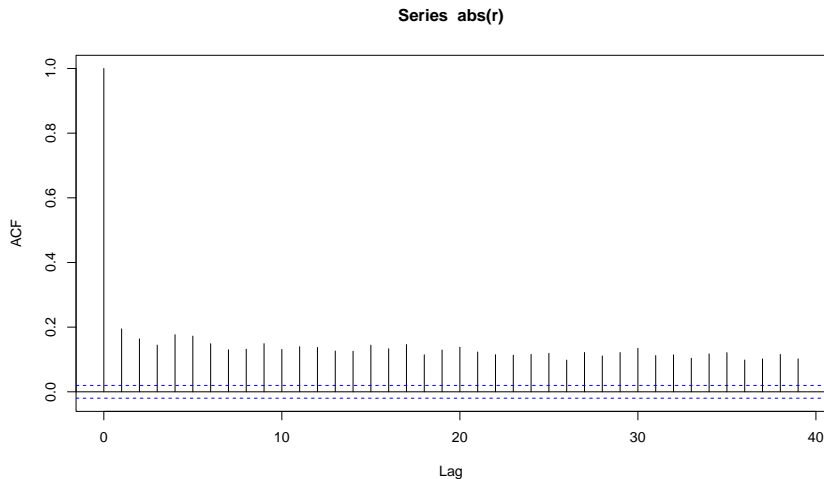
# Fonction d'autocorrélation de $r_t$ (rendement IBM)

acf(r)



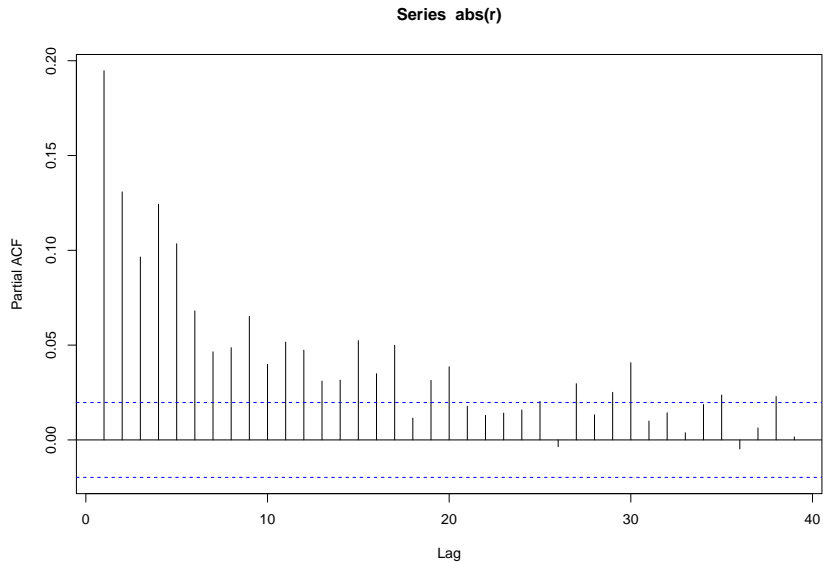
# Fonction d'autocorrélation de $|r_t|$ (rendement IBM)

```
acf(abs(r))
```



# Fonction d'autocorrélation partielle de $|r_t|$

```
pacf(abs(r))
```



## Tests de non-autocorrélation

- ▶ Hypothèse  $H_0$  :  $r_t$  est iid, de variance finie.
- ▶ Les lois ici sont asymptotiques, sous  $H_0$ .
- ▶ Test avec une autocorrélation isolée :

$$\hat{\rho}_k \sim N(0, T^{-1}).$$

- ▶ Test portmanteau (avec plusieurs autocorrélations) Box-Pierce :

$$T \sum_{l=1}^m \hat{\rho}_l^2 \sim \chi^2(m).$$

- ▶ Test portmanteau Ljung-Box (avec correction de biais)

$$T(T+2) \sum_{l=1}^m \frac{\hat{\rho}_l^2}{T-l} \sim \chi^2(m).$$



## Faits empiriques de Cont (2001)

1. Peu d'autocorrelation ( $r_t$ ) sauf à une échelle de moins de 20 minutes.
2. Aplatissement inconditionnel plus grand que 3.
3. Asymétrie négative pour les actifs et les indices.
4. Loi de  $r_t$  "plus gaussienne" à une échelle temporelle plus grande.
5. Variation de volatilité à toute échelle temporelle.
6. Persistance de volatilité.
7. Aplatissement conditionnel moins grand que l'aplatissement inconditionnel, mais toujours plus grand que 3.
8. La fonction d'autocorrélation de  $|r_t|$  décroît selon une loi de puissance (plus lentement que de taux exponentiel). (Longue mémoire)
9. Effet de levier: changement de volatilité négativement corrélé avec le rendement actuel.
10. Corrélacion positive entre volume et volatilité.

# Cours 3, la semaine prochaine

## Plan préliminaire

1. Modèles ARMA(p,q) de base