**Section 3 la procédure de prévision par les modèles ARMA (Méthodologie de Box Jenkins)**

La méthodologie de Box & Jenkins vise à formuler un modèle permettant de représenter une chronique avec comme finalité de prévoir des valeurs futures. De ce fait, l’objet de cette méthodologie est de modéliser une série temporelle en fonction de ses valeurs passées et présentes afin de déterminer le processus ARIMA adéquat par principe de parcimonie.

Cette méthodologie suggère une procédure à quatre étapes : Identification du modèle, Estimation du modèle, Validation du modèle (Test de diagnostique) et prévision

**3.1 Etape d’identification du modèle**: Dans cette première étape, l’objet est de déterminer à partir de l’observation des fonctions d’autocorrélation simple et partielle dans la famille des modèles de types ARIMA (p, d, q) le modèle adéquat. Les tests informels consistent à l’analyse des moments et des plots afin de détecter la stationnarité ; mais ce ne sont que des tests de présomption de stationnarité. Puis une vérification de ces intuitions (tests informels) est faite par l’application des tests formels notamment le test de racine unitaire de Dickey Fuller. A l’issu de cette étape d’identification, on a sélectionné un ou plusieurs modèles. Il convient à présent d’estimer chaque modèle sélectionné, ce qui constitue l’objet de la deuxième étape de la procédure de Box-Jenkins

**3.2 Etape d’estimation du modèle :** Après avoir identifié les valeurs des paramètres "p" et "q" de plusieurs processus, l’étape suivante consiste à estimer les coefficients associés au terme autorégressif et moyenne mobile dans certains cas notamment dans le cas d’un processus AR(P), il est possible d’appliquer les MCO d’une façon générale on utilise la méthode de Max de vraisemblance ou MCO.

**3.3 Etape de validation du modèle**: La validation du modèle se réfère à divers tests statistiques de spécification pour vérifier si le modèle est adéquat c’est-à-dire qu’il ne peut être mis à défaut. Ces tests statistiques consistent à tester que les résidus du modèle estimer ne pas doivent être autocorrélés et ne présentent pas d’hétéroscédasticité.

Il existe aussi plusieurs types de critères pouvant être utilisé afin de comparer entre les modèles :

* **Les critères d’informations** : les critères les plus utilisés sont :
* Critère d’AKAIKE (1969) : $AIC=log\hat{σ}\_{ε}^{2}+\frac{2(p+q)}{T}$
* Critère d’information de Schwartz : $SC=log(\hat{σ}\_{ε}^{2})+\left(p+q\right)\*\frac{logT}{T}$

Plus la valeur prise par les statistiques de AIC et SC est faible plus le modèle est proche de la réalité.

* **Les critères standards** : Les critères standards sont fondés sur le calcul d’une erreur de prévision que l’on cherche à minimiser. Dans ce cadre les critères les plus utilisés sont :
* L’erreur absolue moyenne

MEAN ABSOLUTE ERROR: $MAE=\frac{1}{T}\sum\_{}^{}e\_{t}$

* La racine de l’erreur quadratique moyenne

\* MEAN ABSOLUTE PERCENT ERROR: $MAPE=100\*\frac{1}{T}\sum\_{}^{}\left|\frac{e\_{t}}{X\_{t}}\right|$

Avec T le nombre d’observations et $e\_{t}$ les résidus de l’estimation.

On retient le modèle qui minimise les différents critères.

**3.4 Etape de la Prévision** : la dernière étape de la méthodologie de Box-Jenkins est celle de la prévision. Considérons un processus ARMA (p,q), on note : $\hat{y}\_{t+h}$ la prévision faite en "t" pour la date t+h. Par définition la prévision pour un horizon h est donnée : $\hat{y}\_{t+h}=E\left({y\_{t+h}}/{I\_{t}}\right) $ avec $I\_{t} $: l’information disponible à la date "t". L’espérance conditionnelle représente la meilleure prévision de la série y conditionnellement à l’information disponible à la date t. Dans le cas linéaire, il s’agit d’une fonction de régression. La prévision dans le cas d’un processus ARMA (1,1) est donnée de la manière suivante :

$X\_{t}=ϕ\_{0}+ϕ\_{1}X\_{t-1}+θ\_{1}ε\_{t-1}+V\_{t} \left|ϕ\_{1}\right|<1$

$\hat{X}\_{t+1}=ϕ\_{0}+ϕ\_{1}\hat{X}\_{t}+θ\_{1}ε\_{t}=∝ $

$\hat{X}\_{t+2}=ϕ\_{0}+ϕ\_{1}\hat{X}\_{t+1}+θ\_{1}ε\_{t+1}$

$\hat{X}\_{t+h}=ϕ\_{0}+ϕ\_{1}\hat{X}\_{t+h-1}+θ\_{1}ε\_{t+h-1}$

Nous pouvons résumer les différentes étapes de la méthodologie de Box-Jenkins à partir du schéma suivant**:**

****

# Section 4 : Les modèles des séries temporelles non stationnaire et tests de racine unitaire

Dans la section précédente nous avons vu que la première étape de l’analyse d’une série temporelle consiste à vérifier la stationnarité de processus générateur des données. Dans cette section, nous allons étudier d’une façon plus précise ce qui est un processus non stationnaire, la non stationnarité qui le caractérise et la méthode de stionnarisation pour chaque type de processus.

**4.1 Processus TS (Trend-Stationnaire)** : ce type de processus s’écrit comme la somme d’une fonction déterministe de temps et d’une composante stochastique stationnaire d’espérance mathématique nulle. Un processus de ce type devient stationnaire par écart à une composante déterministe qu’est dans ce cas une fonction linéaire de temps. Formellement, un processus stationnaire $X\_{t}$ est dit TS et s’écrit : $X\_{t}↝ST$ s’il peut s’écrire sous la forme : $X\_{t}=f\left(t\right)+ε\_{t}$

$f:$ Est une fonction de temps et $ε\_{t}$ est un processus stochastique stationnaire. Il est évident que ce processus ne satisfait pas la définition de la stationnarité puisque son espérance dépend du temps. L’exemple le plus simple d’un processus TS est celui d’une tendance linéaire perturbée par un BB : $X\_{t}=∝+βt+ε\_{t}$ où $∝ et β$ sont deux paramètres fixes et $ε\_{t}$ est un BB. Dans ce cas le processus $X\_{t}$ est non stationnaire puisque son espérance dépend du temps. En revanche, le processus $y\_{t}$ défini par l’écart entre $X\_{t}$ et la composante déterministe et stationnaire : $y\_{t}=\left(X\_{t}-∝-β\_{t}\right)⇒y\_{t}=ε\_{t}$

Lorsqu’un processus TS est affecté par un choc aléatoire, l’effet de ce choc tend à disparaitre lorsque le temps passe, c’est la propriété de la non persistance des chocs. La non stationnarité qui caractérise le processus TS est de nature déterministe ou aléatoire.

**4.2 Processus DS (Différence Stationnary) :** c’est un processus dont la non stationnarité est au tour d’une tendance stochastique. Tandis que le processus $X\_{t}$ est caractérisé par une non stationnarité de nature aléatoire, on dit aussi que le processus admet des racines unitaires. Pour le rendre stationnaire on applique l’opération de différentiation à un ordre "d". Formellement, le processus DS s’écrit de la manière suivante :

$X\_{t}=β+X\_{t-1}+ε\_{t}⇒X\_{t}-X\_{t-1}=β+ε\_{t}$ $⇒∆X\_{t}=β+ε\_{t}$ Il est stationnaire

L’introduction de la constante $β$ permet de définir 2 processus différents :

* $β=0 $: le processus DS est sans dérive (marche au hasard). Il s’écrit $X\_{t}=X\_{t-1}+ε\_{t}$

Comme $ε\_{t}$ est un BB, ce processus DS porte le nom de marche au hasard (RANDOM-WALK-MODEL). Il est très utilisé dans la modélisation des marchés financiers.

* $β\ne 0 $: le processus porte le nom de processus DS avec dérive. Il s’écrit :$ X\_{t}=β+X\_{t-1}$

La variance d’un processus DS dépend du temps. Un processus DS est caractérisé par une non stationnarité de nature aléatoire.

**4.3 Test de DICKY-FULLER (test de non stationnarité) :** le test de DICKY-FULLER permet de mettre en évidence le caractère stationnaire ou non stationnaire d’une chronique par la détermination d’une tendance déterministe ou stochastique. Les modèles de base de la construction de ces tests sont du nombre de trois. Le principe de ce test est simple :

Si l’hypothèse $H\_{0}:ϕ=1$ est retenue dans l’un des trois modèles le processus est alors non stationnaire.

$M\left[1\right]: X\_{t}=ϕX\_{t-1}+ε\_{t}$. Modèle autorégressif d’ordre 1

$M\left[2\right]: X\_{t}=C+ϕX\_{t-1}+ε\_{t.}$ Modèle autorégressif d’ordre 1 avec constante

$M\left[3\right]: X\_{t}=C+βt+ϕX\_{t-1}+ε\_{t}.$ Modèle autorégressif d’ordre 1 avec constante et trend.

Si $H\_{0}$ est vérifiée la série $X\_{t}$ n’est pas stationnaire quel que soit le modèle retenu.

Le principe général du test est le suivant :

On estime par les moindres carrés ordinaires (MCO) le paramètre$ϕ$, noté $\hat{∅}$, pour les modèles [1], [2] et [3] ;

L’estimation des coefficients et des écarts types du modèle fourni ;

La statistique de Dickey-Fuller, notée : $t\_{\hat{∅}}=\frac{\hat{∅}}{σ\_{\hat{∅}}}$

Si l’hypothèse $H\_{0}$ est acceptée. Il existe alors une racine unitaire. Le processus n’est donc pas stationnaire.

**4.4 Test de DICKY-FULLER augmenté (ADF) :** dans les modèles précédents utilisés pour les tests de Dickey -Fuller simple le processus $ε\_{t}$ est par hypothèse un BB. Or, il n’y a aucune raison pour que l’erreur soit non corrélée. On appelle test ADF la prise en compte de cette hypothèse. Les tests ADF sont fondés par l’estimation par les MCO des trois modèles :

$M\left[4\right]: X\_{t}=ρX\_{t-1}+\sum\_{j=2}^{P}ϕ\_{j}∆X\_{t-j+1}+ε\_{t}$

$M\left[5\right]: X\_{t}=ρX\_{t-1}+\sum\_{j=2}^{P}ϕ\_{j}∆X\_{t-j+1}+C+ε\_{t}$

$M\left[6\right]: X\_{t}=ρX\_{t-1}+\sum\_{j=2}^{P}ϕ\_{j}∆X\_{t-j+1}+C+βt+ε\_{t}$

Le test se déroule de manière similaire au test de Dickey-Fuller simple seules les tables statistiques diffères. La valeur de $ρ$ peut être déterminée selon les critères AKAIKE (AIC) et SCHWARTS