

Chapitre 1 : Introduction à l'analyse classique de séries chronologiques

Dans l'analyse économique, l'étude des séries chronologiques est d'une importance cruciale en vue de ses domaines d'applications. En effet, la plupart des analyses empiriques effectuées dans les domaines de la macroéconomie et des finances sont basées sur les techniques d'analyse des séries temporelles. Ce chapitre s'intéresse à la présentation des concepts de base de l'analyse classique (traditionnelle) des séries chronologiques.

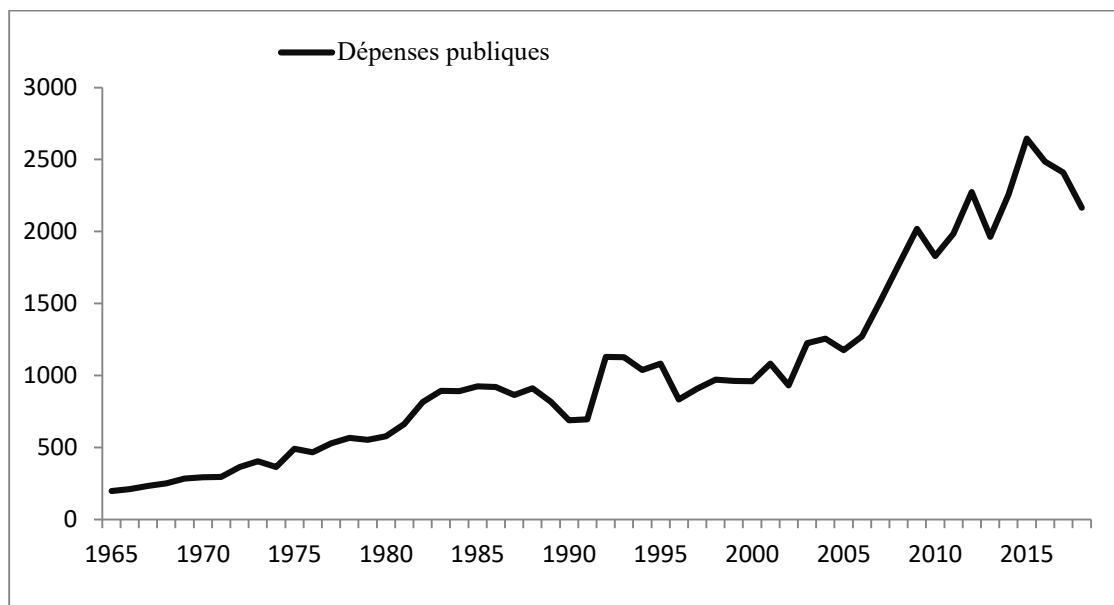
1.1 Présentation d'une série temporelle

Une série chronologique est une suite d'observations prises séquentiellement dans le temps (BOX E J and Al, 2016). L'objectif de l'étude d'une série chronologique est celui d'analyser, de décrire, d'expliquer et de prévoir l'évolution d'un phénomène dans le temps.

Une série chronologique peut prendre divers formes. Lorsque le phénomène observé évolue d'une façon continue dans le temps, la série qui en découle est qualifiée de continue. Tandis que, si le phénomène observé évolue d'une façon discontinue la série est qualifiée de discrète.

Les séries chronologiques disposent aussi de fréquences différentes (journalières, hebdomadaires, mensuelles, trimestrielles, semestrielles, annuelles). En effet, les intervalles de temps d'une série chronologique sont équidistants($t = 1, 2, 3, \dots T$).

Exemple : Evolution des dépenses publiques en Algérie (1965-2018)



Source : ONS 2020.

Il est visible à travers le graphe que les dépenses publiques ont une tendance haussière sur l'ensemble de la période observée.

1.2 Analyse des composantes d'une série temporelle

L'analyse graphique d'une série chronologique montre qu'elle se compose des éléments suivants :

- Une tendance (T) qui représente l'évolution à long terme de la variable observée.
- Une composante saisonnière (S) qui est une caractéristique périodique dont la durée est dépendante de l'environnement dans lequel le phénomène évolue.
- Une composante cyclique (C) qui représente un changement périodique qui se situe à moyen ou à long terme.
- Un résidu (E) qui représente la partie inexpliquée de l'évolution de la série par les composantes citées précédemment.

Nous distinguons deux types de modèles de séries :

Modèle additif : la série se constitue de la somme de ses composantes, $Y = T + S + C + E$

Modèle multiplicatif : la série se constitue du produit de ses composantes $Y = T * S * C * E$

En ce qui concerne la modélisation des séries temporelles, nous distinguons trois types de modèles :

1. Le modèle auto-projectif : La série chronologique dépend de ses propres valeurs passées et d'un terme aléatoire v_t . $Y_t = f(Y_{t-1}, Y_{t-2} \dots Y_{t-p}) + v_t$
2. Le modèle explicatif statique : la série dépend d'un ensemble de variables exogènes et d'un élément résiduel. $Y_t = f(X_t, v_t)$.
3. Le modèle explicatif dynamique : La série chronologique dépend des valeurs passées et présentes d'une ou plusieurs variables exogènes, comme elle peut dépendre aussi de la variable endogène décalée. $Y_t = f(X_t, X_{t-1}, \dots X_{t-q}, Y_{t-1}, \dots Y_{t-p}) + v_t$

Il existe plusieurs techniques afin d'identifier le type du modèle d'une série chronologique :

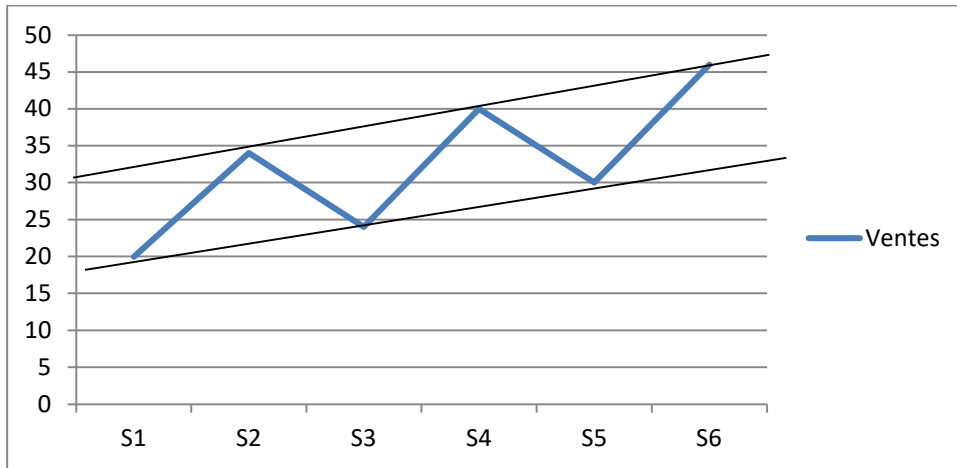
1.2.1 La méthode de la bande

Cette méthode consiste à tracer deux droites ; l'une passe par les valeurs minimales et l'autre passe par les valeurs maximales du graphe. Nous distinguons deux situations :

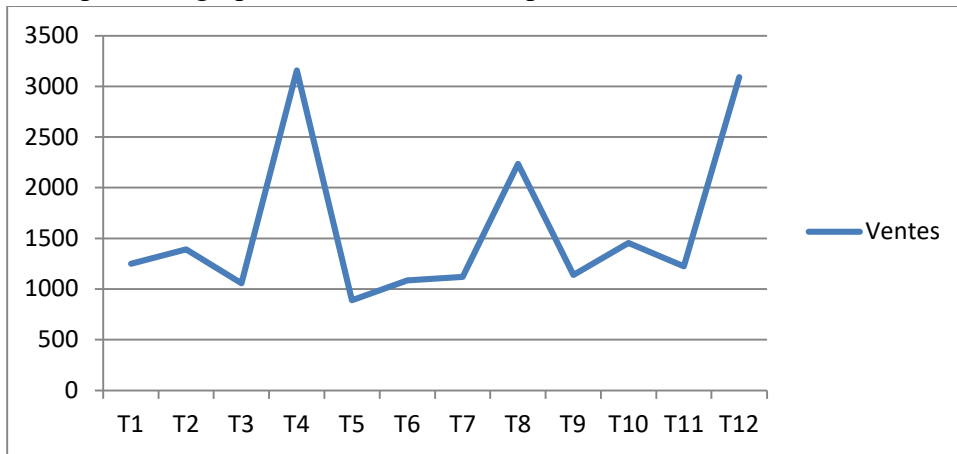
Si les deux droites tracées sont relativement parallèle (l'amplitude de variation est presque la même pour tout le graphe) le modèle est additif

Si les deux droites ne sont pas parallèles (l'amplitude de variation n'est pas la même pour le graphe) le modèle est multiplicatif.

Exemple d'un graphe pour un modèle additif



Exemple d'un graphe d'un modèle multiplicatif



1.2.2 La méthode du profil

Cette méthode consiste à superposer les saisons afin de vérifier si elles sont parallèles. Nous distinguons deux situations :

Si les droites sont parallèles le modèle est additif

Si les droites ne sont pas parallèles le modèle est multiplicatif.

1.2.3 La méthode de Buys et Ballot

Cette méthode se base sur la recherche d'une relation significative entre l'écart type et la moyenne des observations. L'estimation de cette relation est effectuée via la méthode des moindres carrés ordinaires.

Il s'agit d'estimer la relation suivante : $\delta_t = a + b\bar{X}_t$; avec δ_t l'écart type de la série et $\bar{X}_t$ la moyenne de la série (Bourbonnais R et Terraza M. 2016 PP 9-10).

Il s'agit de tester l'hypothèse suivante : $H_0: b = 0$ Le modèle est additif

$H_1: b \neq 0$ Le modèle est multiplicatif

Nous distinguons deux situations :

Si la statistique de Student calculée associée au paramètre b est inférieure à la statistique de Student de la table, nous acceptons H_0 . En effet, l'écart type est indépendant de la moyenne (la pente de la droite des moindres carrées tend vers zéro). Dans ce cas, le modèle est additif

Si la statistique de Student calculée associée au paramètre b est supérieure à la statistique de Student de la table, nous rejetons H_0 et nous acceptons H_1 . En effet, l'écart type est dépendant de la moyenne (la pente de la droite des moindres carrées est différente de zéro). Dans ce cas, le modèle est multiplicatif.

Les composantes d'une série chronologique peuvent être décrites comme suit :

$$Y = T + S + C + E$$

Code Python pour le test de Buys-Ballot

```
import numpy as np
import pandas as pd
from scipy.stats import f

def buys_ballot_test(series, period):
    """
    Test de Buys-Ballot pour la saisonnalité.
    series : liste ou pandas.Series
    period : période saisonnière (ex : 12 pour une série mensuelle)
    """

    series = np.array(series)
    n = len(series)

    # Reshape en matrice (chaque colonne = une saison)
    m = n // period
    truncated = series[:m * period]
    mat = truncated.reshape(m, period)

    # Moyennes saisonnières
    seasonal_means = mat.mean(axis=0)

    # Moyenne générale
    overall_mean = seasonal_means.mean()

    # Sommes des carrés
    SC_between = m * np.sum((seasonal_means - overall_mean) ** 2)
    SC_within = np.sum((mat - seasonal_means) ** 2)

    # Degrés de liberté
    ddl_between = period - 1
    ddl_within = (m - 1) * period

    # Statistique F
```

$$F_stat = (SC_between / ddl_between) / (SC_within / ddl_within)$$

p-value

$$p_value = 1 - f.cdf(F_stat, ddl_between, ddl_within)$$

return F_stat, p_value, seasonal_means

1.2.4 Analyse de la tendance

La méthode la plus efficace de l'analyse de la tendance est celle de l'ajustement linéaire par la méthode des moindres carrées. Sa forme générale est donnée comme suit :

$$Y = a + bX$$

La valeur du coefficient de régression b est : $b = \frac{\sum_{i=1}^n x_i y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n x_i^2 - n \bar{X}^2}$

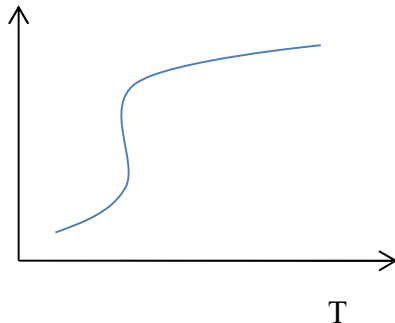
La valeur du paramètre a est : $a = \bar{Y} - b \bar{X}$

Avec : $\bar{X}$ la moyenne de la variable X.

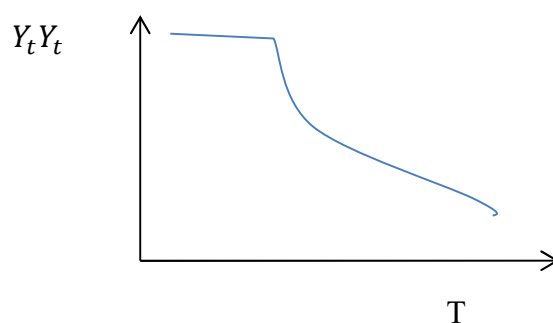
$\bar{Y}$ la moyenne de la variable Y.

n : le nombre d'observations.

Tendance haussière



Tendance baissière

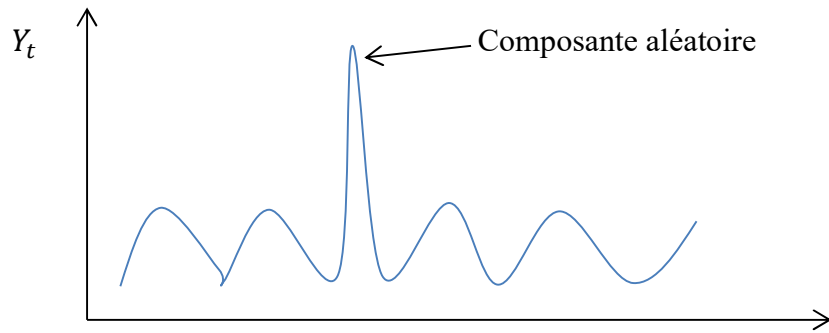


1.2.5 Analyse de la composante aléatoire

La composante aléatoire d'une série chronologique regroupe l'ensemble des paramètres qui déterminent l'évolution du phénomène Y_t hors tendance et saisonnalité.

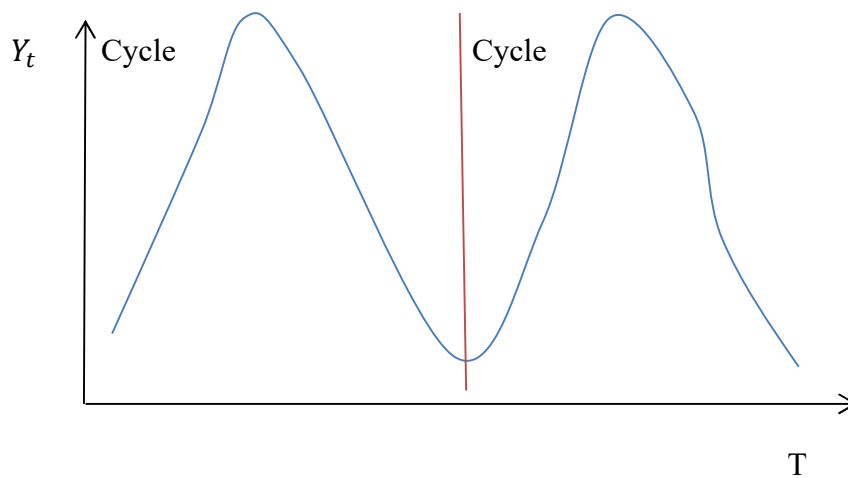
Dans le cas du modèle additif : $E_t = Y_t - (T_t + S_t)$

Dans le cas d'un modèle multiplicatif : $E_t = \frac{Y_t}{(T_t + S_t)}$



1.2.6 Analyse de la composante cyclique

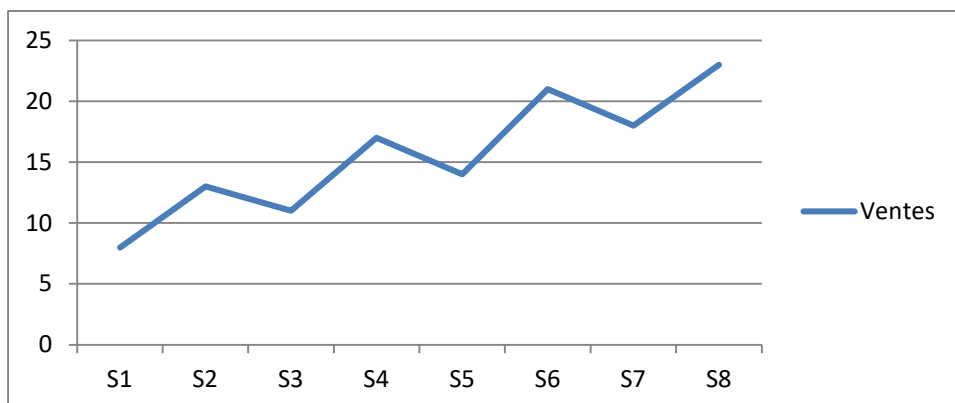
La composante cyclique signifie l'ensemble des variations régulières qui se produit dans une série temporelle. Les variations cycliques se produisent sur des périodes relativement longues.



1.2.7 Analyse de la composante saisonnière

L'analyse de la saisonnalité est primordiale avant l'étude de toute série chronologique. En effet, Il convient de désaisonnaliser la série pour pouvoir analyser les autres composantes de la série.

Exemple : le graphe suivant présente l'évolution des ventes de voitures d'un concessionnaire durant 8 semestres.



Le graphe montre que le volume de vente dans le deuxième semestre de chaque année est supérieur à celui du premier semestre, ce qui confirme l'existence d'une saisonnalité.

Remarque : Les deux composantes la tendance et la saisonnalité sont les plus apparentes lors de l'analyse des séries temporelles liées au domaine de l'analyse économique.

Code Python pour générer les composantes d'une série chronologique

```
import numpy as np
import matplotlib.pyplot as plt

np.random.seed(0)
n = 500
t = np.arange(n)

# --- Composante de tendance ---
trend = 0.01 * t # légère croissance linéaire

# --- Composante saisonnière ---
seasonal = 2 * np.sin(2 * np.pi * t / 50) # saisonnalité de période 50

# --- Bruit ---
noise = np.random.normal(0, 0.5, n)

# --- Série finale ---
series = trend + seasonal + noise

# Affichage
plt.figure(figsize=(12,8))

plt.subplot(4,1,1)
plt.plot(trend)
plt.title("Tendance")

plt.subplot(4,1,2)
plt.plot(seasonal)
plt.title("Saisonnalité")

plt.subplot(4,1,3)
plt.plot(noise)
plt.title("Bruit")

plt.subplot(4,1,4)
plt.plot(series)
plt.title("Série finale")

plt.tight_layout()
plt.show()
```

1.3 Test de Fisher d'analyse de la variance (tests de détection de la tendance et de la saisonnalité)

Ce test permet de détecter l'existence d'une tendance ou d'une saisonnalité dans une série chronologique (Bourbonnais R et Terraza M. 2016. PP 12-15).

Le test s'effectue comme suit :

D'abord, il faut calculer la variance totale de la série : $ST = \sum_{i=1}^N \sum_{j=1}^P (X_{ij} - \bar{X})^2$

Avec N : le nombre d'année.

P : la périodicité ($P = 2, P = 4, P = 12$)

X_{ij} : La valeur observée à la $i^{\text{ème}}$ année et la $j^{\text{ème}}$ période.

$\bar{X}$: la moyenne de la série. $\bar{X} = \frac{1}{N \cdot P} \sum_{i=1}^N \sum_{j=1}^P X_{ij}$

L'étape suivante consiste à calculer la somme des carrées et la variance pour les années et la période :

Somme des carrées (années) : $SA = P \sum_{i=1}^N (X_i - \bar{X})^2$

La variance (années) : $VA = \frac{SA}{N-1}$

Somme des carrées (période) : $SP = N \sum_{j=1}^P (X_j - \bar{X})^2$

La variance (période) : $VP = \frac{SP}{P-1}$

La somme des carrées résiduelle : $SR = \sum (X_{ij} - X_i - X_j + \bar{X})^2$

$$SR = ST - (SA + SP)$$

La variance résiduelle : $VR = \frac{SR}{(N-1) \cdot (P-1)}$

Tableau qui résume l'analyse de la variation

Somme des carrés	Degré de liberté	Type de variation	Valeur de variation
$SP = N \sum_{j=1}^P (X_j - \bar{X})^2$	$p - 1$	Saisonnalité	$VP = \frac{SP}{P - 1}$
$SA = P \sum_{i=1}^N (X_i - \bar{X})^2$	$N - 1$	Annuelle	$VA = \frac{SA}{N - 1}$
$SR = \sum (X_{ij} - X_i - X_j + \bar{X})^2$	$(p - 1)(N - 1)$	Résiduelle	$VR = \frac{SR}{(N - 1) * (P - 1)}$

$ST = \sum_{i=1}^N \sum_{j=1}^P (X_{ij} - \bar{X})^2$	$(Np - 1)$	Totale	$VT = \frac{ST}{Np - 1}$
---	------------	--------	--------------------------

La variance totale : $VT = VA + VP + VT$

- Test de détection de tendance :

H_0 : la série ne possède pas une tendance.

H_1 : la série possède une tendance

La statistique du test est calculée comme suit : $FC = \frac{VA}{VR} = \frac{SA/(n-1)}{SR/(p-1)(n-1)}$

La statistique de la table $F_{(V1,V2): F_{((N-1); (P-1)(N-1))}}$

La règle de décision si : $FC > F_{(V1,V2)}$ nous rejetons H_0 et nous acceptons H_1

Si $FC \leq F_{(V1,V2)}$ nous acceptons H_0

- Test de détection de saisonnalité :

H_0 : La série ne possède pas de saisonnalité

H_1 : La série possède une saisonnalité

$$FC = \frac{VP}{VR} = \frac{SP/(p-1)}{SR/(p-1)(n-1)}$$

La statistique de la table $F_{(V3,V4): F_{((P-1); (P-1)(N-1))}}$

La règle de décision si : $FC > F_{(V3,V4)}$ nous rejetons H_0 et nous acceptons H_1

Si $FC \leq F_{(V3,V4)}$ nous acceptons H_0

Code Python pour le test de décomposition de la variance

```
import numpy as np
import pandas as pd
from scipy.stats import f
from statsmodels.tsa.seasonal import seasonal_decompose

# -----
# Exemple de série chronologique synthétique
# -----
np.random.seed(0)
n = 200
t = np.arange(n)
```

```
trend = 0.02 * t
seasonal = 2 * np.sin(2*np.pi*t/20)
noise = np.random.normal(0, 0.5, n)
series = trend + seasonal + noise

# -----
# Décomposition de la série
# -----
period = 20
ts = pd.Series(series)
result = seasonal_decompose(ts, model='additive', period=period)

T = result.trend.values
S = result.seasonal.values
R = result.resid.values

# On retire les NaN de la tendance (dus au filtrage)
valid = ~np.isnan(T)
T = T[valid]
S = S[valid]
R = R[valid]
X = ts.values[valid]

n = len(X)

# -----
# Sommes de carrés
# -----
SCT = np.sum((X - np.mean(X))**2)
SC_trend = np.sum((T - np.mean(T))**2)
SC_season = np.sum((S - np.mean(S))**2)
SC_res = np.sum(R**2)

# -----
# Degrés de liberté
# -----
ddl_trend = 1          # tendance linéaire
ddl_season = period - 1 # saisons
ddl_res = n - ddl_trend - ddl_season - 1

# -----
# Statistiques F
# -----
F_trend = (SC_trend / ddl_trend) / (SC_res / ddl_res)
F_season = (SC_season / ddl_season) / (SC_res / ddl_res)

# p-values
p_trend = 1 - f.cdf(F_trend, ddl_trend, ddl_res)
p_season = 1 - f.cdf(F_season, ddl_season, ddl_res)

print("==== TEST DE TENDANCE (Fisher) =====")
```

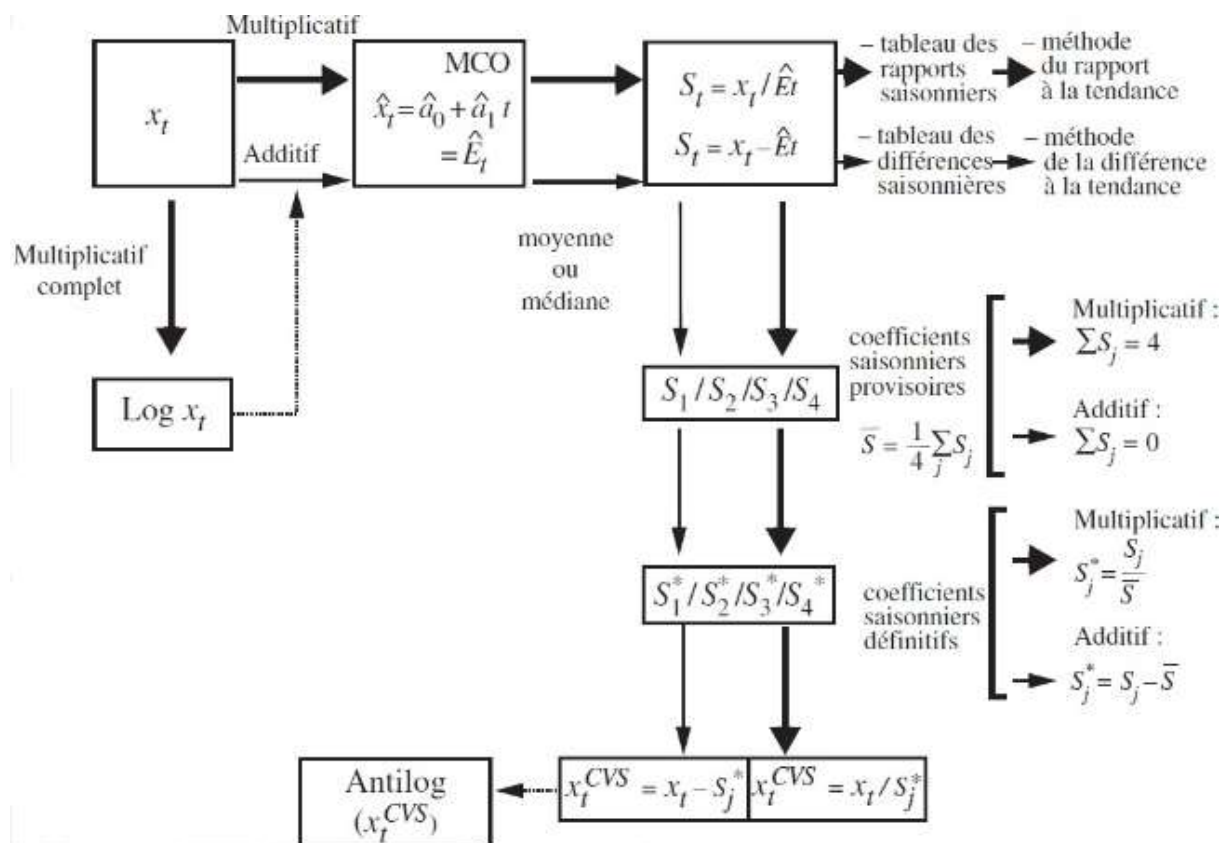
```
print(f"F = {F_trend:.4f}, p-value = {p_trend:.4f}")

print("\n==== TEST DE SAISONNALITÉ (Fisher) =====")
print(f"F = {F_season:.4f}, p-value = {p_season:.4f}")
```

1.4 La désaisonnalisation

La désaisonnalisation d'une série chronologique consiste à éliminer la composante saisonnière sans modifier les autres composantes de la série. Il s'agit donc de corriger les variations saisonnières à l'aide des coefficients saisonniers. Dans cette section, nous allons présenter la méthode de désaisonnalisation par régression dans le cas d'une saisonnalité rigide.

Le schéma ci-dessous présenté par Bourbonnais R et Terraza M résume la procédure de désaisonnalisation :



Source : Bourbonnais R et Terraza M. 2016. Analyse des séries temporelles : Applications à l'économie et à la gestion. Edition DUNOD. Paris. P 23

Le schéma présente la procédure de désaisonnalisation dans le cas du modèle additif et multiplicatif. Cette procédure consiste à éliminer la composante saisonnière de la série sans affecté les autres composantes.

Avec : t : les années, j : les saisons.

S_t : l'effet saisonnier

$S_1^* \dots S_4^*$ la moyenne des saisons

X_t^{CVC} : Série désaisonnalisée

- Dans le cas d'un modèle additif $S = \sum_{j=1}^P S_j$ doit être égale à 0. Dans ce cas, les coefficients saisonniers sont définitifs. Cependant dans le cas où $S \neq 0$ les coefficients saisonniers sont normés pour que leur somme soit égale à 0. $S_j^* = S_j - \bar{S}$.
- Dans le cas d'un modèle multiplicatif $S = \sum_{j=1}^P S_j$ doit être égale à 1. Les coefficients définitifs sont égaux à $S_j^* = S_j / \bar{S}$.

Les prévisions consistent à définir des valeurs futures à partir des observations passées.

Dans le cas d'un modèle additif $X_{t+h} = \hat{X}_{t+h} + Cs$

Dans le cas d'un modèle multiplicatif $X_{t+h} = \hat{X}_{t+h} * Cs$

Où Cs représente le coefficient saisonnier.

Code Python pour la désaisonnalisation d'une série additive

```
import numpy as np
import pandas as pd
from statsmodels.tsa.seasonal import seasonal_decompose
import matplotlib.pyplot as plt

# -----
# Exemple de série chronologique
# -----
np.random.seed(0)
n = 200
t = np.arange(n)

trend = 0.02 * t
seasonal = 2 * np.sin(2*np.pi*t/20)
noise = np.random.normal(0, 0.3, n)
series = trend + seasonal + noise

ts = pd.Series(series)

# -----
# Décomposition additive
# -----
result_add = seasonal_decompose(ts, model="additive", period=20)

# Désaisonnalisation
deseason_add = ts - result_add.seasonal
```

```
# -----  
# Affichage  
# -----  
plt.figure(figsize=(12,6))  
plt.plot(ts, label="Série originale")  
plt.plot(deseason_add, label="Série désaisonnalisée (additif)")  
plt.legend()  
plt.title("Désaisonnalisation - Modèle additif")  
plt.show()
```

Code Python de désaisonnalisation pour un modèle multiplicatif

```
# -----  
# Exemple de série multiplicative  
# -----  
trend = 1 + 0.01 * t  
seasonal = 1 + 0.4 * np.sin(2*np.pi*t/20)  
noise = np.random.normal(1, 0.05, n)  
series_mult = trend * seasonal * noise  
  
ts_mult = pd.Series(series_mult)  
  
# -----  
# Décomposition multiplicative  
# -----  
result_mult = seasonal_decompose(ts_mult, model="multiplicative", period=20)  
  
# Désaisonnalisation  
deseason_mult = ts_mult / result_mult.seasonal  
  
# -----  
# Affichage  
# -----  
plt.figure(figsize=(12,6))  
plt.plot(ts_mult, label="Série originale (multiplicative)")  
plt.plot(deseason_mult, label="Série désaisonnalisée (multiplicatif)")  
plt.legend()  
plt.title("Désaisonnalisation - Modèle multiplicatif")  
plt.show()
```

1.5 Les techniques de prévision dans les séries chronologiques

1.5.1 Le lissage exponentiel

Cette technique de prévision est utilisée lors de la présence d'une tendance aléatoire. Il s'agit d'un filtrage ou d'une transformation de la série d'origine X_t à une nouvelle série chronologique Y_t . L'objectif est celui d'établir des prévisions sur le court terme.

Le principe de base de cette méthode est celui d'attribuer la part la plus importante aux dernières observations effectuées. Ainsi il s'agit de minimiser l'expression suivante :

$$(Y_t - \hat{Y}_t)^2 + (Y_{t-1} - \hat{Y}_{t-1})^2 \alpha + (Y_{t-2} - \hat{Y}_{t-2})^2 \alpha^2 \dots$$

α est une constante, Lorsque $0 < \alpha < 1$ à chaque fois qu'on s'éloigne dans le passé, les observations ont moins de poids dans la prévision.

Il existe deux types de lissages ; lissage exponentiel simple et lissage exponentiel multiple.

1.7.2 Lissage exponentiel simple

Cette méthode de prévision est applicable lorsque la tendance est considérée comme localement constante (constante sur les segments courts de la série) (Cipra T 2020, P 76) : $Tr_t = \alpha_0$.

La prévision se base sur la formule suivante :

$$\hat{Y}_t = \alpha Y_t + (1 - \alpha) \hat{Y}_{t-1} \text{ avec } 0 \leq \alpha \leq 1$$

α représente la constante du lissage applicable à la dernière observation.

$$\text{En effet, } \hat{Y}_{t+h}(n) = \hat{Y}_n$$

En pratique il est nécessaire de définir la valeur initiale $\hat{Y}_0$ et la constante du lissage α .

La valeur initiale $\hat{Y}_0$ est définie par la moyenne arithmétique des premières observations.

La valeur de la constante du lissage α est située dans l'intervalle $0 < \alpha \leq 0,3$. Il existe trois méthodes pour choisir la valeur de cette constante dans cet intervalle (Cipra T 2020, P 77) :

La première méthode consiste à fixer sa valeur d'une façon arbitraire (0,1 ou 0,2).

La deuxième méthode consiste à calculer sa valeur avec la formule suivante :

$\alpha = \frac{1}{m+1}$ avec $m+1$ le rang de moyenne mobile qui serait adéquate avec les données de la série chronologique.

La troisième méthode consiste à effectuer des prévisions avec les valeurs 0,01. 0,02 ...0,3 et de choisir celle qui minimise la somme des résidus (celle méthode est longue en termes de calculs).

1.7.3 Le lissage exponentiel double

Cette méthode de prévision est appliquée lorsque la tendance est linéaire (linéaire sur des courts segments de temps) (Cipra T 2020, P 81). $Tr_{t-1} = \beta_0 + \beta_1(-i)$. L'estimation des paramètres de cette équation est donnée par la minimisation de l'équation suivante :

$$\sum_{i=0}^{\infty} [Y_{t-1} - (\beta_0 + \beta_1(-i))]^2 \beta^i$$

Après transformation on obtient :

$$S_t = (1 - \beta) \sum_{i=0}^{\infty} \beta^i Y_{t-i}$$

S_t représente la valeur de la série chronologique laissée à la période t en appliquant la méthode de lissage exponentiel simple.

En reprenant donc l'équation de prévision donnée par la méthode de lissage simple, on obtient :

$$S_t = \alpha Y_t + (1 - \alpha) S_{t-1}$$

Avec l'hypothèse de $\alpha = 1 - \beta$ et en introduisant le double lissage statistique, on obtient :

$$S_t^{(2)} = (1 - \beta) \sum_{i=0}^{\infty} \beta^i S_{t-i}$$

Et donc $S_t^{(2)} = \alpha S_t + (1 - \alpha) S_{t-1}^{(2)}$

De ce fait la prévision de la valeur de Y_{t-h} par la méthode de lissage double est donnée comme suit :

$$\hat{Y}_{t+h} = \beta_0(t) + \beta_1(t) * h = \left(2 + \frac{\alpha * h}{1 - \alpha}\right) S_t - \left(1 + \frac{\alpha * h}{1 - \alpha}\right) S_t^{(2)}$$

h représente l'horizon de prévision.

Afin de calculer les prévisions il est nécessaire de définir les valeurs initiales S_0 , $S_0^{(2)}$ et la constante du lissage α .

Pour calculer les valeurs initiales S_0 et $S_0^{(2)}$, il est nécessaire d'utiliser les équations de S_t et de $S_t^{(2)}$ pour la période $t = 0$. En effet,

$$S_0 = \beta_0(0) - \frac{1-\alpha}{\alpha} \beta_1(0) \quad ; \quad S_0^{(2)} = \beta_0(0) - \frac{2(1-\alpha)}{\alpha} \beta_1(0)$$

La valeur de la constante du lissage α peut être définie par la même procédure que la méthode de lissage exponentiel simple (Cipra T 2020, P 83).

1.7.4 La méthode de Holt-Winters

Cette méthode de prévision est adaptée à l'existence à la fois d'une tendance linéaire et d'une saisonnalité. Les deux versions de cette méthode pour le modèle additif et multiplicatif développent trois éléments de lissage : La constante de lissage α pour laisser le niveau L_t , le paramètre γ pour le lissage de la tendance T_t et le paramètre δ pour le lissage de l'indice saisonnier I_t . Avec $(0 < \gamma, \alpha, \delta < 1)$. (Cipra T 2020, P 91)

Pour le modèle additif les formules de la méthode de Holt-Winters sont données comme suit :

$$L_t = \alpha(Y_t - I_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1})$$

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$$

$$I_t = \delta(Y_t - L_t) + (1 - \delta)I_{t-s}$$

$$\hat{Y}_t = L_t + I_t$$

En effet, la prévision pour la période (t+h) est donnée :

$$\hat{Y}_{t+h} = L_t + T_t \cdot h + I_{t+h-s} \text{ pour } h = 1, \dots, S$$

$$\hat{Y}_{t+h} = L_t + T_t \cdot h + I_{t+h-2S} \text{ Pour } h = S + 1, \dots, 2S$$

.

Afin d'effectuer la prévision il est nécessaire de fixer les valeurs initiales des paramètres L_0, T_0 et I_{1-s} et aussi les valeurs des constantes du lissage γ, α, δ .

Les valeurs initiales des paramètres L_0, T_0 et I_{1-s} peuvent être estimées par la méthode des moindres carrés ordinaires appliquée sur l'estimation de la composante saisonnière.

Les valeurs des constantes du lissage γ, α, δ peuvent être calculées :

Soit en leur donnant des valeurs fixes recommandées ($\alpha = \delta = 0,4$ et $\gamma = 0,1$)

Soit par la réalisation des prévisions avec les valeurs 0,01, 0,02 ... 0,4 pour α et δ et de 0,01 jusqu'à 0,1 pour γ et de choisir celle qui minimise la somme des résidus (celle méthode est longue en termes de calcul).

Pour le modèle multiplicatif les formules de la méthode de Holt-Winters sont données comme suit :

$$L_t = \alpha(Y_t / I_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1})$$

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$$

$$I_t = \delta(Y_t / L_t) + (1 - \delta)I_{t-s}$$

$$\hat{Y}_t = L_t * I_t$$

En effet, la prévision pour la période (t+h) est donnée :

$$\hat{Y}_{t+h} = (L_t + T_t \cdot h) * I_{t+h-S} \text{ pour } h = 1, \dots, S$$

$$\hat{Y}_{t+h} = (L_t + T_t \cdot h) * I_{t+h-2S} \text{ Pour } h = S + 1, \dots, 2S$$

.

.

.

Pour le cas multiplicatif les valeurs initiales peuvent être trouvées par les formules suivantes :

$$T_0 = \frac{\bar{Y}_P - \bar{Y}_1}{(P-1)S} \quad ; \quad L_0 = \bar{Y}_1 - \frac{S+1}{2} T_0$$

$$I_{i-S} = \frac{1}{P} \sum_{i=0}^{P-1} \frac{Y_{j+Si}}{\bar{Y}_{i+1 - (\frac{S+1}{2} - j)} T_0} \text{ avec } j = 1 \dots S$$

$\bar{Y}_i$ représente la moyenne des observations de la $i^{\text{ème}}$ saison.

P représente le nombre total des saisons.

Les valeurs des constantes du lissage γ, α, δ sont calculées de la même façon que la méthode de Holt-Winters appliquée au modèle additif (Cipra T 2020, P 93).

Code Python pour générer des prévisions par la méthode de Holt-Winters

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.holtwinters import ExponentialSmoothing

# -----
# Exemple : série avec tendance + saisonnalité
# -----
np.random.seed(0)
n = 120
t = np.arange(n)

trend = 0.2 * t
seasonal = 5 * np.sin(2*np.pi*t/12)
noise = np.random.normal(0, 1, n)

series = trend + seasonal + noise
ts = pd.Series(series)

# -----
# Ajustement Holt-Winters
# -----
model = ExponentialSmoothing(
    ts,
```

```
trend='add',
seasonal='add',
seasonal_periods=12
)

fit = model.fit()

# Prévvision pour 24 périodes
forecast = fit.forecast(24)

# -----
# Affichage
# -----
plt.figure(figsize=(12,6))
plt.plot(ts, label="Série originale")
plt.plot(fit.fittedvalues, label="Valeurs ajustées (Holt-Winters)")
plt.plot(forecast, label="Prévisions", linestyle="--")
plt.legend()
plt.title("Méthode de Holt-Winters (trend additif, saison additif)")
plt.show()
```

Séries d'exercices n°1

Exercice n°1

Une entreprise de production de textile décide d'étudier l'évolution de la production de son nouveau produit lancé depuis 4 ans. Les données sont récapitulées dans le tableau suivant (en milliers):

	S1	S2
2015	13	18
2016	15	21
2017	18	24
2018	22	30

1. Représenter graphiquement les données.
2. Décrivez brièvement les phénomènes observés.
3. Détecter les composantes de cette série, en utilisant le test d'analyse de la variance avec un risque $\alpha = 5\%$. $F_{(3,3)}^{5\%} = 9,28$. $F_{(1,3)}^{5\%} = 10,13$.

Exercice n°2

L'évolution des ventes trimestrielles pendant les trois dernières années d'un magasin de vêtement (Milliers d'unités) est donnée comme suit :

Date	T1	T2	T3	T4
Année 1	36	39	30	93
Année 2	24	30	33	87
Année 3	33	42	36	126

L'estimation de la relation $\delta_t = a + b\bar{X}_t$ par la méthode MCO a donné les résultats suivants :

$$\hat{\delta}_i = -15,9 + 0,9\bar{X}$$

Déterminer le schéma de la série à l'aide du test de Buys-Ballot. $t_{n-2}^{\alpha/2} = 12,706$

Exercice n°3

Le tableau ci-dessous montre l'évolution de la demande sur une matière première pour une entreprise de fabrication de cosmétique. Les données représentées sont trimestrielles sur la période 1 :2018 jusqu'à 4 :2021.

	T1	T2	T3	T4
2018	235	290	611	215
2019	290	350	671	275
2020	320	411	732	336
2021	421	471	792	396

1. La série dispose-elle d'une saisonnalité ?
2. La série dispose-elle d'une tendance ?
3. Quel est le type de la série ?

Exercice n°4

Le tableau ci-dessous montre l'évolution des ventes (Tonnes) d'un atelier de recyclage pour les 4 dernières années.

Année	S (T)	Ventes (Y)
2016	1	120
	2	160
2017	3	145
	4	190
2018	5	180
	6	230
2019	7	200
	8	275

1. Tracer le graphe de la série.
2. Donner les prévisions de ventes pour 2020 et 2021.

Exercice n°5

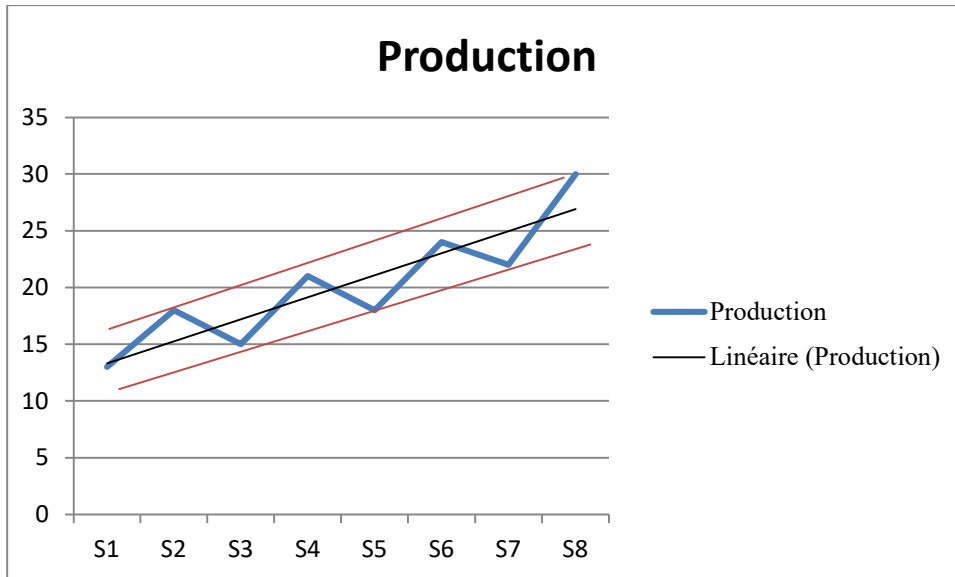
Le tableau ci-dessous représente l'évolution des quantités produites d'une unité de fabrication d'appareils électroniques.

	T1	T2	T3	T4
2020	52	36	69	89
2021	65	45	86	111
2022	81	56	108	139
2023	102	70	135	174

1. Quel est le modèle de décomposition de cette série ? $t_{n-2}^{\alpha/2} = 4,303$
2. Désaisonnaliser cette série.
3. Calculer les prévisions des quantités produites pour l'année 2024.

Solution de l'exercice n°1

1. Représentation graphique



2. Décrire brièvement les phénomènes observés

L'amplitude de variation est la même pour l'ensemble du graphe, le graphe est encadré par deux droites parallèles —> Modèle additif.

Le graphe montre une évolution croissante de la production en fonction de temps, ce qui suggère l'existence d'une tendance.

Le S2 se classe toujours en 1^{er} position et le S1 se classe toujours en seconde position —> présence de saisonnalité.

Le graphe présente une évolution haussière en fonction de temps. —> Tendance

3. Détection des composantes de la série (tendance et saisonnalité)

	S1	S2	Moyenne (A)	Ecart-type (A)
2015	13	18	15,5	3,535
2016	15	21	18	4,242
2017	18	24	21	4,242
2018	22	30	26	5,656
Moyenne (P)	17	23,25	Moyenne Générale	Ecart-type générale
Ecart-type (P)	3,915	5,123	20,125	5,383

- Test de détection de tendance :

H_0 : la série ne possède pas une tendance.

H_1 : la série possède une tendance

$$ST = \sum_{i=1}^N \sum_{j=1}^P (X_{ij} - \bar{X})^2 = SA + SP + SR$$

$$ST = \sum_{i=1}^N \sum_{j=1}^P (X_{ij} - \bar{X})^2$$

$$\begin{aligned} ST &= [(13 - 20,125)^2 + (18 - 20,125)^2 + (15 - 20,125)^2 + (21 - 20,125)^2 \\ &\quad + (18 - 20,125)^2 + (24 - 20,125)^2 + (22 - 20,125)^2 + (30 - 20,125)^2] \\ &= 202,875 \end{aligned}$$

$$SA = P \sum_{i=1}^N (X_i - \bar{X})^2$$

$$\begin{aligned} SA &= 2[(15,5 - 20,125)^2 + (18 - 20,125)^2 + (21 - 20,125)^2 + (26 - 20,125)^2] \\ &= 122,38 \end{aligned}$$

$$SP = N \sum_{j=1}^P (X_j - \bar{X})^2 = 4[(17 - 20,125)^2 + (23,25 - 20,125)^2] = 78,125$$

$$SR = \sum (X_{ij} - X_i - X_j + \bar{X})^2$$

$$SR = ST - (SA + SP) = 2,375$$

$$VA = \frac{SA}{N - 1} = \frac{122,38}{3} = 40,8$$

$$VP = \frac{SP}{P - 1} = \frac{78,12}{1} = 78,12$$

$$VR = \frac{SR}{(N - 1) * (P - 1)} = \frac{2,375}{3} = 0,792$$

$$FC = \frac{VA}{VR} = \frac{40,8}{0,792} = 51,52$$

$FC > F_{(3,3)}^{5\%} = 9,28$ Nous rejetons H_0 et nous acceptons H_1 , donc la série dispose d'une tendance

- Test de détection de saisonnalité :

H_0 : La série ne possède pas de saisonnalité

H_1 : La série possède une saisonnalité

$$FC = \frac{VP}{VR} = \frac{78,12}{0,792} = 98,64$$

$FC > F_{(1,3)}^{5\%} = 10,13$ Nous rejetons H_0 et nous acceptons H_1 , La série dispose d'une saisonnalité.

Solution de l'exercice n°2

L'estimation de l'équation de régression de l'écart type sur la moyenne par la méthode MCO a donné les résultats suivants :

Date	T1	T2	T3	T4	Moy	Ecart type
Année 1	36	39	30	93	49,5	25,32
Année 2	24	30	33	87	43,5	25,32
Année 3	33	42	36	126	59,25	38,67
					50,75	29,77

X_i	δ_i	$X_i * \delta_i$	X_i^2	δ_i estimé	Résidu	Résidu au carré
49,5	25,32	1253,34	2450,25	28,65	-3,33	11,0889
43,5	25,32	1101,42	1892,25	23,25	2,07	4,2849
59,25	38,67	2291,1975	3510,5625	37,42	1,25	1,5625
		4645,9575	7853,0625			16,9363

$$b = \frac{\sum_{i=1}^n X_i \delta_i - n \bar{X} \bar{\delta}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2} = \frac{4645,957 - 3(50,75)(29,77)}{7853,06 - 3(50,75)^2} = \frac{113,474}{126,37} = 0,9$$

$$a = \bar{\delta} - b \bar{X} = 29,77 - (0,9 * 50,75) = -15,9$$

$$\hat{\delta}_i = -15,9 + 0,9 \bar{X}$$

L'écart type de l'erreur :

$$\delta_b^2 = \frac{\sum et^2}{V(X)} = \frac{16,9363}{126,37} = 0,134 \quad \text{l'écart type } \delta_b = \sqrt{0,134} = 0,36$$

La statistique de Student associé au coefficient b est calculée comme suit :

$$T_c = \frac{|b - 0|}{\delta_b} = \frac{|0,9|}{0,36} = 2,5$$

La statistique de Student de la table $T_{n-2}^{\alpha/2} = T_1^{0,025} = 12,706$

Les hypothèses du test :

$H_0: b = 0$ Le modèle est additif

$H_1: b \neq 0$ Le modèle est multiplicatif

Nous avons $T_c < T_{n-2}^{\alpha/2}$ nous acceptons l'hypothèse $H_0 \rightarrow$ La série est générée par un modèle additif.

Solution de l'exercice n°3

1. Test de détection de saisonnalité

H_0 : La série ne possède pas de saisonnalité

H_1 : La série possède une saisonnalité

	T1	T2	T3	T4	Moyenne année X_{i*}
2018	235	290	611	215	337,75
2019	290	350	671	275	396,5
2020	320	411	732	336	449,75
2021	421	471	792	396	520
X_{*j}	316,5	380,5	701,5	305,5	$X_{ij} = 426$

$$SA = P \sum_{i=1}^N (X_i - \bar{X})^2 = 72233,5$$

$$SP = N \sum_{j=1}^N (X_j - \bar{X})^2 = 417924$$

$$SR = \sum (X_{ij} - X_i - X_j + \bar{X})^2 = 766,5$$

$$VP = \frac{SP}{P-1} = 139308$$

$$VA = \frac{SA}{N-1} = 24077,83$$

$$VR = \frac{SR}{(N-1) * (P-1)} = 85,17$$

$$FC = \frac{VP}{VR} = 1635,64$$

$FC > F_{(3,9)}^{5\%} = 3,86$ Nous rejetons H_0 et nous acceptons H_1 , La série dispose d'une saisonnalité.

2. Test de détection d'une tendance

H_0 : la série ne possède pas une tendance.

H_1 : la série possède une tendance

$$FC = \frac{VA}{VR} = 282,71$$

$FC > F_{(3,9)}^{5\%} = 3,86$ Nous rejetons H_0 et nous acceptons H_1 , donc la série dispose d'une tendance

3. Type du modèle de la série

Afin de détecter le type du modèle de la série nous utilisons le test de Buys-Ballot. Il s'agit d'estimer la relation entre l'écart type et la moyenne de la série.

$$b = \frac{\sum_{i=1}^n X_i \delta_i - n \bar{X} \bar{\delta}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2} = 0,0029$$

$$a = \bar{Y} - b \bar{X} = 160,46$$

$$\hat{\delta}_i = 160,46 + 0,0029 \bar{X}$$

La statistique de Student associé au coefficient b est calculée comme suit :

$$T_c = \frac{|b - 0|}{\delta_b} = 0,1$$

La statistique de Student de la table $T_{n-2}^{\alpha/2} = T_2^{0,025} = 4,303$

Les hypothèses du test :

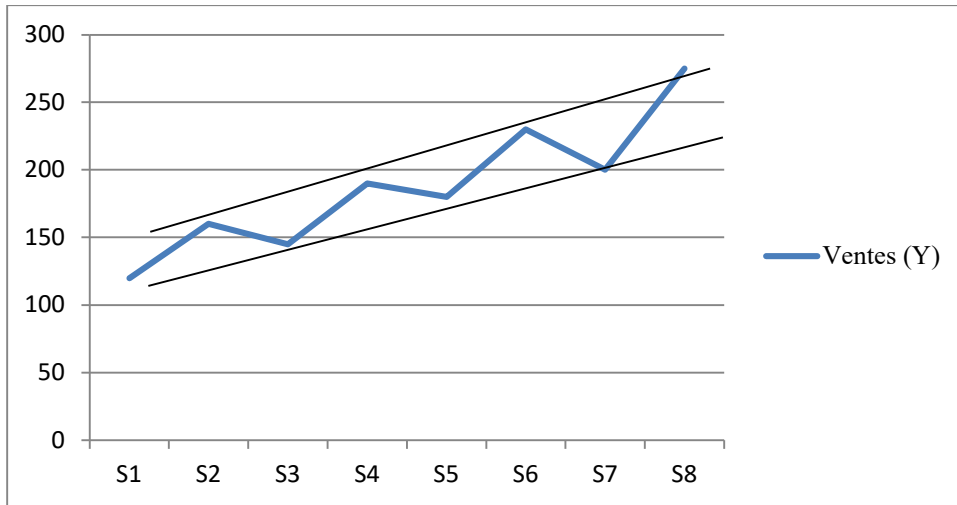
$H_0: b = 0$ Le modèle est additif

$H_1: b \neq 0$ Le modèle est multiplicatif

Nous avons $T_c < T_{n-2}^{\alpha/2}$ nous acceptons l'hypothèse $H_0 \rightarrow$ La série est générée par un modèle additif.

Solution de l'exercice n°4

1. Le graphe



La courbe est encadrée par deux droites parallèles ce que suggère que la série est générée par un modèle additif.

2. Les prévisions

S (T)	Ventes (Y)	T^2	$Y * T$	$\hat{Y}$	S_j	S_j^*	X^{CVS}
1	1,2	1	1,2	1,244	-0,044	-0,11	1,31
2	1,6	4	3,2	1,423	0,177	0,11	1,49
3	1,45	9	4,35	1,6	-0,15	-0,135	1,585
4	1,9	16	7,5	1,781	0,12	0,135	1,765
5	1,8	25	9	1,96	-0,16	-0,16	1,96
6	2,3	36	13,8	2,139	0,16	0,16	2,14
7	2	49	14	2,31	-0,31	-0,28	2,28
8	2,75	64	22	2,5	0,25	0,28	2,47

$$\beta_1 = \frac{\sum_{i=1}^n TY - n\bar{T}\bar{Y}}{\sum_{i=1}^n T^2 - n\bar{T}^2} = \frac{75,05 - 67,5}{204 - 162} = \frac{7,55}{42} = 0,18$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{T} = (1,875) - (0,18 * 4,5) = 1,065$$

$$Y = 1,065 + 0,179 T$$

Dans le cas du modèle additif : $S_j = Y - \hat{Y} \rightarrow$ méthode de la différence de la tendance :

$$\text{Si } \sum_{i=1}^P S_j = 0 \rightarrow S_j^* = S_j$$

$$\text{Si } \sum_{i=1}^P S_j \neq 0 \rightarrow S_j^* = S_j - \bar{S}$$

Dans le cas du modèle multiplicatif $S_j = \frac{Y}{\hat{Y}} \rightarrow$ méthode des rapports à la tendance

Si $\sum_{i=1}^P S_j = P \rightarrow S_j^* = S_j$ avec $\bar{S} = \frac{1}{N} \sum S_j$

Si $\sum_{i=1}^P S_j \neq 0 \rightarrow S_j^* = S_j / \bar{S}$ avec $\bar{S} = \sqrt{\prod_{j=1}^P S_j}$

S_j représente les coefficients saisonniers

S_j^* représente les coefficients saisonniers corrigés.

Prévision dans le cas additif :

$$\hat{Y}_{t+n}^p = \hat{Y}_{t+n} + CS$$

$$CS_{S1} = \frac{-0,11 - 0,135 - 0,16 - 0,28}{4} = -0,171$$

$$CS_{S2} = \frac{0,11 + 0,135 + 0,16 + 0,28}{4} = 0,171$$

	T	H	$\hat{Y}_{t+n}$	CS	$\hat{Y}_{t+n}^p$
2020	9	1	2,676	-0,171	2,505
	10	2	2,855	0,171	3,026
2021	11	3	3,034	-0,171	2,863
	12	4	3,213	0,171	3,384

Solution de l'exercice n°5

1. Le modèle de décomposition de la série

On applique le test de Buys-Ballot

On calcul la moyenne et l'écart type

	T1	T2	T3	T4	$\bar{X}$	δ
2020	52	36	69	89	61,5	19,7
2021	65	45	86	111	76,75	24,51
2022	81	56	108	139	96	30,89
2023	102	70	135	174	120,25	38,61

$$\bar{t} = \frac{61,5 + 76,75 + 96 + 120,25}{4} = 88,625$$

$$\bar{\delta} = \frac{19,7 + 24,51 + 30,89 + 38,61}{4} = 28,43$$

Puis on test s'il y a une relation significative entre la moyenne et l'écart type :

On estime le modèle suivant :

$$b = \frac{\sum_{i=1}^n X_i \delta_i - n \bar{X} \bar{\delta}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2}$$

	σ_i	$\bar{X}_i$	$\sigma_i \bar{X}$	$\bar{X}_i^2$	$\hat{\sigma}_i$	ϵ_i	e_i^2
2013	19.70	61.5	2061.48	3782.25	19.5	0.2	0.04
2014	24.51	76.75	1881.14	5890.56	24.83	0.13	0.01
2015	30.89	96	2965.44	9216	30.54	0.35	0.12
2016	38.61	120.25	4642.85	14460.06	38.4	0.31	0.09
Somme	113.71	354.5	10700.98	33348.87		0.99	0.27

$$b = \frac{10700,98 - 4(28,43)(88,625)}{33348,87 - 4(88,625)^2} = 0,32$$

$$a = \bar{\delta} - b\bar{X} = 28,43 - (0,32 * 88,625) = -0,18$$

$$\hat{\delta}_i = -0,18 + 0,32\bar{X}$$

L'écart type de l'erreur :

$$\delta_b^2 = \frac{\sum e_i^2}{V(X)} = \frac{0,27/2}{33348,87 - (88,625)^2} = 0,000069 \quad \text{l'écart type } \delta_b = \sqrt{0,000069} = 0,0083$$

La statistique de Student associé au coefficient b est calculée comme suit :

$$T_c = \frac{|b - 0|}{\delta_b} = \frac{|0,32|}{0,0083} = 38,55$$

La statistique de Student de la table $T_{n-2}^{\alpha/2} = 4,303$

Les hypothèses du test :

$H_0: b = 0$ Le modèle est additif

$H_1: b \neq 0$ Le modèle est multiplicatif

Nous avons $T_c > T_{n-2}^{\alpha/2}$ nous rejetons l'hypothèse H_0 et nous acceptons $H_1 \rightarrow$ La série est générée par un modèle multiplicatif.

2. Désaisonnalisation de la série

	X_i	t_i	$X_i t_i$	t_i^2	$\hat{x} = \hat{\alpha} + \hat{\beta} t_i$	$S_i = X_i / \hat{x}$		$S^* = S_i / \bar{S}$	$X_{cvs} = X_i / S^*$
2013	52	1	52	1	43.2	1.2037	1.09	1.10	47.27
	36	2	72	4	50.08	0.7188		0.65	55.38
	69	3	207	9	56.96	1.2113		1.11	62.16
	89	4	356	16	63.84	1.3941		1.27	70.07
2014	65	5	325	25	70.72	0.9191	0.90	1.02	63.72
	45	6	270	36	77.6	0.5798		0.64	70.31
	86	7	602	49	84.48	1.0179		1.13	76.10
	111	8	888	64	91.36	1.2149		1.11	100
2015	81	9	729	81	98.24	0.8245	0.93	0.88	92.04
	86	10	860	100	105.12	0.8181		0.87	98.85
	108	11	1188	121	112	0.9642		1.03	104.85
	139	12	1668	144	118.88	1.1692		1.25	111.2
2016	102	13	1326	169	125.76	0.8110	1.005	0.80	127.5
	70	14	980	196	63.84	1.0964		1.09	64.22
	135	15	2025	225	139.52	0.9676		0.96	140.62
	174	16	2784	256	146.4	1.1885		1.18	147.45

Pour désaisonnaliser la série, on estime le modèle suivant :

$$X_i = \alpha + \beta t_i + \varepsilon_i$$

$$b = \frac{\sum_{i=1}^n X_i t_i - N \bar{X} \bar{t}}{\sum_{i=1}^n t_i^2 - n \bar{t}^2} = \frac{13632 - 16(90,5)(7,87)}{1316 - 16(7,87)^2} = 6,88$$

$$a = 90,5 - (6,88 * 7,87) = 36,32$$

$$\hat{x}_i = 36,32 + 6,88 \bar{X}$$

X_{cvs} est la série corrigée des variations saisonnières.

3. Les prévisions

Ainsi, le premier coefficient saisonnier est égal à la racine carrée du produit des premiers trimestres S^* de chaque année.

$$C_{S1} = \sqrt{(1,1) * (1,02) * (0,88) * (0,8)} = 0,75$$

$$C_{S2} = \sqrt{(0,65) * (0,64) * (0,87) * (1,09)} = 0,48$$

$$C_{S3} = \sqrt{(1,11) * (1,11) * (1,03) * (0,96)} = 1,15$$

$$C_{S4} = \sqrt{(1,27) * (1,11) * (1,25) * (1,18)} = 1,84$$

Les prévisions sont données : $X_{t+h} = \hat{X}_{t+h} * C_{Si}$

$$X_{17} = 36,32 + 6,88 (17) = 153,28$$

$$X_{17} = 153,28 * 0,75 = 114,96$$

$$X_{18} = 36,32 + 6,88 (18) = 160,16$$

$$X_{18} = 160,16 * 0,48 = 76,87$$

$$X_{19} = 36,32 + 6,88 (19) = 167,04$$

$$X_{19} = 167,04 * 1,15 = 192,096$$

$$X_{20} = 36,32 + 6,88 (20) = 173,92$$

$$X_{20} = 173,92 * 1,84 = 320,01$$