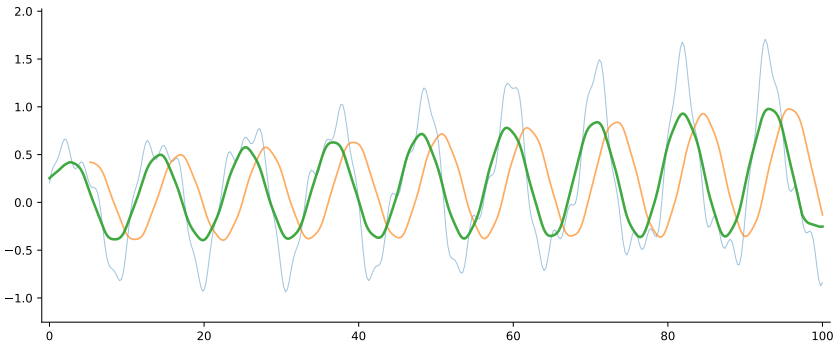
Séries temporelles

Cours 1 – Observer, décomposer, stationnariser



M1 Actuariat – Université Gustave Eiffel

Solesne Bourguin – 2026–2027

Fil conducteur de la séance

Aujourd'hui, on part d'une question très simple : devant une série temporelle, que faut-il regarder avant de choisir un modèle ?

Fil conducteur de la séance

Aujourd'hui, on part d'une question très simple : devant une série temporelle, que faut-il regarder avant de choisir un modèle ?

1. Observer la série : chronogramme, tendance, saisonnalité, variabilité.
2. Séparer ce qui est lent, régulier, et irrégulier : décomposition et résidus.
3. Comparer une observation à son passé : lag plot et opérateur retard.
4. Transformer la série : différences ordinaires et saisonnières.
5. Introduire la stationnarité : le cadre probabiliste pour la suite du cours.

Fil conducteur de la séance

Aujourd'hui, on part d'une question très simple : devant une série temporelle, que faut-il regarder avant de choisir un modèle ?

1. Observer la série : chronogramme, tendance, saisonnalité, variabilité.
2. Séparer ce qui est lent, régulier, et irrégulier : décomposition et résidus.
3. Comparer une observation à son passé : lag plot et opérateur retard.
4. Transformer la série : différences ordinaires et saisonnières.
5. Introduire la stationnarité : le cadre probabiliste pour la suite du cours.

Le but n'est pas de tout modéliser aujourd'hui. Le but est d'apprendre une méthode de lecture qui servira dans tout le cours.

Organisation générale du cours

Le poly contient beaucoup plus que ce qui sera écrit au tableau. En cours, on va alterner :

Support	**Rôle**	**Exemples**
Poly	référence complète	définitions, preuves, détails
Tableau	construction lente	calculs, démonstrations, schémas
Projection	fil conducteur	figures, questions, résultats clés
R / TD–TP	pratique	graphiques, diagnostics, modèles

Organisation générale du cours

Le poly contient beaucoup plus que ce qui sera écrit au tableau. En cours, on va alterner :

Support	**Rôle**	**Exemples**
Poly	référence complète	définitions, preuves, détails
Tableau	construction lente	calculs, démonstrations, schémas
Projection	fil conducteur	figures, questions, résultats clés
R / TD–TP	pratique	graphiques, diagnostics, modèles

Les slides ne remplacent pas le poly : elles servent à guider le cours et à révéler progressivement ce qu'on peut conclure des exemples.

Rythme prévu sur les 9 séances de cours

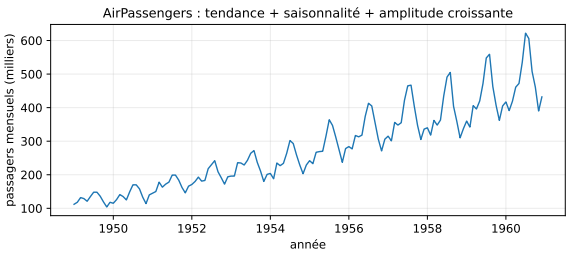
Séance	Contenu visé
1	Chapitre 1 : séries observées, décomposition, retards/différences, stationnarité, bruit blanc.
2	Chapitre 2 et début chapitre 3 : objets R, graphiques, filtrage fini.
3	Fin chapitre 3 : filtres infinis, convolution, bruit blanc filtré, séries linéaires.
4	Chapitre 4 : MA, AR, ARMA, causalité, inversibilité.
5	Fin chapitre 4 : ACF/PACF, identification, simulations.
6	Chapitre 5 : non-stationnarité, différences, ARIMA/SARIMA.
7	Chapitre 6 : estimation, choix de modèle, diagnostic, prévision.
8	Chapitre 7 : lissage exponentiel, Holt, Holt–Winters.
9	Chapitre 8 + synthèse : spectral léger, volatilité conditionnelle, méthode complète.

Rythme prévu sur les 9 séances de cours

Séance	Contenu visé
1	Chapitre 1 : séries observées, décomposition, retards/différences, stationnarité, bruit blanc.
2	Chapitre 2 et début chapitre 3 : objets R, graphiques, filtrage fini.
3	Fin chapitre 3 : filtres infinis, convolution, bruit blanc filtré, séries linéaires.
4	Chapitre 4 : MA, AR, ARMA, causalité, inversibilité.
5	Fin chapitre 4 : ACF/PACF, identification, simulations.
6	Chapitre 5 : non-stationnarité, différences, ARIMA/SARIMA.
7	Chapitre 6 : estimation, choix de modèle, diagnostic, prévision.
8	Chapitre 7 : lissage exponentiel, Holt, Holt–Winters.
9	Chapitre 8 + synthèse : spectral léger, volatilité conditionnelle, méthode complète.

Ce premier cours installe le vocabulaire. Les calculs de modèle commencent vraiment dans les chapitres suivants.

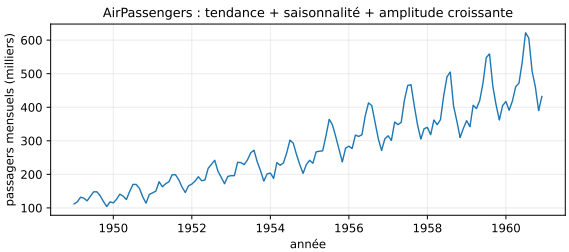
Le problème de départ : une observation dépend du temps



Avant de parler de modèle, on regarde la série.

Que voyez-vous sur ce graphique ?

Le problème de départ : une observation dépend du temps

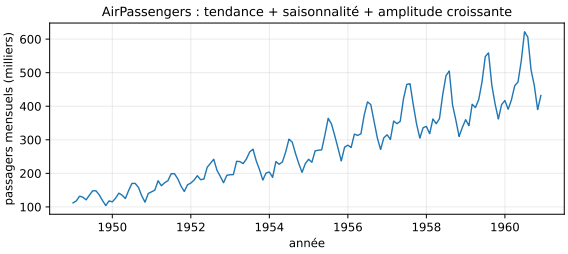


Avant de parler de modèle, on regarde la série.

Ce que l'on doit repérer :

- les observations sont ordonnées dans le temps ;
- le niveau moyen augmente ;
- un motif annuel se répète ;
- l'amplitude du motif augmente.

Le problème de départ : une observation dépend du temps



Avant de parler de modèle, on regarde la série.

Ce que l'on doit repérer :

- les observations sont ordonnées dans le temps ;
- le niveau moyen augmente ;
- un motif annuel se répète ;
- l'amplitude du motif augmente.

Une série temporelle n'est pas seulement une liste de nombres : l'ordre des observations fait partie de l'information.

Pourquoi un échantillon i.i.d. ne suffit plus ?

Dans un échantillon i.i.d., permuter les observations ne change pas l'objet statistique. Pour une série temporelle, cela détruit une partie de l'information.

Pourquoi un échantillon i.i.d. ne suffit plus ?

Dans un échantillon i.i.d., permuter les observations ne change pas l'objet statistique. Pour une série temporelle, cela détruit une partie de l'information.

On note

$$x_1, \dots, x_n \quad \text{ou} \quad (x_t)_{t \in T}$$

les **données observées**. Ici, x_t est un nombre : la valeur mesurée à la date t.

On note

$$X = (X_t)_{t \in T}$$

un **processus stochastique**. Ici, X_t est une variable aléatoire : le modèle théorique avant observation.

Pourquoi un échantillon i.i.d. ne suffit plus ?

Dans un échantillon i.i.d., permuter les observations ne change pas l'objet statistique. Pour une série temporelle, cela détruit une partie de l'information.

On note

$$x_1, \dots, x_n \quad \text{ou} \quad (x_t)_{t \in T}$$

les **données observées**. Ici, x_t est un nombre : la valeur mesurée à la date t.

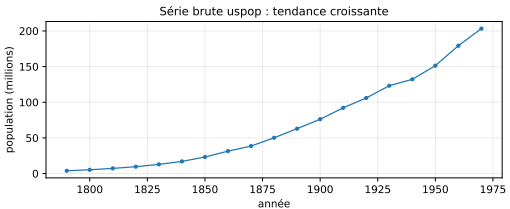
On note

$$X = (X_t)_{t \in T}$$

un **processus stochastique**. Ici, X_t est une variable aléatoire : le modèle théorique avant observation.

La distinction sera constante : x_t désigne une valeur observée ; X_t désigne la variable aléatoire que l'on utilise pour modéliser cette valeur.

Le chronogramme : premier diagnostic

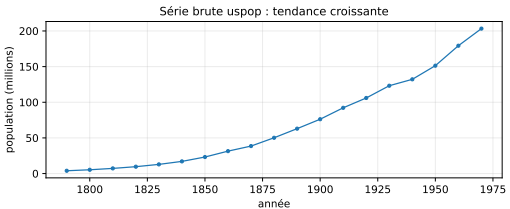


Un **chronogramme** trace les observations dans l'ordre du temps.

Il ne donne pas encore un modèle, mais il permet de formuler les premières questions :

- le niveau moyen semble-t-il constant ?
- voit-on une tendance ?
- voit-on un motif qui revient ?
- la variabilité change-t-elle avec le niveau ?

Le chronogramme : premier diagnostic



Un **chronogramme** trace les observations dans l'ordre du temps.

Il ne donne pas encore un modèle, mais il permet de formuler les premières questions :

- le niveau moyen semble-t-il constant ?
- voit-on une tendance ?
- voit-on un motif qui revient ?
- la variabilité change-t-elle avec le niveau ?

Sur cette série, le point dominant est une croissance de long terme. Une méthode de série temporelle devra d'abord traiter cette évolution du niveau.

Pourquoi décomposer une série ?

Sur une série brute, plusieurs phénomènes sont superposés. Les mélanger dans une seule quantité empêche souvent de comprendre la dynamique.

Pourquoi décomposer une série ?

Sur une série brute, plusieurs phénomènes sont superposés. Les mélanger dans une seule quantité empêche souvent de comprendre la dynamique.

Une hausse lente du niveau n'a pas la même signification qu'un pic exceptionnel ou qu'un effet saisonnier régulier.

Pour prévoir, il est souvent plus simple de prévoir séparément un niveau, un effet saisonnier, puis un reste.

Pour diagnostiquer un modèle, on regarde ce qu'il n'explique pas encore : les résidus.

Pour appliquer un modèle stationnaire, il faut souvent enlever d'abord tendance et saisonnalité.

Pourquoi décomposer une série ?

Sur une série brute, plusieurs phénomènes sont superposés. Les mélanger dans une seule quantité empêche souvent de comprendre la dynamique.

Une hausse lente du niveau n'a pas la même signification qu'un pic exceptionnel ou qu'un effet saisonnier régulier.

Pour prévoir, il est souvent plus simple de prévoir séparément un niveau, un effet saisonnier, puis un reste.

Pour diagnostiquer un modèle, on regarde ce qu'il n'explique pas encore : les résidus.

Pour appliquer un modèle stationnaire, il faut souvent enlever d'abord tendance et saisonnalité.

Décomposer, c'est transformer une série confuse en trois questions plus simples : évolution lente, motif régulier, reste à modéliser.

Un modèle exploratoire : tendance + saisonnalité + reste

On part d'une série observée (x_t) et on imagine un processus (X_t) dont elle est une réalisation. Pour séparer les phénomènes visibles, on utilise souvent l'écriture

$$X_t = m_t + s_t + Z_t.$$

Un modèle exploratoire : tendance + saisonnalité + reste

On part d'une série observée (x_t) et on imagine un processus (X_t) dont elle est une réalisation. Pour séparer les phénomènes visibles, on utilise souvent l'écriture

$$X_t = m_t + s_t + Z_t.$$

X_t : variable aléatoire qui modélise la valeur à la date t.
m_t : **tendance**, évolution lente du niveau moyen.

s_t : **saisonnalité**, motif régulier qui se répète avec une période connue.

Z_t : **bruit** ou reste théorique, partie non expliquée par tendance et saisonnalité.

Un modèle exploratoire : tendance + saisonnalité + reste

On part d'une série observée (x_t) et on imagine un processus (X_t) dont elle est une réalisation. Pour séparer les phénomènes visibles, on utilise souvent l'écriture

$$X_t = m_t + s_t + Z_t.$$

X_t : variable aléatoire qui modélise la valeur à la date t.
m_t : **tendance**, évolution lente du niveau moyen.

s_t : **saisonnalité**, motif régulier qui se répète avec une période connue.

Z_t : **bruit** ou reste théorique, partie non expliquée par tendance et saisonnalité.

Cette écriture décrit un modèle idéal. Dans les données, on ne voit pas séparément m_t, s_t et Z_t : il faut les reconstruire ou les estimer.

Les objets que l'on va manipuler

Dans cette séance, les notations suivantes seront utilisées de façon stable.

Notation	Statut	Interprétation
x_t	observé	valeur réellement mesurée à la date t ; c'est une donnée.
X_t	modèle	variable aléatoire qui modélise la valeur à la date t.
y_t	observé / transformé	série transformée, par exemple $y_t = \log(x_t)$ quand $x_t > 0$.
$\widehat{m}_t$	estimé	tendance reconstruite à partir des données.
$\widehat{s}_t$	estimé	effet saisonnier reconstruit à partir des données.
$\widehat{x}_t$ ou $\widehat{y}_t$	estimé	valeur ajustée, c'est-à-dire valeur reconstruite par la méthode choisie.
$\widehat{z}_t$	calculé	résidu empirique : observation moins valeur ajustée.

Les objets que l'on va manipuler

Dans cette séance, les notations suivantes seront utilisées de façon stable.

Notation	Statut	Interprétation
x_t	observé	valeur réellement mesurée à la date t ; c'est une donnée.
X_t	modèle	variable aléatoire qui modélise la valeur à la date t.
y_t	observé / transformé	série transformée, par exemple $y_t = \log(x_t)$ quand $x_t > 0$.
$\widehat{m}_t$	estimé	tendance reconstruite à partir des données.
$\widehat{s}_t$	estimé	effet saisonnier reconstruit à partir des données.
$\widehat{x}_t$ ou $\widehat{y}_t$	estimé	valeur ajustée, c'est-à-dire valeur reconstruite par la méthode choisie.
$\widehat{z}_t$	calculé	résidu empirique : observation moins valeur ajustée.

Le chapeau signifie : « obtenu à partir des données ». Le résidu empirique n'est donc pas observé directement ; il apparaît seulement après un ajustement.

De la décomposition théorique à l'ajustement

Modèle idéal
On écrit formellement

$$X_t = m_t + s_t + Z_t.$$

Cette ligne dit quels types de phénomènes on veut séparer.

Données
On part seulement de

$$x_1, \ldots, x_n.$$

La méthode choisie produit $\widehat{m}_t$ et $\widehat{s}_t$.

De la décomposition théorique à l'ajustement

Modèle idéal
On écrit formellement

$$X_t = m_t + s_t + Z_t.$$

Cette ligne dit quels types de phénomènes on veut séparer.

Données
On part seulement de

$$x_1, \dots, x_n.$$

La méthode choisie produit $\widehat{m}_t$ et $\widehat{s}_t$.

Une fois l'ajustement fait, on pose généralement

$$\widehat{x}_t = \widehat{m}_t + \widehat{s}_t, \qquad \widehat{z}_t = x_t - \widehat{x}_t.$$

De la décomposition théorique à l'ajustement

Modèle idéal
On écrit formellement

$$X_t = m_t + s_t + Z_t.$$

Cette ligne dit quels types de phénomènes on veut séparer.

Données
On part seulement de

$$x_1, \dots, x_n.$$

La méthode choisie produit $\widehat{m}_t$ et $\widehat{s}_t$.

Une fois l'ajustement fait, on pose généralement

$$\widehat{x}_t = \widehat{m}_t + \widehat{s}_t, \qquad \widehat{z}_t = x_t - \widehat{x}_t.$$

La question pratique est donc : quelle méthode utiliser pour obtenir $\widehat{m}_t$ et $\widehat{s}_t$?

Deux façons d'obtenir les composantes ajustées

On a besoin d'une règle qui transforme la série observée en composantes estimées.

Deux façons d'obtenir les composantes ajustées

On a besoin d'une règle qui transforme la série observée en composantes estimées.

Décomposition exploratoire
Un algorithme, comme `stl()` dans R, extrait directement une tendance, une saisonnalité et un reste. On l'utilise pour comprendre la série.

Modèle paramétrique
On choisit une forme explicite, par exemple une droite pour la tendance et un effet par mois pour la saisonnalité, puis on estime les paramètres.

Deux façons d'obtenir les composantes ajustées

On a besoin d'une règle qui transforme la série observée en composantes estimées.

Décomposition exploratoire
Un algorithme, comme `stl()` dans R, extrait directement une tendance, une saisonnalité et un reste. On l'utilise pour comprendre la série.

Modèle paramétrique
On choisit une forme explicite, par exemple une droite pour la tendance et un effet par mois pour la saisonnalité, puis on estime les paramètres.

Dans les deux cas, les composantes obtenues dépendent de la méthode. Il ne faut pas croire qu'une tendance estimée est une grandeur donnée par la nature.

Une méthode exploratoire : `stl()`

`stl()` signifie *Seasonal-Trend decomposition using Loess*. On l'utilise ici comme une boîte à outils pratique pour séparer une série saisonnière en trois morceaux.

Sur une série transformée y_t, par exemple $y_t = \log(x_t)$, l'appel

```
stl(y, s.window = "periodic")
```

renvoie trois sorties appelées `trend`, `seasonal` et `remainder`.

Une méthode exploratoire : `stl()`

`stl()` signifie *Seasonal-Trend decomposition using Loess*. On l'utilise ici comme une boîte à outils pratique pour séparer une série saisonnière en trois morceaux.

Sur une série transformée y_t, par exemple $y_t = \log(x_t)$, l'appel

```
stl(y, s.window = "periodic")
```

renvoie trois sorties appelées `trend`, `seasonal` et `remainder`.

On les identifie respectivement à la tendance ajustée, à la saisonnalité ajustée et au résidu empirique de la série y_t :

$$y_t = \widehat{m}_t + \widehat{s}_t + \widehat{z}_t.$$

Une méthode exploratoire : `stl()`

`stl()` signifie *Seasonal-Trend decomposition using Loess*. On l'utilise ici comme une boîte à outils pratique pour séparer une série saisonnière en trois morceaux.

Sur une série transformée y_t, par exemple $y_t = \log(x_t)$, l'appel

```
stl(y, s.window = "periodic")
```

renvoie trois sorties appelées `trend`, `seasonal` et `remainder`.

On les identifie respectivement à la tendance ajustée, à la saisonnalité ajustée et au résidu empirique de la série y_t :

$$y_t = \widehat{m}_t + \widehat{s}_t + \widehat{z}_t.$$

L'idée de `stl()` est de lisser localement la tendance, puis d'estimer un motif saisonnier, et de recommencer jusqu'à obtenir une séparation stable.

Une méthode explicite : tendance linéaire

Pour comprendre ce que signifie « ajuster », on peut choisir une tendance très simple :

$$m_t = a_0 + a_1 t,$$

où a_0 et a_1 sont deux nombres inconnus.

Une méthode explicite : tendance linéaire

Pour comprendre ce que signifie « ajuster », on peut choisir une tendance très simple :

$$m_t = a_0 + a_1 t,$$

où a_0 et a_1 sont deux nombres inconnus.

Les moindres carrés choisissent $\widehat{a}_0$ et $\widehat{a}_1$ de manière à minimiser

$$\sum_{t=1}^{n} \big(x_t - a_0 - a_1 t\big)^2.$$

La tendance ajustée obtenue est

$$\widehat{m}_t = \widehat{a}_0 + \widehat{a}_1 t.$$

Une méthode explicite : tendance linéaire

Pour comprendre ce que signifie « ajuster », on peut choisir une tendance très simple :

$$m_t = a_0 + a_1 t,$$

où a_0 et a_1 sont deux nombres inconnus.

Les moindres carrés choisissent $\widehat{a}_0$ et $\widehat{a}_1$ de manière à minimiser

$$\sum_{t=1}^{n} \big(x_t - a_0 - a_1 t\big)^2.$$

La tendance ajustée obtenue est

$$\widehat{m}_t = \widehat{a}_0 + \widehat{a}_1 t.$$

Ici, $\widehat{m}_t$ est simplement la valeur donnée par la droite qui se rapproche le plus des données au sens des moindres carrés.

Une méthode explicite : saisonnalité mensuelle

Pour des données mensuelles, on peut donner un effet propre à chaque mois.

Une écriture possible est

$$s_t = \alpha_1 \mathbf{1}_{\{t \text{ en janvier}\}} + \cdots + \alpha_{12} \mathbf{1}_{\{t \text{ en décembre}\}}.$$

Les coefficients $\alpha_1, \ldots, \alpha_{12}$ mesurent les effets mensuels.

Une méthode explicite : saisonnalité mensuelle

Pour des données mensuelles, on peut donner un effet propre à chaque mois.

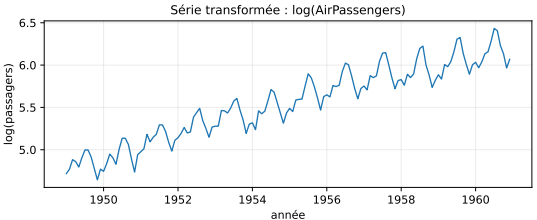
Une écriture possible est

$$s_t = \alpha_1 \mathbf{1}_{\{t \text{ en janvier}\}} + \cdots + \alpha_{12} \mathbf{1}_{\{t \text{ en décembre}\}}.$$

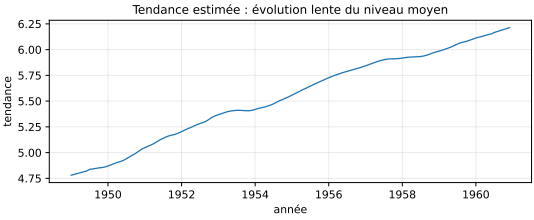
Les coefficients $\alpha_1, \ldots, \alpha_{12}$ mesurent les effets mensuels.

Un mois systématiquement haut aura un effet saisonnier estimé positif ; un mois systématiquement bas aura un effet estimé négatif. Cette idée sera utile pour interpréter les décompositions et les modèles saisonniers.

Décomposer, concrètement : que montrent les deux premières sorties ?

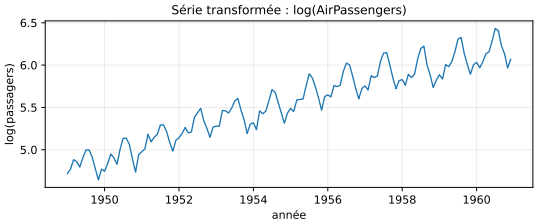


série transformée $y_t = \log(x_t)$

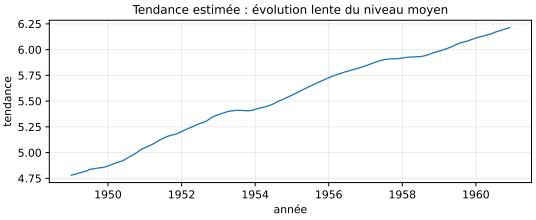


tendance ajustée $\widehat{m}_t$

Décomposer, concrètement : que montrent les deux premières sorties ?



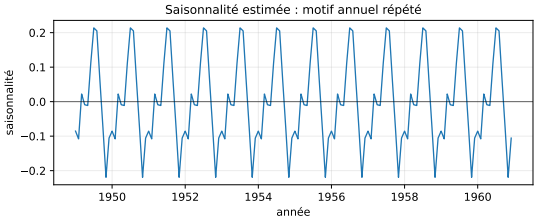
série transformée $y_t = \log(x_t)$



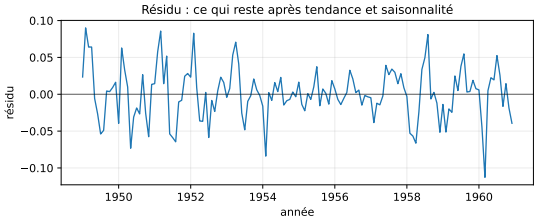
tendance ajustée $\widehat{m}_t$

À gauche, tout est encore mélangé. À droite, la tendance suit le mouvement lent du niveau moyen sans chercher à reproduire chaque oscillation mensuelle.

Décomposer, concrètement : saisonnalité et résidu

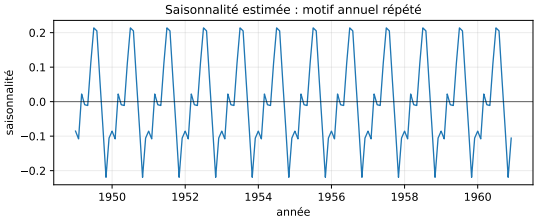


saisonnalité ajustée $\widehat{s_t}$

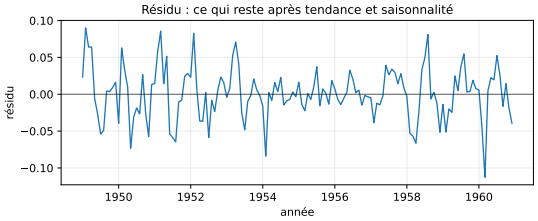


résidu empirique $\widehat{z_t}$

Décomposer, concrètement : saisonnalité et résidu



saisonnalité ajustée $\widehat{s_t}$



résidu empirique $\widehat{z_t}$

La composante saisonnière contient le motif qui revient chaque année. Le résidu sert de contrôle qualité : s'il contient encore une structure visible, la décomposition n'a pas tout expliqué.

Pourquoi inspecter les résidus ?

Après un ajustement, le résidu empirique mesure ce qui reste inexpliqué à chaque date.

Si les résidus ressemblent à un bruit désorganisé autour de zéro, l'ajustement a capturé les structures visibles principales.

Si les résidus montrent encore une tendance, une saisonnalité ou une dépendance nette, il reste quelque chose à modéliser.

Pourquoi inspecter les résidus ?

Après un ajustement, le résidu empirique mesure ce qui reste inexpliqué à chaque date.

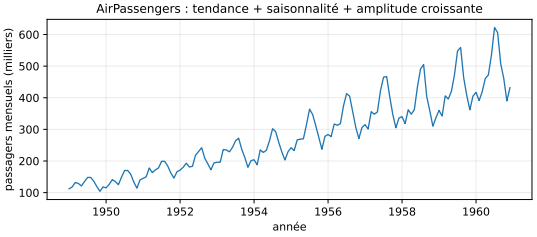
Si les résidus ressemblent à un bruit désorganisé autour de zéro, l'ajustement a capturé les structures visibles principales.

Si les résidus montrent encore une tendance, une saisonnalité ou une dépendance nette, il reste quelque chose à modéliser.

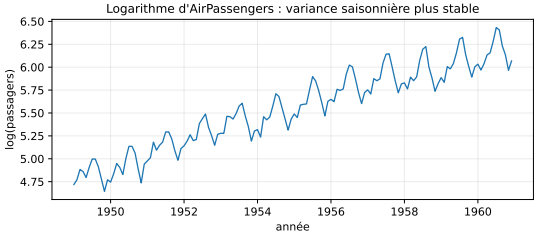
On ne regarde donc pas seulement la courbe ajustée. On regarde aussi ce que cette courbe laisse derrière elle.

Pourquoi prendre le logarithme ?

Certaines séries positives ont des oscillations dont l'amplitude augmente avec le niveau. Il faut alors distinguer variations absolues et variations relatives.



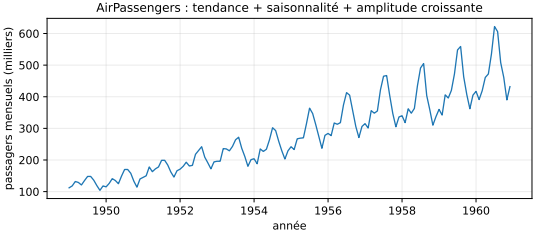
x_t : nombre observé de passagers



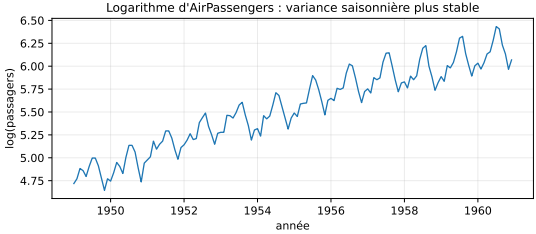
$\log(x_t)$: même série sur échelle logarithmique

Pourquoi prendre le logarithme ?

Certaines séries positives ont des oscillations dont l'amplitude augmente avec le niveau. Il faut alors distinguer variations absolues et variations relatives.



x_t : nombre observé de passagers



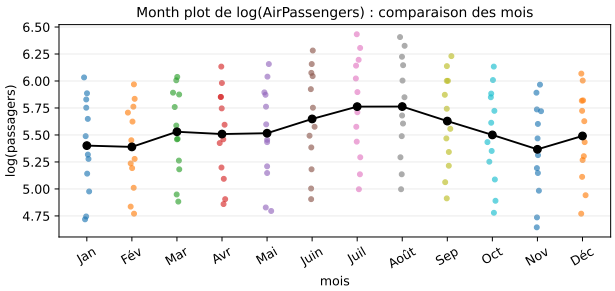
$\log(x_t)$: même série sur échelle logarithmique

Ici x_t est une observation positive, m_t représente son niveau de référence, et u_t une variation relative autour de ce niveau :

$$x_t = m_t(1 + u_t) \quad \Longrightarrow \quad \log x_t = \log m_t + \log(1 + u_t) \simeq \log m_t + u_t.$$

Le logarithme transforme approximativement les variations relatives en variations additives : une hausse de 10% a presque la même taille, que le niveau soit 100 ou 500.

Comparer les saisons : le month plot

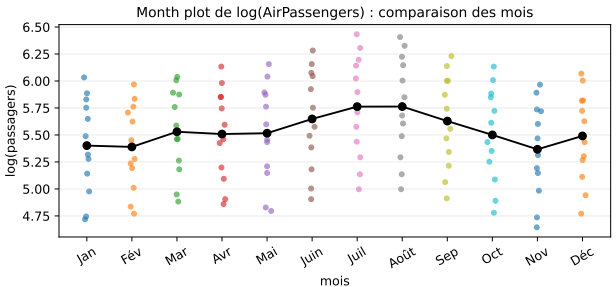


Le chronogramme suit le temps. Le **month plot** réorganise les observations par mois.

Il répond à d'autres questions :

- quels mois sont typiquement hauts ?
- quels mois sont typiquement bas ?
- le motif saisonnier est-il stable ?

Comparer les saisons : le month plot



Le chronogramme suit le temps. Le **month plot** réorganise les observations par mois.

Il répond à d'autres questions :

- quels mois sont typiquement hauts ?
- quels mois sont typiquement bas ?
- le motif saisonnier est-il stable ?

Ici, on lit la saisonnalité de façon plus directe que sur le chronogramme : les mois d'été sont systématiquement plus hauts que les mois d'hiver.

Comparer une valeur à son passé : idée du lag plot

Un chronogramme montre l'évolution au cours du temps, mais il ne montre pas directement si le passé aide à deviner le présent.

Comparer une valeur à son passé : idée du lag plot

Un chronogramme montre l'évolution au cours du temps, mais il ne montre pas directement si le passé aide à deviner le présent.

On part des données observées $x_1, \ldots, x_n$. Pour un retard $h \geq 1$, le **lag plot** représente les points

$$(x_{t-h}, x_t), \qquad t = h+1, \ldots, n.$$

Ici h est le nombre de pas de temps qui sépare les deux observations comparées.

Comparer une valeur à son passé : idée du lag plot

Un chronogramme montre l'évolution au cours du temps, mais il ne montre pas directement si le passé aide à deviner le présent.

On part des données observées $x_1, \ldots, x_n$. Pour un retard $h \geq 1$, le **lag plot** représente les points

$$(x_{t-h}, x_t), \qquad t = h+1, \ldots, n.$$

Ici h est le nombre de pas de temps qui sépare les deux observations comparées.

Nuage sans forme : connaître x_{t-h} aide peu à deviner x_t.

Nuage structuré : le passé contient une information sur le présent.

Comparer une valeur à son passé : idée du lag plot

Un chronogramme montre l'évolution au cours du temps, mais il ne montre pas directement si le passé aide à deviner le présent.

On part des données observées $x_1, \ldots, x_n$. Pour un retard $h \geq 1$, le **lag plot** représente les points

$$(x_{t-h}, x_t), \qquad t = h+1, \ldots, n.$$

Ici h est le nombre de pas de temps qui sépare les deux observations comparées.

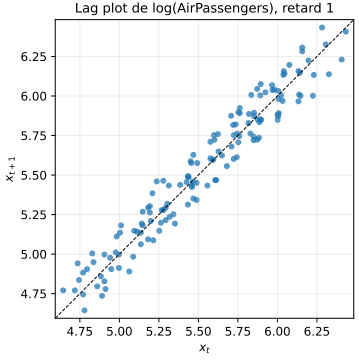
Nuage sans forme : connaître x_{t-h} aide peu à deviner x_t.

Nuage structuré : le passé contient une information sur le présent.

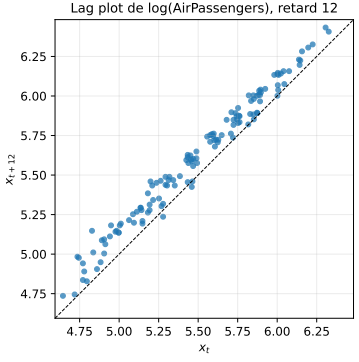
Le lag plot est une version graphique de la question : « la valeur passée explique-t-elle partiellement la valeur courante ? »

Lag plot : retard 1 et retard 12

Sur une série mensuelle, deux retards ont des interprétations différentes : $h = 1$ compare deux mois consécutifs ; $h = 12$ compare un mois avec le même mois de l'année précédente.



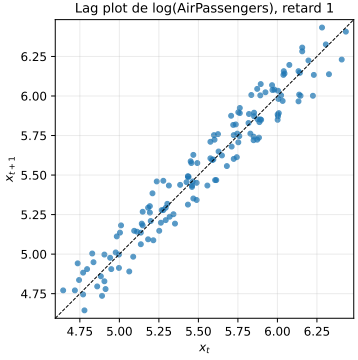
x_t contre x_{t-1} : persistance de court terme



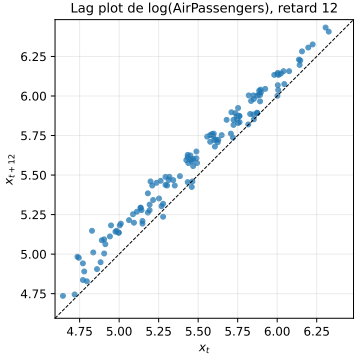
x_t contre x_{t-12} : effet annuel

Lag plot : retard 1 et retard 12

Sur une série mensuelle, deux retards ont des interprétations différentes : $h = 1$ compare deux mois consécutifs ; $h = 12$ compare un mois avec le même mois de l'année précédente.



x_t contre x_{t-1} : persistance de court terme



x_t contre x_{t-12} : effet annuel

Les deux nuages sont très structurés : les valeurs passées sont informatives. Le retard 12 est particulièrement naturel ici, car la série a une saisonnalité annuelle.

Notation : pourquoi introduire l'opérateur retard ?

Les transformations utiles en séries temporelles consistent souvent à comparer une valeur à une valeur passée. Sans notation compacte, les formules deviennent vite lourdes.

Notation : pourquoi introduire l'opérateur retard ?

Les transformations utiles en séries temporelles consistent souvent à comparer une valeur à une valeur passée. Sans notation compacte, les formules deviennent vite lourdes.

Soit $X = (X_t)_{t\in\mathbb{Z}}$ une série. L'**opérateur retard** B transforme X en la série BX définie par

$$(BX)_t = X_{t-1}.$$

Par récurrence, $(B^k X)_t = X_{t-k}$.

Notation : pourquoi introduire l'opérateur retard ?

Les transformations utiles en séries temporelles consistent souvent à comparer une valeur à une valeur passée. Sans notation compacte, les formules deviennent vite lourdes.

Soit $X = (X_t)_{t\in\mathbb{Z}}$ une série. L'**opérateur retard** B transforme X en la série BX définie par

$$(BX)_t = X_{t-1}.$$

Par récurrence, $(B^k X)_t = X_{t-k}$.

Dans les calculs, on écrit souvent $BX_t = X_{t-1}$. C'est une abréviation : rigoureusement, B agit sur toute la série.

Opérateur retard : premiers exemples

L'intérêt de B est de transformer des opérations temporelles en calculs algébriques.

Opération	**Écriture avec B**	**Interprétation**
Valeur précédente	$(BX)_t = X_{t-1}$	comparer au passé immédiat
Même mois un an avant	$(B^{12}X)_t = X_{t-12}$	comparer la même saison
Différence ordinaire	$((1-B)X)_t = X_t - X_{t-1}$	passer du niveau à l'accroissement
Différence saisonnière	$((1-B^{12})X)_t = X_t - X_{t-12}$	enlever un motif annuel simple

Opérateur retard : premiers exemples

L'intérêt de B est de transformer des opérations temporelles en calculs algébriques.

Opération	**Écriture avec B**	**Interprétation**
Valeur précédente	$(BX)_t = X_{t-1}$	comparer au passé immédiat
Même mois un an avant	$(B^{12}X)_t = X_{t-12}$	comparer la même saison
Différence ordinaire	$((1-B)X)_t = X_t - X_{t-1}$	passer du niveau à l'accroissement
Différence saisonnière	$((1-B^{12})X)_t = X_t - X_{t-12}$	enlever un motif annuel simple

Cette notation deviendra essentielle pour les modèles ARMA et ARIMA : leurs équations sont des polynômes en B.

Différencier : comparer au passé immédiat

Pour un processus (X_t), X_t est le niveau modélisé à la date t. La différence ordinaire remplace ce niveau par son accroissement :

$$(\Delta X)_t = X_t - X_{t-1} = ((1 - B)X)_t.$$

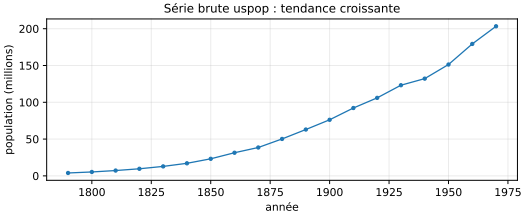
Sur les données, on calcule de même $\Delta x_t = x_t - x_{t-1}$.

Différencier : comparer au passé immédiat

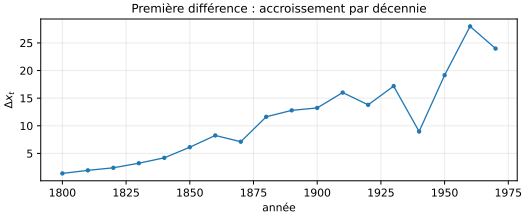
Pour un processus (X_t), X_t est le niveau modélisé à la date t. La différence ordinaire remplace ce niveau par son accroissement :

$$(\Delta X)_t = X_t - X_{t-1} = ((1 - B)X)_t.$$

Sur les données, on calcule de même $\Delta x_t = x_t - x_{t-1}$.



x_t : niveau de population



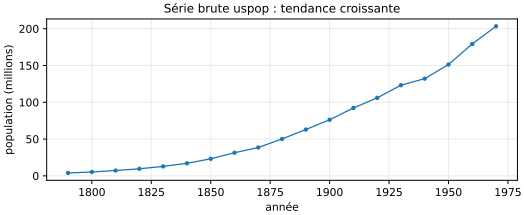
Δx_t : augmentation entre deux dates

Différencier : comparer au passé immédiat

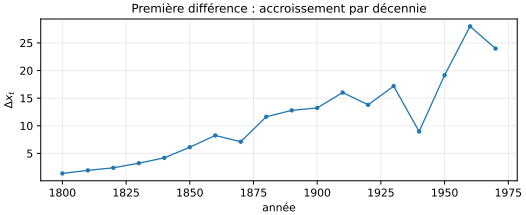
Pour un processus (X_t), X_t est le niveau modélisé à la date t. La différence ordinaire remplace ce niveau par son accroissement :

$$(\Delta X)_t = X_t - X_{t-1} = ((1 - B)X)_t.$$

Sur les données, on calcule de même $\Delta x_t = x_t - x_{t-1}$.



x_t : niveau de population



Δx_t : augmentation entre deux dates

La série brute répond à « quel est le niveau ? ». La différence répond à « de combien a-t-on augmenté depuis la date précédente ? ».

À quoi sert la différence première ?

La différence première est utile quand une série évolue surtout par accumulation.

À quoi sert la différence première ?

La différence première est utile quand une série évolue surtout par accumulation.

Pour une tendance déterministe $x_t = a + bt$, les paramètres a et b sont des constantes :

$$x_t - x_{t-1} = b.$$

La différence retire la partie linéaire.

Pour une marche aléatoire, Z_t est le choc nouveau entre $t-1$ et t :

$$X_t = X_{t-1} + Z_t \quad \Rightarrow \quad \Delta X_t = Z_t.$$

La différence isole les innovations.

À quoi sert la différence première ?

La différence première est utile quand une série évolue surtout par accumulation.

Pour une tendance déterministe $x_t = a + bt$, les paramètres a et b sont des constantes :

$$x_t - x_{t-1} = b.$$

La différence retire la partie linéaire.

Pour une marche aléatoire, Z_t est le choc nouveau entre $t-1$ et t :

$$X_t = X_{t-1} + Z_t \quad \Rightarrow \quad \Delta X_t = Z_t.$$

La différence isole les innovations.

La différence ne sert pas à « rendre les données jolies ». Elle change la question : on ne modélise plus le niveau, mais l'évolution d'une date à la suivante.

Différence saisonnière : comparer à la même saison précédente

Pour une série mensuelle, comparer février à janvier peut mélanger évolution et saisonnalité. Pour retirer une saisonnalité annuelle, on compare plutôt février avec février de l'année précédente.

Si s est la période saisonnière, par exemple $s = 12$ pour des données mensuelles, la différence saisonnière est

$$(\Delta_s X)_t = X_t - X_{t-s} = ((1 - B^s)X)_t.$$

Sur les données observées : $\Delta_s x_t = x_t - x_{t-s}$.

Différence saisonnière : comparer à la même saison précédente

Pour une série mensuelle, comparer février à janvier peut mélanger évolution et saisonnalité. Pour retirer une saisonnalité annuelle, on compare plutôt février avec février de l'année précédente.

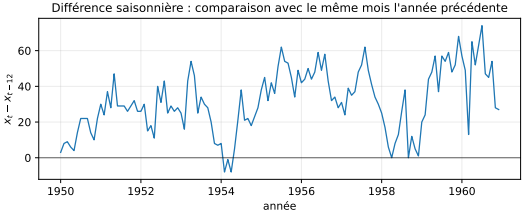
Si s est la période saisonnière, par exemple $s = 12$ pour des données mensuelles, la différence saisonnière est

$$(\Delta_s X)_t = X_t - X_{t-s} = ((1 - B^s)X)_t.$$

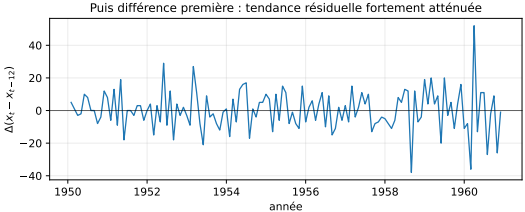
Sur les données observées : $\Delta_s x_t = x_t - x_{t-s}$.

Si un effet saisonnier se répète d'une année sur l'autre, la différence saisonnière le réduit fortement, car on soustrait deux observations prises dans la même position du cycle.

Différence saisonnière : ce que montrent les graphiques

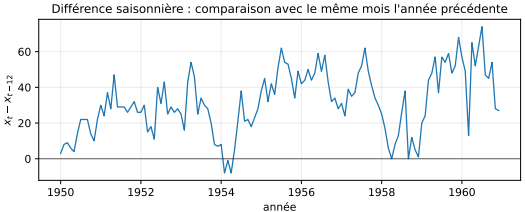


$\Delta_{12} \log(x_t)$: comparaison au même mois l'année précédente

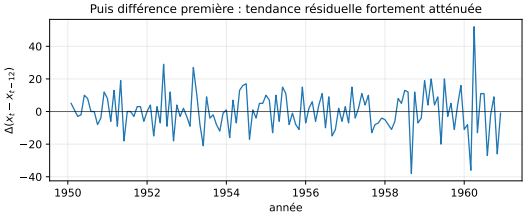


$\Delta\Delta_{12} \log(x_t)$: on ajoute une différence ordinaire

Différence saisonnière : ce que montrent les graphiques



$\Delta_{12} \log(x_t)$: comparaison au même mois l'année précédente



$\Delta\Delta_{12} \log(x_t)$: on ajoute une différence ordinaire

À gauche, la saisonnalité annuelle est fortement réduite, mais il peut rester une évolution lente. À droite, on enlève aussi une partie de cette évolution : la série devient plus comparable d'une date à l'autre.

Pourquoi parler de stationnarité ?

Les transformations précédentes ont un objectif commun : obtenir une série dont le comportement est suffisamment stable pour être modélisé.

Pourquoi parler de stationnarité ?

Les transformations précédentes ont un objectif commun : obtenir une série dont le comportement est suffisamment stable pour être modélisé.

Idée intuitive : une série stationnaire n'a pas de date privilégiée. Elle peut fluctuer, mais ses règles de fluctuation ne changent pas avec le temps.

Pourquoi parler de stationnarité ?

Les transformations précédentes ont un objectif commun : obtenir une série dont le comportement est suffisamment stable pour être modélisé.

Idée intuitive : une série stationnaire n'a pas de date privilégiée. Elle peut fluctuer, mais ses règles de fluctuation ne changent pas avec le temps.

Exemples à garder en tête :

- un bruit blanc centré de variance constante est stationnaire ;
- une marche aléatoire n'est pas stationnaire car sa variance augmente avec le temps ;
- une série avec tendance déterministe visible n'est généralement pas stationnaire.

Stationnarité forte

Un processus $X = (X_t)_{t\in\mathbb{Z}}$ est **strictement stationnaire** si, pour tous $k \geq 1$, tous $t_1, \ldots, t_k \in \mathbb{Z}$ et tout $h \in \mathbb{Z}$,

$$(X_{t_1}, \ldots, X_{t_k}) \stackrel{\mathcal{L}}{=} (X_{t_1+h}, \ldots, X_{t_k+h}).$$

Stationnarité forte

Un processus $X = (X_t)_{t\in\mathbb{Z}}$ est **strictement stationnaire** si, pour tous $k \geq 1$, tous $t_1, \ldots, t_k \in \mathbb{Z}$ et tout $h \in \mathbb{Z}$,

$$(X_{t_1}, \ldots, X_{t_k}) \overset{\mathcal{L}}{=} (X_{t_1+h}, \ldots, X_{t_k+h}).$$

Cette définition dit que déplacer toute la fenêtre d'observation ne change pas la loi jointe. Elle est très naturelle, mais souvent difficile à vérifier directement.

Stationnarité faible : ce qu'on utilisera le plus

Un processus $X = (X_t)_{t\in\mathbb{Z}}$ est **stationnaire du second ordre** si :

1. $\mathbb{E}[X_t^2] < \infty$ pour tout t ;
2. $\mathbb{E}[X_t] = \mu$ ne dépend pas de t ;
3. $\mathrm{Cov}(X_{t+h}, X_t)$ ne dépend que du retard h.

Stationnarité faible : ce qu'on utilisera le plus

Un processus $X = (X_t)_{t\in\mathbb{Z}}$ est **stationnaire du second ordre** si :

1. $\mathbb{E}[X_t^2] < \infty$ pour tout t ;
2. $\mathbb{E}[X_t] = \mu$ ne dépend pas de t ;
3. $\mathrm{Cov}(X_{t+h}, X_t)$ ne dépend que du retard h.

La **fonction d'autocovariance** est

$$\gamma(h) = \mathrm{Cov}(X_{t+h}, X_t).$$

Si $\gamma(0) > 0$, l'**autocorrélation** est

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}.$$

Comment interpréter l'autocovariance ?

$\gamma(0) = \mathrm{Var}(X_t)$ mesure l'amplitude moyenne des fluctuations autour de la moyenne.

$\gamma(h) > 0$: deux valeurs séparées par h pas de temps ont tendance à être du même côté de la moyenne.

$\gamma(h) < 0$: une valeur haute tend à être associée, au retard h, à une valeur basse.

$\rho(h)$ met cette dépendance sur une échelle normalisée entre -1 et 1.

Comment interpréter l'autocovariance ?

$\gamma(0) = \mathrm{Var}(X_t)$ mesure l'amplitude moyenne des fluctuations autour de la moyenne.

$\gamma(h) < 0$: une valeur haute tend à être associée, au retard h, à une valeur basse.

$\gamma(h) > 0$: deux valeurs séparées par h pas de temps ont tendance à être du même côté de la moyenne.

$\rho(h)$ met cette dépendance sur une échelle normalisée entre -1 et 1.

Le lag plot montre visuellement la dépendance. L'autocorrélation en donne une mesure numérique.

Quelques contraintes sur une autocovariance

Si X est stationnaire du second ordre, alors

$$\gamma(0) \geq 0, \qquad \gamma(-h) = \gamma(h), \qquad |\gamma(h)| \leq \gamma(0).$$

De plus, pour tout n et tous $a_1, \ldots, a_n \in \mathbb{R}$,

$$\sum_{i,j=1}^{n} a_i a_j \gamma(i-j) \geq 0.$$

Quelques contraintes sur une autocovariance

Si X est stationnaire du second ordre, alors

$$\gamma(0) \geq 0, \qquad \gamma(-h) = \gamma(h), \qquad |\gamma(h)| \leq \gamma(0).$$

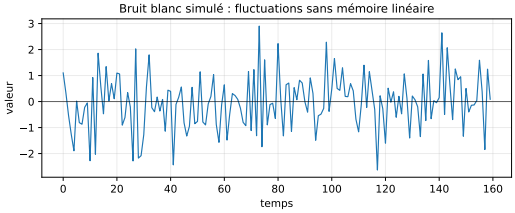
De plus, pour tout n et tous $a_1, \ldots, a_n \in \mathbb{R}$,

$$\sum_{i,j=1}^{n} a_i a_j \, \gamma(i-j) \geq 0.$$

Pourquoi ? Pour $Y = \sum_i a_i X_i$,

$$\mathrm{Var}(Y) = \sum_{i,j=1}^{n} a_i a_j \, \mathrm{Cov}(X_i, X_j) = \sum_{i,j=1}^{n} a_i a_j \gamma(i-j) \geq 0.$$

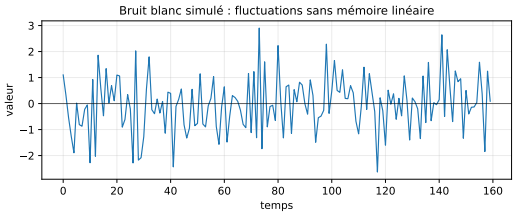
Bruit blanc : absence de mémoire linéaire



Un **bruit blanc faible** de moyenne 0 et variance σ^2 est un processus (Z_t) tel que

$$\mathbb{E}[Z_t] = 0, \qquad \gamma_Z(0) = \sigma^2,$$
$$\gamma_Z(h) = 0 \quad (h \neq 0).$$

Bruit blanc : absence de mémoire linéaire

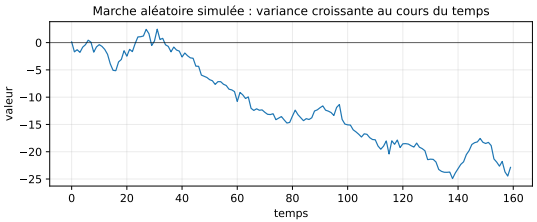


Un **bruit blanc faible** de moyenne 0 et variance σ^2 est un processus (Z_t) tel que

$$\mathbb{E}[Z_t] = 0, \qquad \gamma_Z(0) = \sigma^2,$$
$$\gamma_Z(h) = 0 \quad (h \neq 0).$$

Sur le graphique, il n'y a ni tendance, ni saisonnalité, ni structure évidente. Le bruit blanc sert de brique de base aux modèles linéaires.

Contre-exemple important : la marche aléatoire



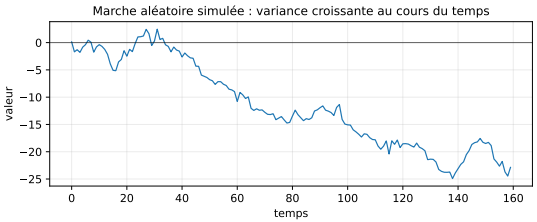
Si

$$X_t = X_{t-1} + Z_t,$$

alors

$$X_t = X_0 + Z_1 + \cdots + Z_t.$$

Contre-exemple important : la marche aléatoire



Si

$$X_t = X_{t-1} + Z_t,$$

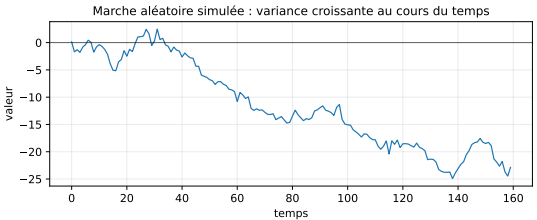
alors

$$X_t = X_0 + Z_1 + \cdots + Z_t.$$

Si les Z_t sont centrés de variance σ^2 et non corrélés,

$$\mathrm{Var}(X_t) = \mathrm{Var}(X_0) + t\sigma^2.$$

Contre-exemple important : la marche aléatoire



Si

$$X_t = X_{t-1} + Z_t,$$

alors

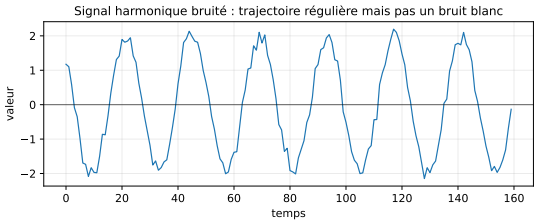
$$X_t = X_0 + Z_1 + \cdots + Z_t.$$

Si les Z_t sont centrés de variance σ^2 et non corrélés,

$$\mathrm{Var}(X_t) = \mathrm{Var}(X_0) + t\sigma^2.$$

Le graphique montre une trajectoire qui dérive. Mathématiquement, sa variabilité dépend de t : ce n'est pas stationnaire du second ordre.

Exemple stationnaire non trivial : oscillations aléatoires



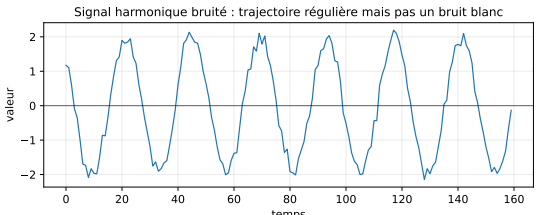
Un processus peut avoir un aspect périodique sans être déterministe.

Exemple typique :

$$X_t = A\cos(\omega t) + B\sin(\omega t),$$

où A et B sont deux variables aléatoires centrées, de même variance, non corrélées, et ω fixe la fréquence d'oscillation.

Exemple stationnaire non trivial : oscillations aléatoires



Un processus peut avoir un aspect périodique sans être déterministe.

Exemple typique :

$$X_t = A\cos(\omega t) + B\sin(\omega t),$$

où A et B sont deux variables aléatoires centrées, de même variance, non corrélées, et ω fixe la fréquence d'oscillation. Alors

$$\mathbb{E}[X_t] = 0, \qquad \gamma(h) = \sigma^2 \cos(\omega h).$$

Le graphique oscille, mais ses propriétés probabilistes ne changent pas quand on décale le temps.

Méthode de lecture à retenir

Pour une nouvelle série, on ne commence pas par choisir un modèle. On suit une chaîne de questions.

1. Chronogramme : tendance ? saisonnalité ? variance stable ? rupture ?
2. Transformation : logarithme utile ? différence utile ? différence saisonnière ?
3. Décomposition : que reste-t-il après tendance et saisonnalité ?
4. Dépendance : les retards semblent-ils informatifs ?
5. Stationnarité : peut-on raisonnablement chercher un modèle stationnaire pour le reste transformé ?

Méthode de lecture à retenir

Pour une nouvelle série, on ne commence pas par choisir un modèle. On suit une chaîne de questions.

1. Chronogramme : tendance ? saisonnalité ? variance stable ? rupture ?
2. Transformation : logarithme utile ? différence utile ? différence saisonnière ?
3. Décomposition : que reste-t-il après tendance et saisonnalité ?
4. Dépendance : les retards semblent-ils informatifs ?
5. Stationnarité : peut-on raisonnablement chercher un modèle stationnaire pour le reste transformé ?

Les modèles des prochains chapitres servent à formaliser et automatiser ces questions.

Avant le TD/TP

Le TD/TP 1 reprend une partie de ce qu'on vient de voir, sur d'autres séries fournies avec R.

- lire un chronogramme ;
- repérer tendance, saisonnalité, variabilité ;
- comprendre pourquoi on regarde le logarithme ;
- ajuster une tendance simple ;
- calculer et lire des résidus.

Avant le TD/TP

Le TD/TP 1 reprend une partie de ce qu'on vient de voir, sur d'autres séries fournies avec R.

- lire un chronogramme ;
- repérer tendance, saisonnalité, variabilité ;
- comprendre pourquoi on regarde le logarithme ;
- ajuster une tendance simple ;
- calculer et lire des résidus.

Le but n'est pas de trouver le bon modèle dès la première séance. Le but est d'apprendre à poser les bonnes questions devant une série temporelle.