UNIVERSITÀ
DI SIENA
1240

Serie storiche economiche

Stagionalità e Medie Mobili

Dott. Damiano Leo

Università di Siena
Corso di Statistica per l'economia e l'impresa

damiano.leo@unifi.it

Struttura della lezione

1. **Recap**: cosa è una serie storica e come la scomponiamo
2. **La componente stagionale**
3. **Medie mobili**

Che cos'è una serie storica?

Una **serie storica** è una sequenza di osservazioni ordinate nel tempo:

$$y_1, y_2, \ldots, y_n$$

Idea chiave

L'ordine temporale conta: le osservazioni **non sono intercambiabili**.

Esempi

PIL annuale, tasso di disoccupazione mensile, vendite trimestrali, consumi energetici giornalieri.

Il modello generale: approccio classico e moderno

$$Y_t = f(t) + \varepsilon_t$$

Approccio classico

L'attenzione è rivolta soprattutto a:

$$f(t)$$

cioè alla parte deterministica della serie. La componente accidentale è trattata come residuo non sistematico, un rumore bianco.

Approccio moderno

L'attenzione si sposta anche sulla struttura stocastica:

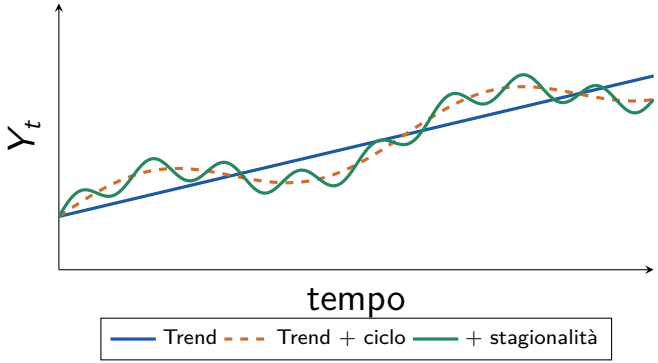
$$\varepsilon_t$$

che può essere autocorrelata e modellata direttamente.

Le componenti di una serie storica

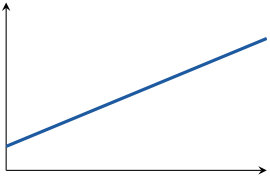
$$Y_t = T_t + C_t + S_t + \varepsilon_t$$

- T_t: trend
- C_t: ciclo
- S_t: stagionalità
- ε_t: componente di errore



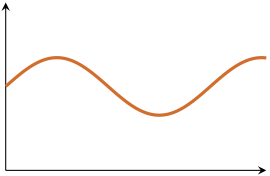
Trend, ciclo e stagionalità: come distinguerli?

Trend



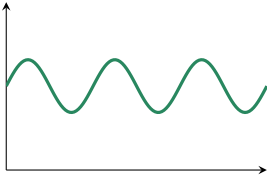
Tendenza di fondo di lungo periodo.

Ciclo



Fluttuazioni di medio-lungo periodo attorno al trend.

Stagionalità



Movimento regolare che si ripete nello stesso periodo dell'anno.

Tipi di composizione delle componenti

Modello additivo

$$Y_t = T_t + C_t + S_t + \varepsilon_t$$

Si usa quando le oscillazioni stagionali hanno ampiezza circa costante.

Modello misto

$$Y_t = T_t \cdot S_t + C_t + \varepsilon_t$$

Alcune componenti interagiscono moltiplicativamente, altre additivamente.

Modello moltiplicativo

$$Y_t = T_t \cdot C_t \cdot S_t \cdot a_t$$

Si usa quando le oscillazioni crescono con il livello della serie.

Logaritmi

$$\log Y_t = \log T_t + \log C_t + \log S_t + \log a_t$$

Il modello moltiplicativo diventa additivo nei logaritmi.

Il trend: come rappresentiamo $f(t)$?

Se assumiamo che la parte sistematica sia composta dal solo trend:

$$Y_t = f(t) + \varepsilon_t$$

il problema diventa:

come scegliamo $f(t)$?

Trend noti a meno di parametri

- Polinomiale;
- Esponenziale;
- Logistico;
- Gompertz.

Trend non specificati

- Approssimazioni locali;
- Procedure di smoothing;
- Medie mobili.

Take-away del recap

1. Una serie storica è ordinata nel tempo: l'ordine delle osservazioni conta.
2. Nell'approccio classico:
$$Y_t = f(t) + \varepsilon_t$$
3. La parte sistematica può essere scomposta in trend, ciclo e stagionalità.
4. Il modello può essere additivo, moltiplicativo o misto.
5. Il trend può essere rappresentato con funzioni diverse: polinomiali, esponenziali, logistiche, Gompertz.

La componente stagionale

Che cos'è la componente stagionale?

La **componente stagionale** descrive i movimenti di una serie storica che si ripetono con regolarità nel tempo.

Esempio

Con dati mensili possiamo osservare comportamenti simili nello stesso mese di anni diversi:

- le vendite di panettoni aumentano ogni dicembre;
- il turismo cresce nei mesi estivi;
- i consumi energetici aumentano in inverno.

Punto importante

La stagionalità non è una variazione casuale: è una componente sistematica che tende a ripetersi.

Il periodo stagionale

Indichiamo con s il **periodo stagionale**, cioè il numero di osservazioni dopo il quale il comportamento tende a ripetersi.

Dati mensili

$$s = 12$$

Il ciclo stagionale si completa in 12 mesi.

Dati trimestrali

$$s = 4$$

Il ciclo stagionale si completa in 4 trimestri.

Una componente stagionale soddisfa:

$$S_{t+s} = S_t$$

A distanza di s periodi si ripete lo stesso effetto stagionale, non necessariamente lo stesso valore osservato della serie.

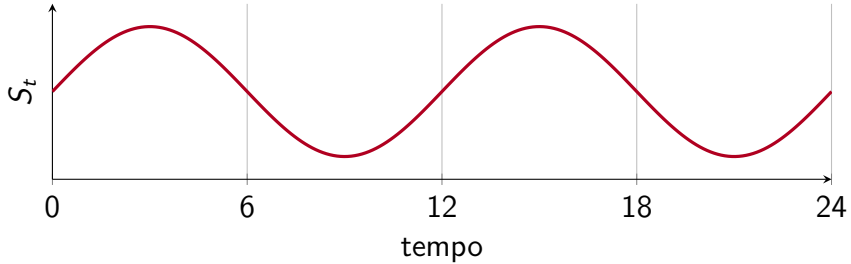
Come rappresentare graficamente la stagionalità?

Una componente stagionale:

- sale e scende nel tempo;
- si ripete dopo un periodo s;
- presenta picchi e minimi in periodi ricorrenti.

Idea naturale

Usiamo funzioni periodiche, come seno e coseno, perché generano onde regolari.



Che cos'è un'armonica?

Una **armonica** è una singola onda regolare, costruita con seno o coseno, che compie un certo numero di oscillazioni dentro il periodo stagionale.

Se $s = 12$, cioè con dati mensili:

$$\text{prima armonica} \Rightarrow 1 \text{ ciclo in } 12 \text{ mesi}$$

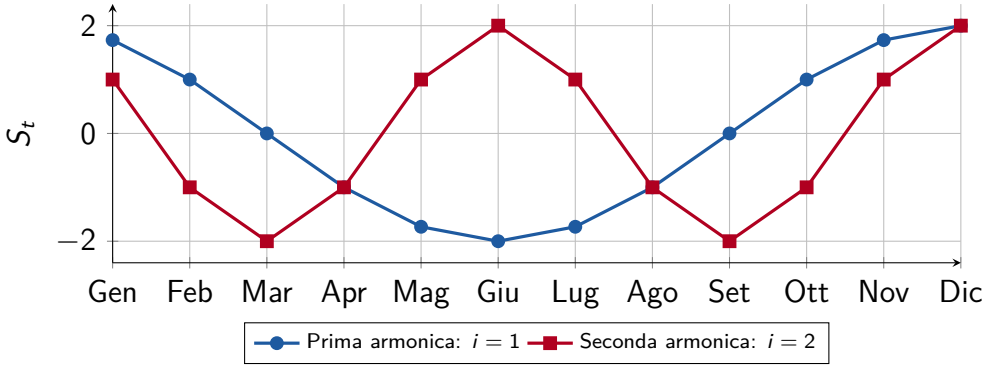
$$\text{seconda armonica} \Rightarrow 2 \text{ cicli in } 12 \text{ mesi}$$

$$\text{terza armonica} \Rightarrow 3 \text{ cicli in } 12 \text{ mesi}$$

Attenzione

Dire "seconda armonica" non significa "due armoniche". Significa una sola onda che compie due cicli nel periodo stagionale.

Prima e seconda armonica



L'indice i indica quante oscillazioni complete l'onda compie nel periodo s.

Perché una sola armonica può non bastare?

Una sola armonica può essere troppo semplice per descrivere la stagionalità osservata nei dati.

Nei dati reali possiamo avere:

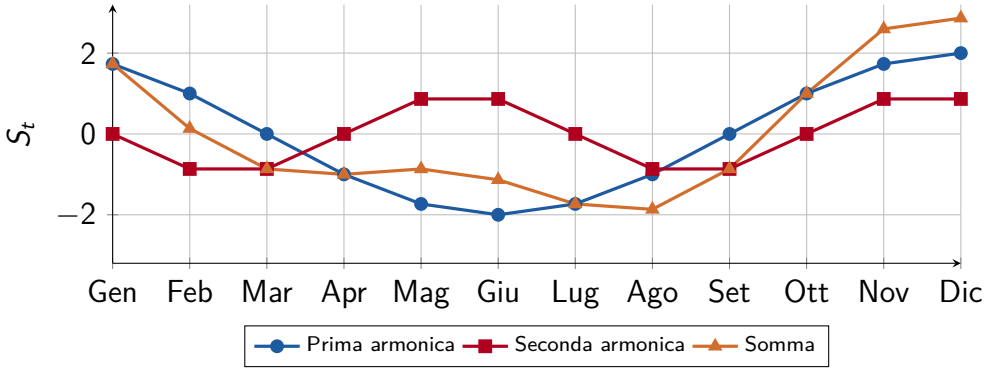
- più picchi nello stesso anno;
- aumenti rapidi e diminuzioni lente;
- mesi particolari con effetti molto forti;
- forme non perfettamente simmetriche.

Idea

Per ottenere una stagionalità più flessibile possiamo sommare più armoniche.

Questa è l'idea alla base della rappresentazione di Fourier.

Somma di armoniche



La somma produce una forma stagionale più articolata.

La serie di Fourier: idea generale

La serie di Fourier rappresenta una funzione periodica come somma di onde elementari:

$$f(t) = a_0 + \sum_{i=1}^{m} [a_i \cos(\omega_i t) + b_i \sin(\omega_i t)]$$

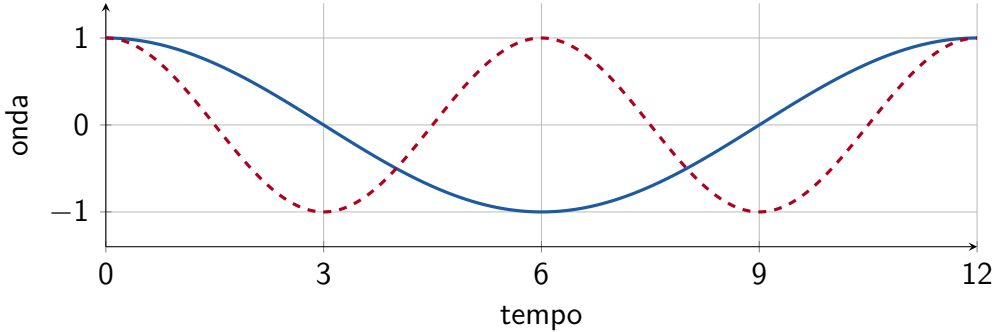
Interpretazione

- a_0: livello medio della funzione;
- i: indice dell'armonica;
- ω_i: frequenza angolare dell'armonica, ovvero quanto velocemente seno e coseno completano il loro ciclo;
- a_i, b_i: coefficienti che determinano il peso delle singole onde.

Una forma stagionale complessa può essere costruita sommando più onde semplici.

Frequenza angolare: idea visiva

La frequenza angolare indica quanto velocemente l'onda percorre il suo ciclo.



Interpretazione

A parità di intervallo temporale, una curva con frequenza angolare più alta oscilla più rapidamente ovvero: la curva completa più cicli nel periodo s.

Fourier adattata alla stagionalità

Per una componente stagionale di periodo s deve valere:

$$S_t = S_{t+s}$$

cioè il comportamento stagionale si ripete dopo s osservazioni. Dal punto di vista grafico questo vuol dire che: nel grafico a sinistra il livello in $t + s$ è tornato a quello in t, mentre nel grafico a destra la funzione seno/coseno ha compiuto un giro completo.

Fourier adattata alla stagionalità

Dal punto di vista matematico questo vuol dire che, partendo dal coseno:

$$S_{t+s} = S_t \quad \Longleftrightarrow \quad \cos\big(\omega_i(t+s)\big) = \cos(\omega_i t)$$

Questo accade se, dopo s osservazioni, l'argomento aumenta di un numero intero di giri completi:

$$\omega_i(t+s) = \omega_i t + 2\pi i, \qquad i = 1, 2, \ldots$$

Da cui:

$$\omega_i t + \omega_i s = \omega_i t + 2\pi i \quad \Longrightarrow \quad \omega_i s = 2\pi i \quad \Longrightarrow \quad \omega_i = \frac{2\pi i}{s}$$

Rappresentazione della componente stagionale

$$S_t = a_0 + \sum_{i=1}^{m} \left[a_i \cos\left(\frac{2\pi i}{s} t\right) + b_i \sin\left(\frac{2\pi i}{s} t\right) \right]$$

Il termine a_0

Nella formula:

$$S_t = a_0 + \sum_{i=1}^{m} \left[a_i \cos\left(\frac{2\pi i}{s} t\right) + b_i \sin\left(\frac{2\pi i}{s} t\right) \right]$$

il termine a_0 rappresenta il **livello medio** della componente periodica.

La stagionalità viene considerata come neutra su un intero periodo:

$$\sum_{j=1}^{s} S_j = 0$$

Per questo, nella rappresentazione della sola componente stagionale, spesso si pone:

$$a_0 = 0$$

Riscrittura della formula

La forma attuale della componente stagionale è:

$$S_t = \sum_{i=1}^{m} \left[a_i \cos\left(\frac{2\pi i}{s} t\right) + b_i \sin\left(\frac{2\pi i}{s} t\right) \right]$$

Obiettivo

Riscrivere la stessa espressione in una forma equivalente, ma più semplice da interpretare.

Una formula più intuitiva

$$S_t = \sum_{i=1}^{m} A_i \cos\left(\frac{2\pi i}{s} t - \phi_i\right)$$

- A_i: ampiezza dell'armonica;
- $\frac{2\pi i}{s}$: frequenza angolare;
- ϕ_i: fase, cioè spostamento temporale.

Così ogni armonica ha un'interpretazione più chiara.

Dalla forma seno-coseno alla forma ampiezza-fase

Consideriamo una singola armonica:

$$a_i \cos\left(\frac{2\pi i}{s} t\right) + b_i \sin\left(\frac{2\pi i}{s} t\right)$$

Per semplificare poniamo:

$$x_i = \frac{2\pi i}{s} t$$

Quindi abbiamo:

$$a_i \cos(x_i) + b_i \sin(x_i)$$

Vogliamo riscriverla nella forma più intuitiva:

$$A_i \cos(x_i - \phi_i)$$

Dimostrazione: identità trigonometrica

Usiamo:

$$\cos(x_i - \phi_i) = \cos(x_i)\cos(\phi_i) + \sin(x_i)\sin(\phi_i)$$

Moltiplicando per A_i:

$$A_i \cos(x_i - \phi_i) = A_i \cos(x_i)\cos(\phi_i) + A_i \sin(x_i)\sin(\phi_i)$$

Raccogliendo:

$$A_i \cos(x_i - \phi_i) = (A_i \cos\phi_i)\cos(x_i) + (A_i \sin\phi_i)\sin(x_i)$$

Ponendo $a_i = A_i \cos\phi_i$ e $b_i = A_i \sin\phi_i$ otteniamo:

$$A_i \cos(x_i - \phi_i) = a_i \cos(x_i) + b_i \sin(x_i)$$

Forma finale della componente stagionale

Applicando il passaggio a tutte le armoniche:

$$S_t = \sum_{i=1}^{m} \left[a_i \cos\left(\frac{2\pi i}{s} t\right) + b_i \sin\left(\frac{2\pi i}{s} t\right) \right]$$

può essere riscritta come:

$$S_t = \sum_{i=1}^{m} A_i \cos\left(\frac{2\pi i}{s} t - \phi_i\right)$$

Perché questa forma è più intuitiva?

Fa vedere direttamente:

- quanto è forte l'oscillazione;
- quante oscillazioni avvengono nel periodo;
- dove cadono massimi e minimi.

Interpretazione dei parametri

Per una singola armonica:

$$S_t = A_i \cos\left(\frac{2\pi i}{s} t - \phi_i\right)$$

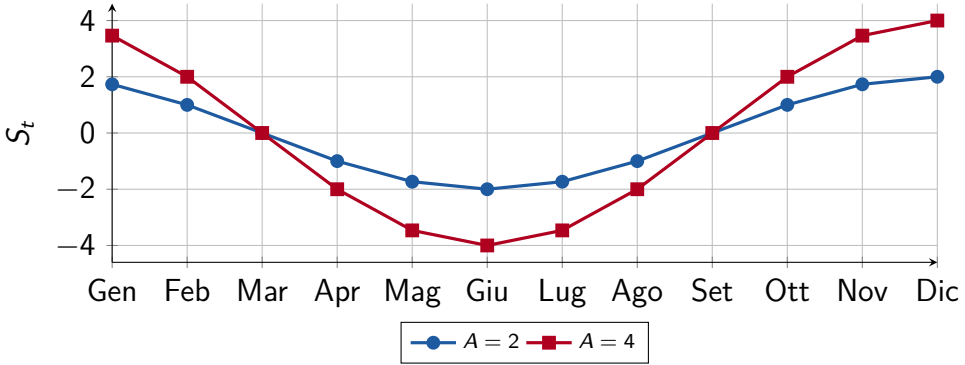
Ampiezza	Frequenza angolare	Fase
A_i	$\frac{2\pi i}{s}$	ϕ_i
Misura l'intensità dell'oscillazione.	Indica la velocità dell'oscillazione.	Sposta l'onda nel tempo.

Questi tre elementi determinano forma, intensità e posizione della componente stagionale.

Effetto dell'ampiezza

Confrontiamo:

$$S_t = 2\cos\left(\frac{2\pi}{12}t\right) \qquad \text{e} \qquad S_t = 4\cos\left(\frac{2\pi}{12}t\right)$$



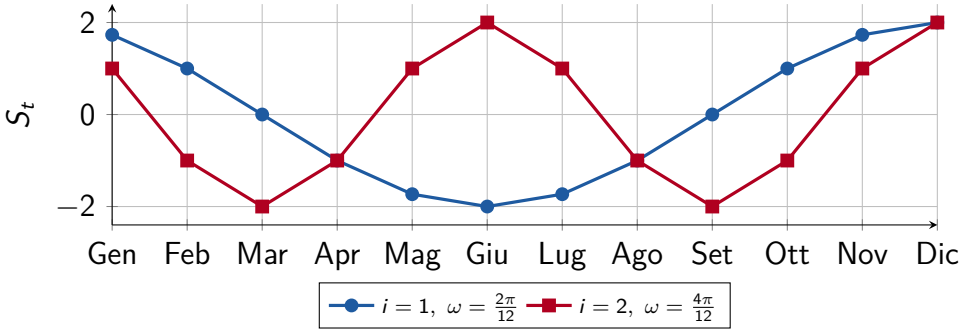
Aumentando A, la stagionalità diventa più intensa.

Effetto della frequenza angolare

Confrontiamo due armoniche:

$$S_t = 2\cos\left(\frac{2\pi}{12}t\right) \qquad \text{e} \qquad S_t = 2\cos\left(\frac{4\pi}{12}t\right)$$

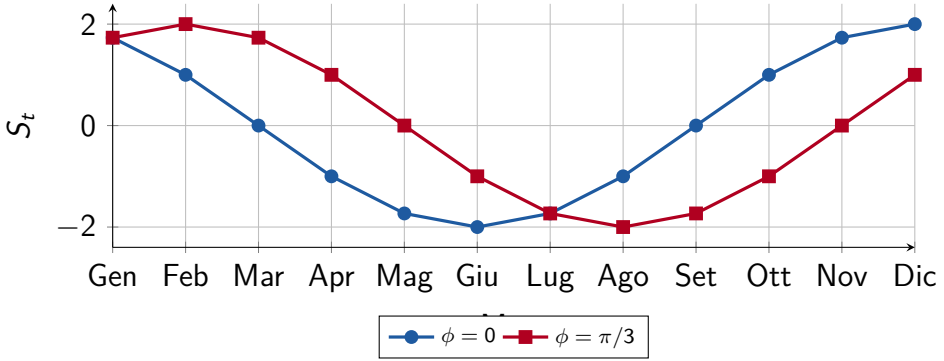
La frequenza angolare è $\frac{2\pi i}{s}$: aumentando i, la funzione compie più oscillazioni nel periodo s.



Effetto della fase

Confrontiamo:

$$S_t = 2\cos\left(\frac{2\pi}{12}t\right) \qquad \text{e} \qquad S_t = 2\cos\left(\frac{2\pi}{12}t - \frac{\pi}{3}\right)$$



La fase non cambia l'altezza dell'onda: sposta nel tempo massimi e minimi.

Perché il meno nella fase sposta a destra?

In generale, per:

$$\cos(t)$$

Il massimo si ha quando:

$$t = 0,\ 2\pi,\ 4\pi, \ldots$$

Ora consideriamo:

$$\cos(t-2)$$

Il massimo si ha quando:

$$t - 2 = 0 \quad \Longrightarrow \quad t = 2$$

Interpretazione

La funzione $\cos(t-2)$ raggiunge in $t = 2$ il valore che $\cos(t)$ raggiungeva in $t = 0$.

Consideriamo il nostro caso

$$S_t = 2\cos\left(\frac{2\pi}{12}t - \frac{\pi}{3}\right)$$

Per capire dove cade il massimo, poniamo l'argomento del coseno uguale a zero:

$$\frac{2\pi}{12}t - \frac{\pi}{3} = 0$$

Poiché $\frac{2\pi}{12} = \frac{\pi}{6}$, otteniamo:

$$\frac{\pi}{6}t - \frac{\pi}{3} = 0 \quad \Longrightarrow \quad \frac{\pi}{6}t = \frac{\pi}{3} \quad \Longrightarrow \quad t = \frac{\pi}{3} \cdot \frac{6}{\pi} = 2$$

Quindi il massimo della curva con fase cade in $t = 2$, cioè a febbraio.

Quante armoniche consideriamo?

Il numero massimo di armoniche utili dipende dal periodo stagionale s.

Se s è pari:

$$m_{\max} = \frac{s}{2}$$

Se s è dispari:

$$m_{\max} = \frac{s-1}{2}$$

Dati mensili

$$s = 12 \implies m_{\max} = \frac{12}{2} = 6$$

Con dati mensili possiamo arrivare al massimo alla sesta armonica.

Dati quadrimestrali

$$s = 3 \implies m_{\max} = \frac{3-1}{2} = 1$$

Con dati quadrimestrali possiamo arrivare al massimo ad una armonica.

Take-away sulla stagionalità

1. La stagionalità è un movimento che si ripete dopo un periodo s.
2. Le armoniche sono delle curve il cui grado indica il numero di oscillazioni all'interno del periodo s.
3. La somma di più armoniche permette di rappresentare forme stagionali più flessibili.
4. La formula di Fourier adatta alla stagionalità è:

$$S_t = \sum_{i=1}^{m} A_i \cos\left(\frac{2\pi i}{s} t - \phi_i\right)$$

5. A_i controlla l'ampiezza, $\frac{2\pi i}{s}$ la frequenza angolare, ϕ_i la fase.

Medie mobili

Perché introduciamo le medie mobili?

Finora abbiamo visto che una serie storica può essere pensata come somma o prodotto di componenti:
Noi però osserviamo solo:

$$Y_t$$

cioè il valore complessivo della serie al tempo t.

Problema

Guardando solo Y_t, non sappiamo distinguere immediatamente:

- quanto dipende dal movimento di fondo;
- quanto dipende dalla stagionalità;
- quanto dipende dalla componente accidentale.

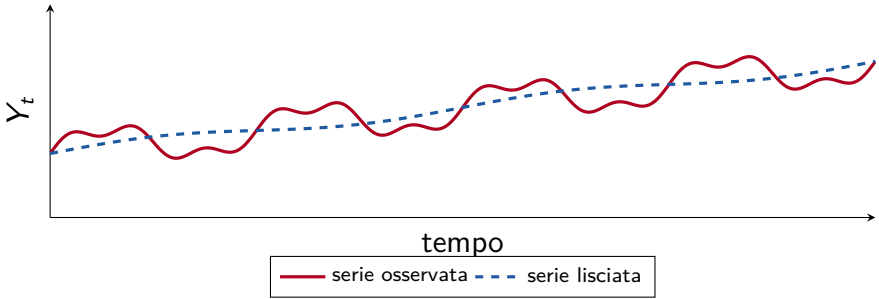
La media mobile è un primo strumento per lisciare la serie e far emergere il trend-ciclo.
Ci permetterà di calcolare poi una serie destagionalizzata.

Idea intuitiva

Una media mobile sostituisce il valore della serie originale con una media dei valori vicini.

Intuizione

Se un valore è molto alto o molto basso per ragioni occasionali, facendo la media con i valori vicini l'effetto viene attenuato.



Definizione generale di media mobile

Una media mobile è una combinazione lineare di osservazioni vicine nel tempo:

$$Y_t^* = \sum_{i=-m_1}^{m_2} \vartheta_i Y_{t+i} \qquad \text{dove} \qquad \vartheta_i = \frac{w_i}{\sum w_i}$$

Elementi della formula

- Y_t^*: valore lisciato al tempo t
- Y_{t+i}: valori osservati vicini a t
- ϑ_i: pesi della media mobile
- w_i: pesi grezzi
- m_1: valori precedenti
- m_2: valori successivi

La media mobile trasforma la serie originale in una serie più liscia.

Ordine della media mobile

L'**ordine** della media mobile indica quanti valori entrano nel calcolo.

Nella formula generale:

$$Y_t^* = \sum_{i=-m_1}^{m_2} \vartheta_i Y_{t+i}$$

il numero totale di osservazioni utilizzate è:

$$m_1 + m_2 + 1$$

Interpretazione

- m_1: valori passati usati nella media;
- m_2: valori futuri usati nella media;
- $+1$: valore corrente Y_t.

Un semplice esempio numerico

Consideriamo una serie osservata in più anni:

Anno	1999	2000	2001	2002	2003
Y_t	7	4	3	11	14

Calcoliamo una media mobile a 3 termini con pesi uguali, considerando per ciascun periodo un valore precedente, il valore corrente e un valore successivo:

$$Y^*_{2000} = \frac{1}{3} \cdot 7 + \frac{1}{3} \cdot 4 + \frac{1}{3} \cdot 3 = 4.67$$

$$Y^*_{2001} = \frac{1}{3} \cdot 4 + \frac{1}{3} \cdot 3 + \frac{1}{3} \cdot 11 = 6$$

$$Y^*_{2002} = \frac{1}{3} \cdot 3 + \frac{1}{3} \cdot 11 + \frac{1}{3} \cdot 14 = 9.33$$

Media mobile con l'operatore ritardo

Poiché $Y_{t+i} = B^{-i} Y_t$, possiamo scrivere:

$$Y_t^* = \sum_{i=-m_1}^{m_2} \vartheta_i Y_{t+i} = \sum_{i=-m_1}^{m_2} \vartheta_i B^{-i} Y_t$$

e quindi:

$$Y_t^* = \left[\sum_{i=-m_1}^{m_2} \vartheta_i B^{-i} \right] Y_t$$

Interpretazione

Questa notazione mostra in maniera più evidente come la media mobile sia un **filtro lineare**, ovvero prende la serie originale e la trasforma mediante una combinazione lineare dei suoi valori ritardati e anticipati

Media mobile semplice

Una **media mobile semplice** è una media mobile in cui tutte le osservazioni considerate hanno lo stesso peso.

Esempio a 3 termini

Consideriamo tre valori vicini:

$$Y_{t-1}, \qquad Y_t, \qquad Y_{t+1}$$

Poiché i pesi sono uguali, ciascuno pesa $\frac{1}{3}$. Quindi:

$$Y_t^* = \frac{1}{3}Y_{t-1} + \frac{1}{3}Y_t + \frac{1}{3}Y_{t+1}$$

cioè:

$$Y_t^* = \frac{Y_{t-1} + Y_t + Y_{t+1}}{3}$$

È detta semplice perché tutti i valori nella finestra hanno lo stesso peso.

Media mobile centrata

Una **media mobile centrata** è una media mobile in cui il valore lisciato al tempo t è calcolato usando osservazioni disposte simmetricamente intorno a t.

Esempio a 3 termini:

$$Y_t^* = \frac{Y_{t-1} + Y_t + Y_{t+1}}{3}$$

Interpretazione

Quando la finestra della media mobile è disposta simmetricamente attorno a t, il valore lisciato viene naturalmente attribuito al periodo centrale t.

La centratura è importante perché evita di spostare artificialmente la serie avanti o indietro.

Medie mobili simmetriche

Una media mobile centrata è **simmetrica** se i pesi delle osservazioni equidistanti dal centro sono uguali.
Consideriamo una media mobile:

$$Y_t^* = \sum_{i=-m}^{m} \vartheta_i Y_{t+i}$$

La condizione di simmetria è: $\vartheta_{-i} = \vartheta_i$ per $i = 1, \ldots, m$

Esempio: media mobile simmetrica a 5 termini

$$Y_t^* = \frac{1}{8} Y_{t-2} + \frac{2}{8} Y_{t-1} + \frac{2}{8} Y_t + \frac{2}{8} Y_{t+1} + \frac{1}{8} Y_{t+2}$$

Vettore dei pesi: $\left(\frac{1}{8}, \frac{2}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}\right)$

I pesi agli estremi sono uguali e i pesi interni sono uguali.

Media mobile non centrata

Una **media mobile non centrata** è una media in cui il valore ottenuto non è naturalmente riferito al tempo t.

Retrospettiva

$$Y_t^* = \frac{Y_t + Y_{t-1} + Y_{t-2}}{3}$$

Usa il valore corrente e valori passati.

Ordine pari

$$Y_t^* = \frac{Y_t + Y_{t+1} + Y_{t+2} + Y_{t+3}}{4}$$

Usa valori prima e dopo t, ma non ha un centro unico.

Attenzione

Una media non centrata può introdurre uno spostamento temporale della serie lisciata: nel caso di ordine pari, il centro cade tra due tempi osservati.

Composizione di medie mobili

Le medie mobili godono della **proprietà di composizione**.

Questo significa che è possibile applicare più medie mobili a una stessa serie, anche in sequenza.

Siano M_1 e M_2 due medie mobili. Applicare prima M_1 e poi M_2 equivale a considerare la media mobile composta:

$$M_2(M_1 Y_t) = (M_2 M_1) Y_t$$

In molti casi, l'ordine di applicazione non cambia il risultato:

$$M_2 M_1 = M_1 M_2$$

Perché questa proprietà è utile?

La composizione è utile soprattutto quando partiamo da una media mobile di ordine pari.

Problema

Se la media mobile ha ordine pari, non esiste un valore centrale unico.

Per esempio, una media mobile a 4 termini:

$$Y_t^* = \frac{Y_t + Y_{t+1} + Y_{t+2} + Y_{t+3}}{4}$$

ha centro tra $t+1$ e $t+2$.

Quindi il valore lisciato non è riferito a un tempo osservato preciso.

Per riportare il risultato su un tempo osservato, applichiamo una seconda media mobile a 2 termini.

Esempio: media mobile centrata di ordine pari

Calcoliamo due medie mobili a 4 termini:

$$M_t = \frac{Y_{t-1} + Y_t + Y_{t+1} + Y_{t+2}}{4}$$

Il centro di M_t cade tra t e $t+1$, cioè in $t+0.5$.

$$M_{t+1} = \frac{Y_t + Y_{t+1} + Y_{t+2} + Y_{t+3}}{4}$$

Il centro di M_{t+1} cade tra $t+1$ e $t+2$, cioè in $t+1.5$.

Il valore centrato in $t+1$ si ottiene con una media a 2 termini:

$$Y^*_{t+1} = \frac{M_t + M_{t+1}}{2}$$

Invarianza e conservazione del trend

Quando usiamo una media mobile, vogliamo ottenere una serie più liscia.

Ma attenzione: lisciare non significa modificare tutto.

Idea

Una buona media mobile deve:

- attenuare oscillazioni stagionali e accidentali;
- non deformare il trend di fondo;
- non spostare artificialmente la serie nel tempo.

Se la serie fosse già perfettamente liscia, per esempio una retta,

$$Y_t = a + bt,$$

una buona media mobile dovrebbe restituire la stessa retta.

Invarianza e conservazione del trend: condizioni sui pesi

Consideriamo una media mobile centrata:

$$Y_t^* = \sum_{i=-m}^{m} \vartheta_i Y_{t+i}$$

Per conservare un trend polinomiale fino al grado p, i pesi devono soddisfare:

$$\sum_{i=-m}^{m} \vartheta_i = 1 \qquad \text{(somma dei pesi)}$$

$$\sum_{i=-m}^{m} i^r \vartheta_i = 0, \quad r = 1, \ldots, p \qquad \text{(momenti nulli)}$$

Così la media mobile lascia inalterato un polinomio di grado p.

Invarianza e conservazione del trend: come interpretare le condizioni sui momenti?

Le quantità $\sum_i i^r \vartheta_i$ sono i **momenti dei pesi** rispetto al tempo centrale t.

Interpretazione

- $r = 1$: $\sum_i i\vartheta_i = 0$ conserva una un trend lineare
- $r = 2$: $\sum_i i^2\vartheta_i = 0$ conserva un trend quadratico
- $r = 1, \ldots, p$: $\sum_i i^r\vartheta_i = 0$ conserva un polinomio di grado p

Più alto è il grado del trend, più condizioni imponiamo sui pesi.

Esempio: media mobile centrata e simmetrica

Consideriamo la media mobile centrata a 3 termini:

$$Y_t^* = \frac{1}{3}Y_{t-1} + \frac{1}{3}Y_t + \frac{1}{3}Y_{t+1}$$

Verifichiamo le condizioni:

$$\sum_i \vartheta_i = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$$

$$\sum_i i\vartheta_i = (-1)\frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0$$

Questa media è in grado di conservare un trend lineare.

La stessa media conserva anche un trend quadratico?

Per conservare un trend quadratico servirebbe anche:

$$\sum_i i^2 \vartheta_i = 0$$

Calcoliamo il momento di ordine 2:

$$\sum_i i^2 \vartheta_i = (-1)^2 \frac{1}{3} + 0^2 \frac{1}{3} + 1^2 \frac{1}{3}$$

$$= \frac{1}{3} + 0 + \frac{1}{3} = \frac{2}{3} \neq 0$$

Conclusione

Questa specifica media mobile centrata a 3 termini conserva il trend lineare, ma non conserva un trend quadratico. Se il trend è curvo, questa media mobile può modificarlo leggermente

Effetto della media mobile sul rumore bianco

Quando applichiamo una media mobile alla serie, la applichiamo anche alla componente accidentale, la quale ricordiamo essere considerata un rumore bianco $\varepsilon_t \sim WN(0, \sigma^2)$:

$$\varepsilon_t^* = \sum_{i=-m}^{m} \vartheta_i \varepsilon_{t+i}$$

Il valore atteso della componente filtrata è:

$$E(\varepsilon_t^*) = E\left(\sum_{i=-m}^{m} \vartheta_i \varepsilon_{t+i}\right) = \sum_{i=-m}^{m} \vartheta_i E(\varepsilon_{t+i})$$

Poiché $E(\varepsilon_{t+i}) = 0$, allora:

$$E(\varepsilon_t^*) = 0$$

La media mobile non modifica il valore atteso del rumore, che resta nullo.

Effetto della media mobile sul rumore bianco

Passando invece alla varianza, poiché gli errori sono incorrelati, la varianza è:

$$Var(\varepsilon_t^*) = \sum_{i=-m}^{m} \vartheta_i^2 Var(\varepsilon_{t+i}) = \sigma^2 \sum_{i=-m}^{m} \vartheta_i^2$$

Poniamo $\sigma_*^2 = Var(\varepsilon_t^*)$. Allora:

$$\frac{\sigma_*^2}{\sigma^2} = \sum_{i=-m}^{m} \vartheta_i^2$$

Se $\sum_i \vartheta_i^2 < 1$, allora $\sigma_*^2 < \sigma^2$.

Ma attenzione: effetto Slutsky-Yule

La media mobile può ridurre la variabilità del rumore, ma non è neutrale: può introdurre correlazione tra i valori trasformati.

Perché?

Medie mobili successive condividono alcune osservazioni della serie originale.

Per esempio, con una media mobile a 3 termini:

$$Y_t^* = \frac{Y_{t-1} + Y_t + Y_{t+1}}{3}$$

$$Y_{t+1}^* = \frac{Y_t + Y_{t+1} + Y_{t+2}}{3}$$

I due valori lisciati condividono Y_t e Y_{t+1}.

Questo fenomeno è noto come **effetto Slutsky-Yule**.

Come ridurre il problema?

Non è possibile evitare completamente l'effetto Slutsky-Yule.

Tuttavia, quando si lavora con serie stagionali, si scelgono spesso medie mobili coerenti con il periodo stagionale.

Idea

Se il filtro introduce oscillazioni spurie, è preferibile che abbiano periodo vicino alla stagionalità già presente nella serie.

In questo modo:

- si attenua la componente stagionale;
- si evita di creare oscillazioni spurie con periodicità completamente nuova;
- si ottiene una stima più interpretabile del trend-ciclo.

La media mobile può attenuare la stagionalità?

Supponiamo il seguente modello:

$$Y_t = T_t + S_t + \varepsilon_t$$

Se calcoliamo una media mobile lunga quanto il periodo stagionale s:

$$Y_t^* = \frac{1}{s}\left(Y_t + Y_{t+1} + \cdots + Y_{t+s-1}\right)$$

e se la stagionalità si compensa nel periodo:

$$\sum_{j=1}^{s} S_j = 0$$

allora la parte stagionale viene attenuata o eliminata.

La media mobile lascia emergere il movimento di fondo della serie.

Costruzione della serie destagionalizzata: caso additivo

Supponiamo un modello additivo:

$$Y_t = T_t + C_t + S_t + \varepsilon_t$$

Se la media mobile attenua la stagionalità e riduce il rumore, la serie lisciata può essere usata come stima del trend-ciclo:

$$Y_t^* \approx T_t + C_t$$

Per isolare la componente stagionale, confrontiamo la serie osservata con la serie lisciata e otteniamo l'indice specifico di stagionalità:

$$IS_t = Y_t - Y_t^*$$

L'indice specifico grezzo approssima l'effetto stagionale della singola osservazione.

Dagli indici specifici ai coefficienti stagionali

Per stimare l'effetto stagionale della posizione j, si fa la media degli indici specifici riferiti alla stessa posizione del ciclo, ottenendo i **coefficienti stagionali grezzi**:

$$\widehat{S}_j^* = \frac{1}{N_j} \sum_{t \in \mathcal{T}_j} IS_t, \qquad j = 1, \ldots, s$$

Dove:

$$\mathcal{T}_j = \{\text{osservazioni nella posizione stagionale } j\}$$

Normalizzazione dei coefficienti stagionali

Nel modello additivo, gli effetti stagionali ideali devono compensarsi nel ciclo:

$$\sum_{j=1}^{s} S_j = 0$$

Tuttavia, **noi non osserviamo direttamente i veri effetti stagionali S_j, ma stimiamo dei coefficienti stagionali grezzi $\widehat{S}_j^*$**, che non necessariamente sommano a zero.
Ci calcoliamo quindi una versione corretta dei coefficienti stagionali, sottraendo la loro media:

$$\widehat{S}_j = \widehat{S}_j^* - \frac{1}{s}\sum_{k=1}^{s} \widehat{S}_k^*$$

In questo modo i nuovi coefficienti stagionali soddisfano: $\sum_{j=1}^{s} \widehat{S}_j = 0$.

Costruzione della serie destagionalizzata

Una volta ottenuti i coefficienti stagionali corretti:

$$\widehat{S}_1, \widehat{S}_2, \ldots, \widehat{S}_s$$

la serie destagionalizzata si ottiene sottraendo alla serie osservata l'effetto stagionale corrispondente:

$$Y_t^d = Y_t - \widehat{S}_j$$

dove j è la posizione stagionale dell'osservazione t.

Costruzione della serie destagionalizzata: caso moltiplicativo

Supponiamo un modello moltiplicativo:

$$Y_t = T_t \cdot C_t \cdot S_t \cdot \varepsilon_t$$

Se la media mobile stima il trend-ciclo:

$$Y_t^* \approx T_t \cdot C_t$$

allora l'indice specifico grezzo di stagionalità è:

$$IS_t = \frac{Y_t}{Y_t^*}$$

In fatti

Coefficienti stagionali moltiplicativi

Per il mese j, calcoliamo la media degli indici specifici dello stesso mese in anni diversi:

$$\widehat{S}_j^* = \frac{1}{N} \sum_{T=1}^{N} IS_{T,j}$$

Nel modello moltiplicativo, i coefficienti stagionali ideali devono essere neutri.
Una normalizzazione possibile è:

$$\prod_{j=1}^{s} S_j = 1$$

Quindi:

$$\widehat{S}_j = \frac{\widehat{S}_j^*}{\left(\prod_{k=1}^{s} \widehat{S}_k^*\right)^{1/s}}$$

Destagionalizzazione moltiplicativa

Una volta stimati i coefficienti stagionali moltiplicativi:

$$\widehat{S}_1, \ldots, \widehat{S}_s$$

la serie destagionalizzata è:

$$Y_t^d = \frac{Y_t}{\widehat{S}_j}$$

dove j è il mese o trimestre corrispondente al tempo t.

Interpretazione

Nel modello moltiplicativo, la stagionalità è un fattore: per rimuoverla, dividiamo la serie osservata per il coefficiente stagionale stimato.

Take-away: media mobile

1. Una media mobile è uno strumento di lisciamento.
2. Non è direttamente la serie destagionalizzata.
3. Serve a stimare il trend-ciclo.
4. Può essere centrata o non centrata.
5. Se l'ordine è pari, spesso serve una seconda media per centrare il risultato.
6. Una buona media mobile deve attenuare oscillazioni stagionali e accidentali senza deformare il trend.

Formula chiave

$$Y_t^* = \sum_i \vartheta_i Y_{t+i}$$