



# STATISTICA PER L'ECONOMIA E L'IMPRESA [PROF. GIANNI BETTI]

## Analisi delle Serie Storiche – Approccio Classico e Trend

Federico Bacchi

Dipartimento di Economia Politica e Statistica (DEPS)  
Università degli Studi di Siena

4 maggio 2026

# Contenuti

- 1 Serie Storiche
- 2 Modelli di Composizione
- 3 Trend
- 4 Componente Stagionale

# Contenuti

- 
- 1 Serie Storiche
  - 2 Modelli di Composizione
  - 3 Trend
  - 4 Componente Stagionale

# Dati Econometrici

I **dati econometrici** sono insiemi di osservazioni quantitative utilizzate per stimare, testare e validare *modelli economici* attraverso *metodi statistici*.

Si distinguono tre categorie principali:

1. i **dati sezionali** (o **cross-sectional data**) si riferiscono all'osservazione di **più unità statistiche** in un **unico istante temporale**;
2. le **serie storiche** (o **time series data**) si riferiscono all'osservazione di **un'unica unità statistica** in **più istanti temporali**;
3. i **dati longitudinali** (o **panel data**) si riferiscono all'osservazione di **più unità statistiche** in **più istanti temporali**.

In questa lezione e nelle successive studieremo alcune delle principali tecniche di **analisi delle serie storiche**.

# Modello Stocastico Generale

Si supponga di voler descrivere il **processo generatore** dei dati di una **serie storica**  $y_1, \dots, y_n$ , relativa a una variabile  $Y_t$ .

Un modello stocastico generale è:

$$Y_t = f(t) + u_t \quad (1)$$

- ▶  $f(t)$ : componente **deterministica** (parte sistematica)
- ▶  $u_t$ : componente **stocastica** (variabile casuale)

Le due componenti *non* sono *osservabili separatamente* e devono essere stimate dal **campione**.

## Approccio Classico

Il modello (1) può essere studiato con la lente dell'**approccio classico**.

### Assunzioni

1. Esiste una **legge di evoluzione temporale** del fenomeno, rappresentata da  $f(t)$ .
2. La componente casuale  $u_t$  rappresenta **circostanze accidentali** di entità trascurabile.
3. I residui sono imputati al caso e assimilati a **errori accidentali**.

### Conseguenze

1. La procedura di **stima** si concentra sulla parte deterministica  $f(t)$ .
2. La parte stocastica  $u_t$  è generata da un processo **white noise**: le variabili casuali  $u_1, \dots, u_n$  sono i.i.d. con

$$\begin{aligned} E[u_t] &= 0, & t = 1, \dots, n; \\ \text{Var}[u_t] &= \sigma_u^2, & t = 1, \dots, n. \end{aligned}$$

# Approccio Moderno

In alternativa, il modello (1) può essere studiato con la lente dell'**approccio moderno**.

## Assunzioni

1. La parte deterministica  $f(t)$  è assente o già *eliminata*.
2. La componente stocastica  $u_t$  *non* è dovuta a circostanze accidentali.

## Conseguenze

1. La parte stocastica  $u_t$  è modellata secondo un **processo a componenti correlate**:

$$u_t = g(Y_{t-1}, Y_{t-2}, \dots, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) + \varepsilon_t,$$

dove  $\varepsilon_{t-k}$  sono **shock passati** non osservabili ma concettualmente presenti nel modello.

2. Necessità di tecniche statistiche avanzate (AR, MA, ARMA, ecc.).

# Obiettivi

Indipendentemente dall'approccio adottato, l'utilità dell'analisi delle serie storiche risiede nella possibilità di rispondere a diversi **obiettivi** e **bisogni conoscitivi**.

1. **Descrizione:** identificare pattern, trend, stagionalità e irregolarità nei dati osservati.
2. **Spiegazione:** comprendere i meccanismi che generano la serie e le relazioni dinamiche sottostanti.
3. **Previsione:** stimare valori futuri sulla base dell'evoluzione passata e del modello adottato.
4. **Filtraggio:** isolare componenti specifiche (trend, ciclo, stagionalità, rumore) per analisi mirate.
5. **Controllo:** monitorare il processo nel tempo per individuare anomalie o cambiamenti strutturali.

# Contenuti

- 
- 1 Serie Storiche
  - 2 Modelli di Composizione
  - 3 Trend
  - 4 Componente Stagionale

# Modelli di Composizione

Supponendo di voler adottare l'**approccio classico**, una delle tecniche più utilizzate per analizzare le serie storiche è quella di **modelli di composizione**.

La serie osservata viene rappresentata come combinazione di **componenti strutturali** che vengono trattate come entità **separabili**.

## Obiettivi

1. **Scomporre** la serie nelle componenti strutturali.
2. **Stimare** le componenti strutturali con **tecniche deterministiche**.

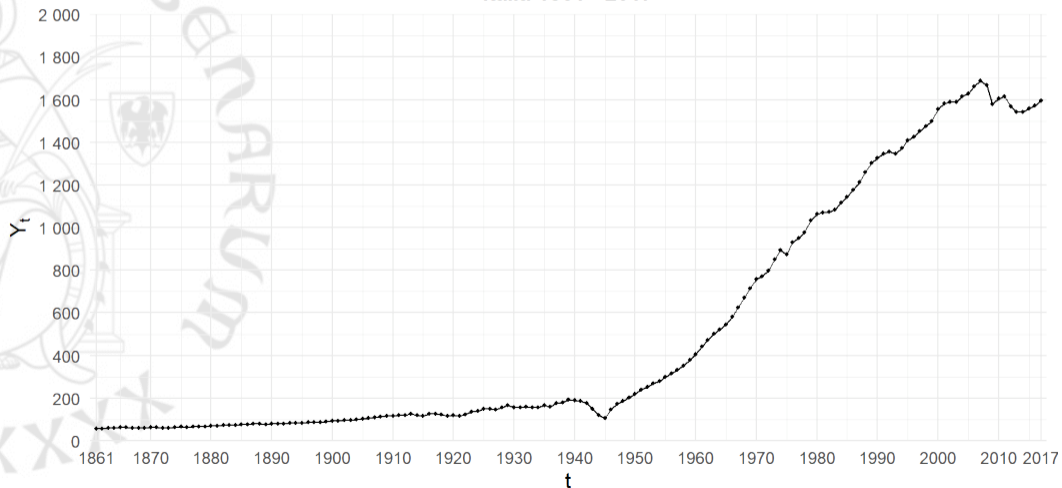
I modelli di composizione forniscono una **lettura descrittiva** e **interpretativa** della struttura della serie.

# Componenti

1. **Trend** ( $T_t$ ): **tendenza di fondo** del fenomeno oggetto di studio, riferita a un lungo periodo di tempo.
  - ▶ Crescita di lungo periodo del PIL italiano (3% annuo).
2. **Ciclo** ( $C_t$ ): **fluttuazioni** attribuibili al succedersi nel fenomeno considerato di *fasi ascendenti* e di *fasi discendenti*.
  - ▶ Crescita del PIL italiano caratterizzata da fasi di espansione (per esempio, il miracolo economico) e fasi di recessione (per esempio, crisi del debito sovrano).
3. **Stagionalità** ( $S_t$ ): movimenti del fenomeno nel corso dell'anno che, per effetto dell'influenza di fattori climatici e sociali, **tendono a ripetersi** in maniera pressoché analoga nel medesimo periodo.
  - ▶ PIL italiano: tutte le grandi fabbriche sono chiuse nel mese di agosto.
4. **Componente accidentale** ( $a_t$ ): tiene conto del fatto che non c'è mai una relazione perfetta tra la variabile sotto osservazione e le diverse componenti e del comportamento **non** perfettamente **prevedibile** degli agenti economici.

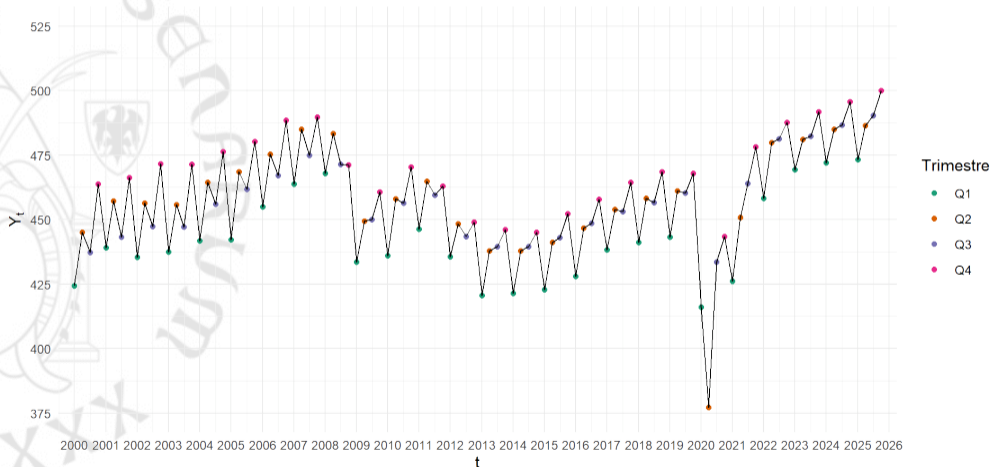
# PIL Italiano

**PIL annuale (miliardi di euro)**  
Italia 1861 - 2017



# PIL Italiano

**PIL trimestrale (miliardi di euro)**  
Italia 2000 - 2025



# Tipi di Composizione

I modelli di composizione possono essere distinti in base al **tipo di composizione** che viene assunta tra le diverse componenti.

## 1. Composizione Additiva

- ▶ **Ipotesi:** le componenti sono tra loro **indipendenti**.
- ▶ **Equazione:**  $Z_t = T_t + C_t + S_t + a_t$

## 2. Composizione Moltiplicativa

- ▶ **Ipotesi:** le componenti **non** sono tra loro **indipendenti**.
- ▶ **Equazione:**  $Z_t = T_t \times C_t \times S_t \times a_t$
- ▶ Si riduce al caso di composizione additiva considerando i logaritmi:

$$Z_t = T_t \times C_t \times S_t \times a_t \quad \Rightarrow \quad \log Z_t = \log T_t + \log C_t + \log S_t + \log a_t$$

## 3. Composizione Mista

- ▶ **Ipotesi:** **coesistenza** di componenti indipendenti e non indipendenti.
- ▶ **Equazione:**  $Z_t = T_t \times S_t + C_t + a_t$

# Modelli di Composizione e Modello Stocastico Generale

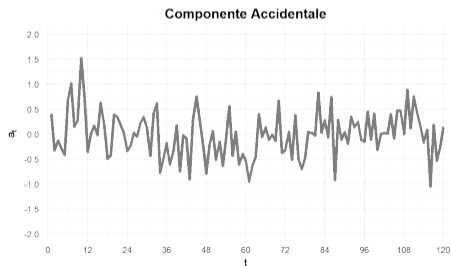
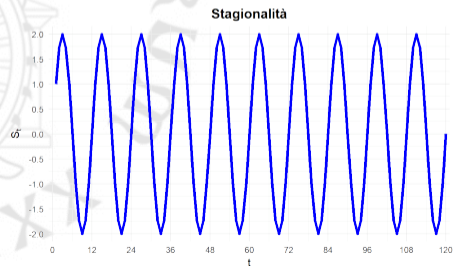
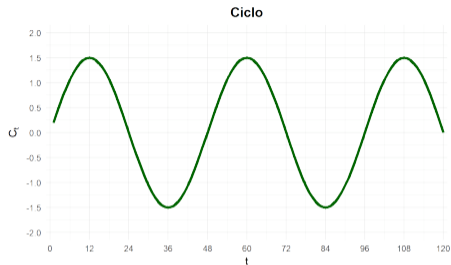
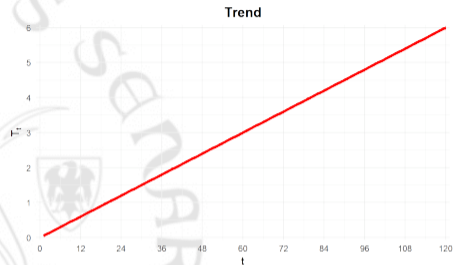
Si noti come sia stato introdotto un **cambio di notazione**: la variabile oggetto di studio è stata indicata nei modelli di composizione con  $Z_t$ .

Alla base di questa scelta, apparentemente stilistica, c'è una **motivazione formale**:

- ▶  $Y_t$  viene utilizzata nel modello stocastico generale ed è definita come la somma di una **componente deterministica**  $f(t)$  e una **componente stocastica**  $u_t$ ;
- ▶  $Z_t$  viene utilizzata nei modelli di composizione ed è definita come combinazione di **componenti strutturali di natura deterministica**  $(T_t, C_t, S_t, a_t)$ .

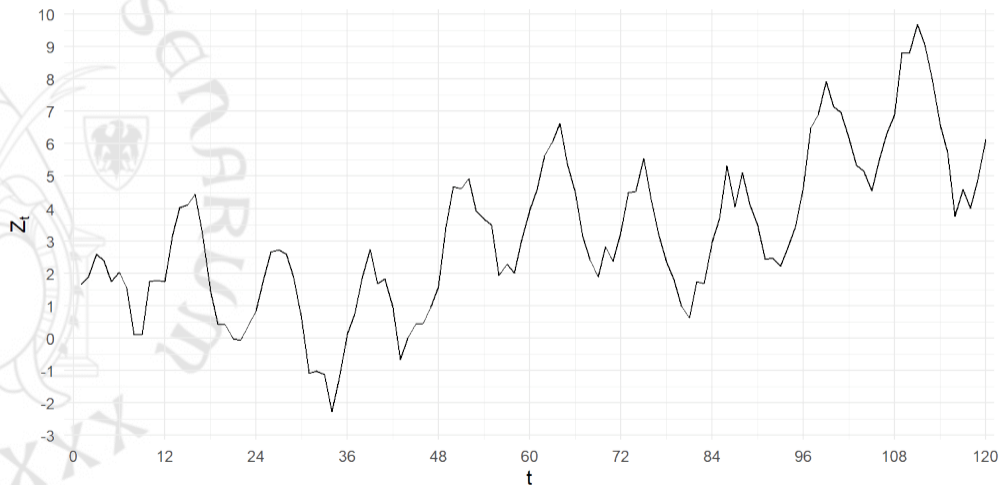
In sintesi, l'utilizzo di  $Z_t$  deriva dalla volontà di rendere esplicito che *non* si sta considerando una vera e propria componente stocastica.

# Serie (mensile) a componenti additive: Esempio

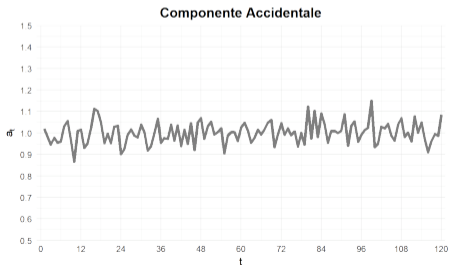
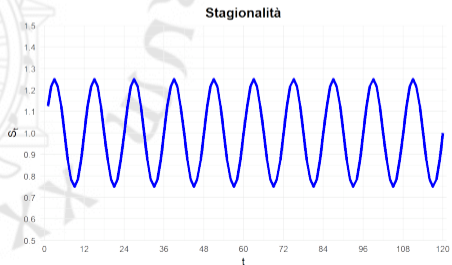
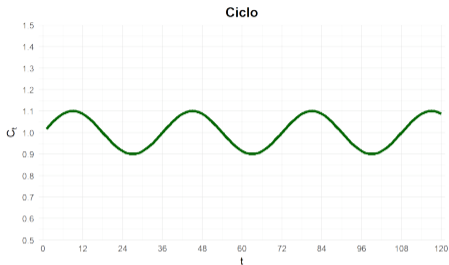
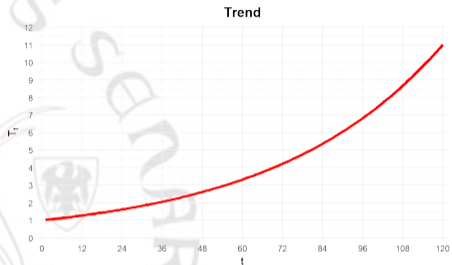


## Serie (mensile) a componenti additive: Esempio

$$Z_t = T_t + C_t + S_t + a_t$$

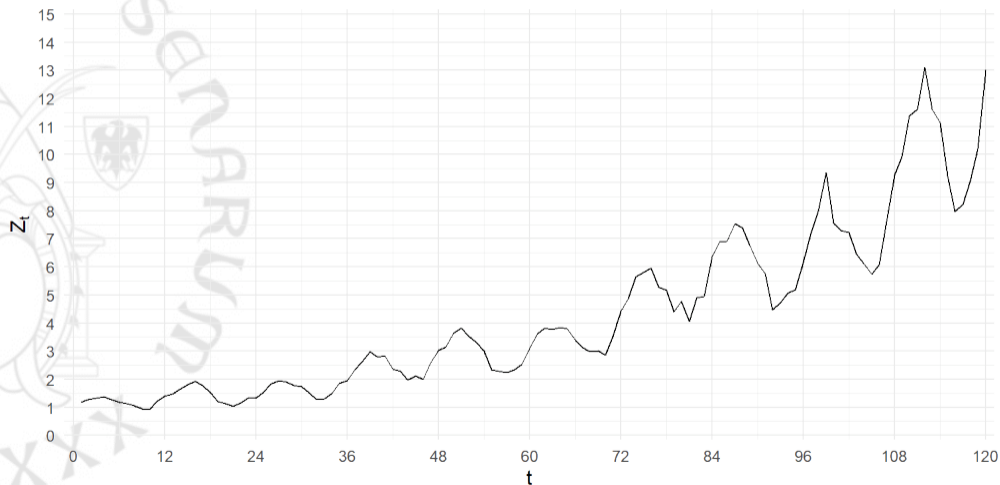


# Serie (mensile) a componenti moltiplicative: Esempio



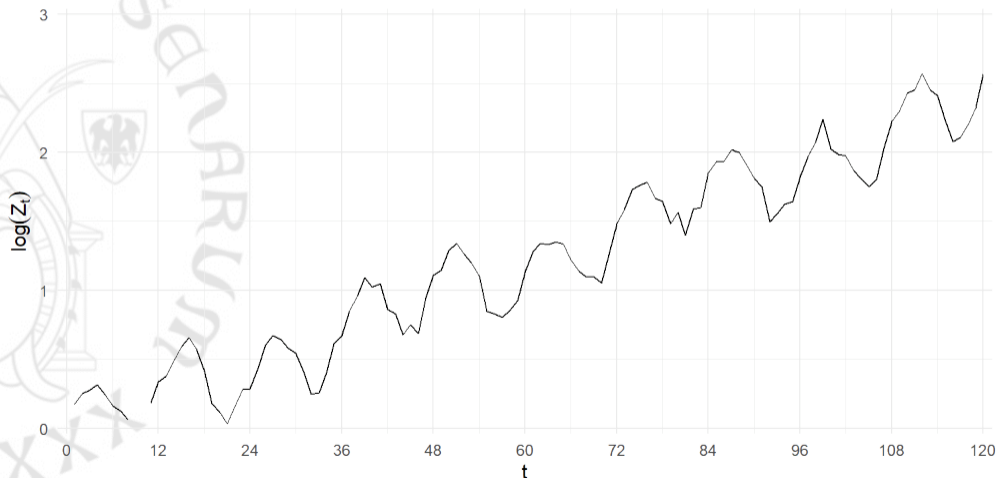
## Serie (mensile) a componenti moltiplicative: Esempio

$$Z_t = T_t \times C_t \times S_t \times a_t$$



## Serie (mensile) a componenti moltiplicative: Esempio

$$\log(Z_t) = \log(T_t) + \log(C_t) + \log(S_t) + \log(a_t)$$



# Pregi e difetti del Modello di Composizione

## Pregi

- ▶ **Semplicità:** struttura intuitiva e facilmente interpretabile.
- ▶ **Applicabile anche a serie corte:** non richiede grandi quantità di dati.
- ▶ **Prima approssimazione:** utile come punto di partenza per analisi più complesse.

## Difetti

- ▶ **Pluralità di soluzioni:** diverse decomposizioni possibili per la stessa serie.
- ▶ **Assunzione modellistica rigida:** ipotesi di indipendenza o forma fissa delle componenti.
- ▶ **Visione settorializzata:** analizza le componenti separatamente, trascurando le interazioni dinamiche.

# Contenuti

- 1 Serie Storiche
- 2 Modelli di Composizione
- 3 Trend
  - Trend Lineare
  - Trend Non Lineare
- 4 Componente Stagionale

# Trend

Il **trend** di una serie storica è stato definito come la tendenza di fondo del fenomeno nel lungo periodo.

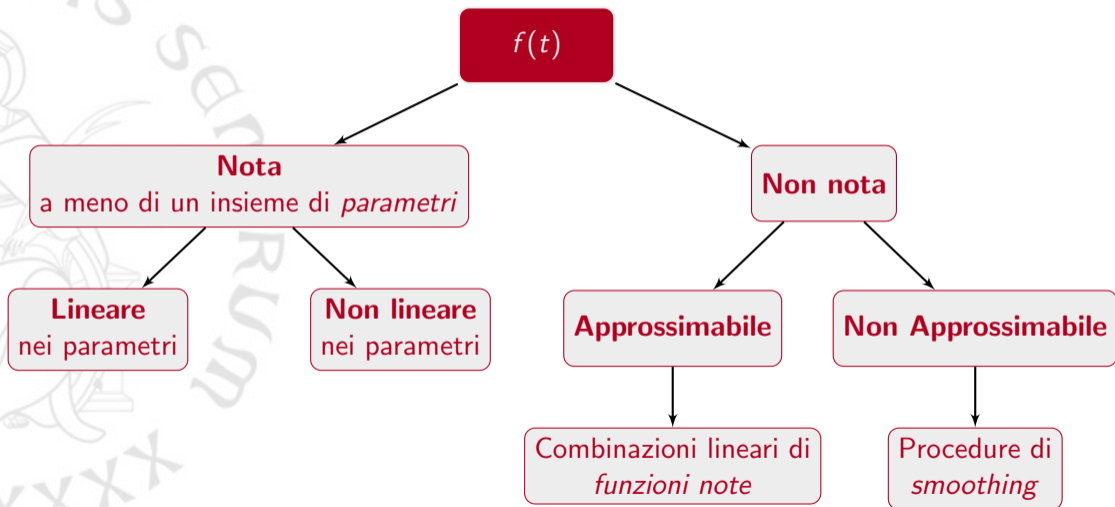
- ▶ Si caratterizza per una **evoluzione lenta** e **regolare** nel tempo.
- ▶ In genere, è rappresentabile attraverso una **funzione del tempo** da stimare.

Si consideri una serie storica  $\{y_t\}_{t=1}^n$  per la quale si ipotizza un modello stocastico:

$$Y_t = f(t) + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma_\varepsilon^2) \quad (2)$$

Si supponga, inoltre, che la parte sistematica  $f(t)$  sia composta **solo** dal **trend**.

# Come rappresentare $f(t)$ ?



## Modelli con Trend Lineare

Si supponga che la parte deterministica del modello  $f(t)$  sia costituita solamente dal trend e che sia nota a meno di un vettore di parametri  $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_q)^T$ .

I modelli con **trend lineare** si basano sull'ipotesi che  $f(t)$  sia una funzione lineare rispetto a  $\alpha_0, \alpha_1, \dots, \alpha_q$ .

La stima di tali modelli si ottiene attraverso le procedure derivate dal **modello di regressione lineare**.

Di seguito, si considereranno due modelli con trend lineare (o *linearizzabile*) nei parametri:

1. il modello con **trend polinomiale**;
2. il modello con **trend esponenziale**.

## Trend Polinomiale

Un esempio di modello con trend lineare nei parametri è il **trend polinomiale**:

$$f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_q t^q$$

Di conseguenza, l'equazione stocastica (2) diventa a tutti gli effetti un modello di regressione lineare:

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_q t^q + \varepsilon_t, \quad t = 1, \dots, n,$$

esprimibile anche in **forma matriciale**:

$$\begin{matrix} \mathbf{y} \\ n \times 1 \end{matrix} = \begin{matrix} \mathbf{P} \\ n \times q \end{matrix} \begin{matrix} \boldsymbol{\alpha} \\ q \times 1 \end{matrix} + \begin{matrix} \boldsymbol{\varepsilon} \\ n \times 1 \end{matrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1^2 & \dots & 1^q \\ 1 & 2 & 2^2 & \dots & 2^q \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & n & n^2 & \dots & n^q \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

## Stima dei parametri

Di conseguenza, le stime dei parametri si ottengono attraverso il **metodo dei minimi quadrati ordinari** (OLS) attraverso la formula:

$$\hat{\alpha} = \left( \mathbf{P}^T \mathbf{P} \right)^{-1} \mathbf{P}^T \mathbf{y}$$

Il polinomio stimato attraverso il metodo OLS

$$y_t = \hat{\alpha}_0 + \hat{\alpha}_1 t + \hat{\alpha}_2 t^2 + \cdots + \hat{\alpha}_q t^q + e_t, \quad t = 1, \dots, n$$

può essere usato per **fini interpolativi**, mentre serve cautela per fini previsivi.

## Ordine del Polinomio

Una scelta cruciale nella definizione del modello con trend polinomiale è quella relativa all'**ordine**  $q$ .

L'ordine  $q$  ottimale dipende dal **comportamento di fondo** della serie storica.

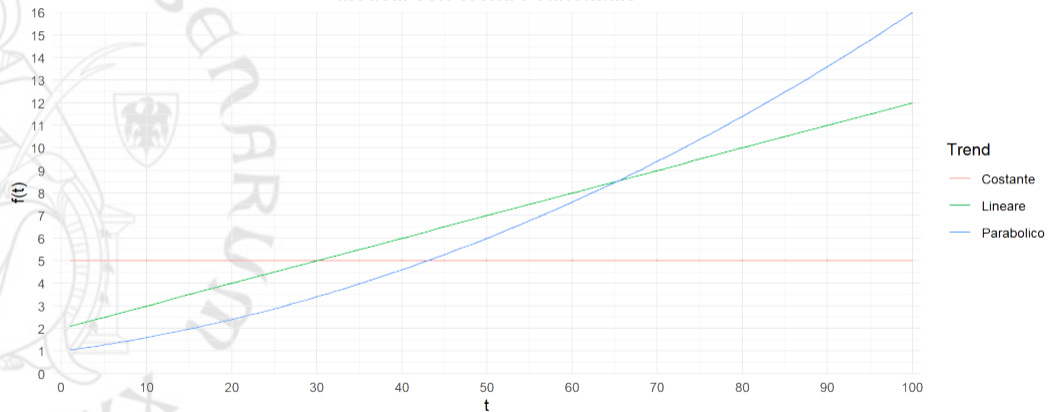
I casi particolari più utilizzati sono i seguenti:

- ▶ **trend costante:**  $q = 0 \Rightarrow Y_t = \alpha_0 + \varepsilon_t$ ;
- ▶ **trend lineare:**  $q = 1 \Rightarrow Y_t = \alpha_0 + \alpha_1 t + \varepsilon_t$ ;
- ▶ **trend parabolico:**  $q = 2 \Rightarrow Y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \varepsilon_t$ .

In generale, è preferibile la scelta di un ordine  $q$  abbastanza **piccolo** per evitare di perdere troppi gradi di libertà.

# Modelli con Trend Polinomiale

## Modelli con Trend Polinomiale



## Criterio delle Differenze Successive I

Tuttavia, non è detto che la scelta di  $q$  sia intuitiva. Dunque, è necessario adottare il **criterio delle differenze successive**.

1. Si indichi con  $\mathbf{B}$  l'**operatore ritardo** che trasforma la serie  $y_t$  in

$$y_{t-1} = \mathbf{B} y_t,$$

da cui segue:

$$\mathbf{B}^h y_t = y_{t-h};$$

2. Si definisca **operatore identità** l'operatore ritardo  $\mathbf{B}^0 = \mathbf{I}$  che lascia immutata la serie:

$$\mathbf{B}^0 y_t = \mathbf{I} y_t = y_t;$$

3. Si introduca l'**operatore differenza prima all'indietro**  $(\mathbf{I} - \mathbf{B})$  per il quale

$$(\mathbf{I} - \mathbf{B}) y_t = y_t - y_{t-1};$$

## Criterio delle Differenze Successive II

4. Si calcolino le differenze successive della serie;
5. L'operazione si arresta quando, per un determinato valore  $r$ , la serie  $(I - B)^r y_t$  è approssimativamente **costante**.
6. L'ordine ottimale del polinomio è  $q \equiv r$ .

### Esempio

Si supponga che il trend osservato sia lineare, ovvero  $y_t = \alpha_0 + \alpha_1 t$ .

Allora:

$$\begin{aligned}(I - B)y_t &= y_t - y_{t-1} = \\ &= \alpha_0 + \alpha_1 t - (\alpha_0 + \alpha_1(t - 1)) = \\ &= \alpha_0 + \alpha_1 t - \alpha_0 - \alpha_1 t + \alpha_1 = \\ &= \alpha_1\end{aligned}$$

$\alpha_1$  è un valore costante ed è stato ottenuto per  $r = 1$ ; dunque, l'ordine ottimale del trend polinomiale è  $q = 1$ .

## Modelli con Trend Esponenziale

Si supponga di voler studiare una serie storica relativa a un fenomeno caratterizzato da una **crescita** nel tempo **esplosiva**.

In questo contesto, un tipo di modelli più idoneo rispetto a quelli con trend polinomiale sono i **modelli con trend esponenziale**:

$$f(t) = \alpha_0 e^{\alpha_1 t} \quad (3)$$

- ▶ Se  $\alpha_0 > 0$ , l'andamento del trend dipende da  $\alpha_1$ .
- ▶ Il **tasso di crescita** relativo è **costante**, in quanto:

$$\frac{\frac{df(t)}{dt}}{f(t)} = \alpha_1$$

## Linearizzazione del Modello

Si noti come il modello con trend esponenziale **non** sia **lineare** nei parametri, in quanto  $\alpha_1$  è un esponente.

Tuttavia, è possibile ottenere la **linearizzazione** del modello rispetto ai parametri con l'aiuto della trasformata logaritmica di (3):

$$\log \{f(t)\} = \log \alpha_0 + \alpha_1 t$$

La linearizzazione rispetto ai parametri  $\alpha_0$  e  $\alpha_1$  può, dunque, essere ottenuta tramite l'adozione di un **modello stocastico moltiplicativo**

$$y_t = f(t) \cdot \varepsilon_t, \quad t = 1, \dots, n,$$

mentre *non* può essere ottenuta dalla specificazione additiva

$$y_t = f(t) + \varepsilon_t, \quad t = 1, \dots, n.$$

## Stima del Modello

La stima dei parametri  $\alpha_0$  e  $\alpha_1$  è possibile con la specificazione moltiplicativa:

$$y_t = \alpha_0 e^{\alpha_1 t} \cdot \varepsilon_t, \quad t = 1, \dots, n,$$

da cui si ricava

$$\log y_t = \log \alpha_0 + \alpha_1 t + \log \varepsilon_t, \quad t = 1, \dots, n;$$

$$\log y_t = \alpha_0^* + \alpha_1 t + \varepsilon_t^*, \quad t = 1, \dots, n;$$

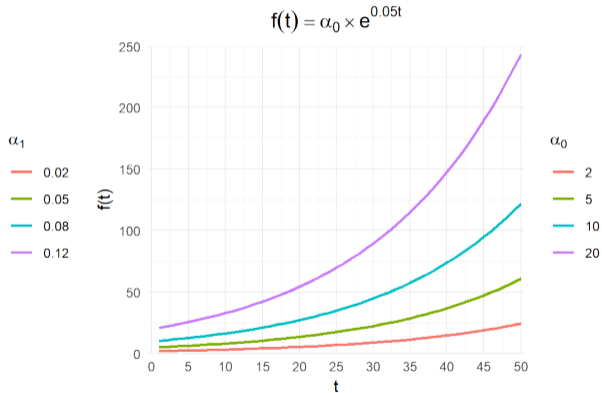
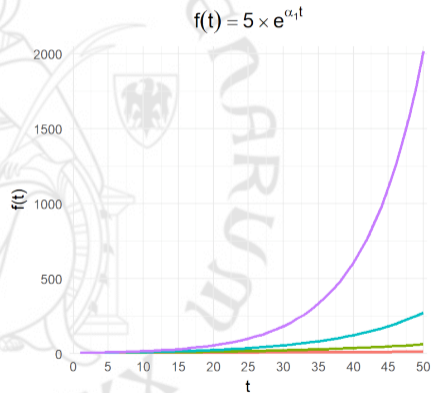
dove i parametri  $\alpha_0^*$ ,  $\alpha_1$  possono essere stimati tramite il **metodo OLS**.

### Problema

- ▶ La stima di  $y_t$  è **distorta**, in quanto  $E(y_t) \neq f(t)$ .

# Modelli con Trend Esponenziale

## Trend Esponenziale



## Trend Non Lineare

Si supponga, come in precedenza, che la parte deterministica del modello  $f(t)$  sia costituita solamente dal trend e che sia nota a meno di un vettore di parametri  $\theta$ .

I modelli con **trend non lineare** nei parametri si basano sull'ipotesi che  $f(t)$  sia una funzione **non lineare** e **non linearizzabile** rispetto a  $\theta$ .

In altre parole, il trend  $f(t)$  non può essere ricondotto attraverso alcuna trasformazione al trend polinomiale.

I modelli con trend non lineare nei parametri sono buone rappresentazioni di fenomeni caratterizzati da **crescite molto accelerate**.

## Curva Esponenziale Modificata

Un primo esempio di modello con trend non lineare nei parametri e il modello con **curva di crescita esponenziale modificata**:

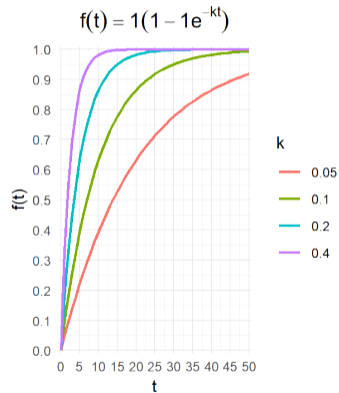
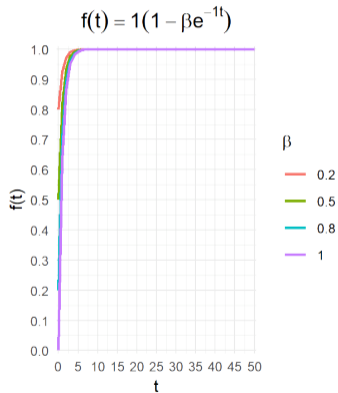
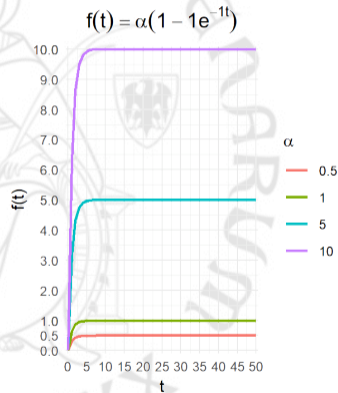
$$f(t) = \alpha \left( 1 - \beta e^{-kt} \right),$$

dove:

- ▶  $\alpha > 0$  determina il **valore limite** della crescita;
- ▶  $\beta > 0$  determina, insieme ad  $\alpha$ , l'**intersezione** della curva **con l'asse verticale**;
- ▶  $k > 0$  è la **costante di proporzionalità**.

# Curva Esponenziale Modificata

## Curva Esponenziale Modificata



## Formulazione alternativa

Il modello con curva di crescita esponenziale modificata può essere riscritto anche come:

$$f(t) = \beta_0 + \beta_1 e^{\beta_2^* t},$$

dove, con riferimento alla formula originaria:

- ▶  $\beta_0 = \alpha > 0$ ;
- ▶  $\beta_1 = -\alpha \beta < 0$ ;
- ▶  $\beta_2^* = -k < 0$ .

Questa formulazione rende maggiormente evidente la differenza con il trend esponenziale semplice, dove è assente l'**asintoto superiore**  $\beta_0$ .

La presenza di  $\beta_0 > 0$  rende **impossibile** la **linearizzazione** tramite trasformazione logaritmica.

Tuttavia, la differenza  $\beta_0 + f(t)$  è, a tutti gli effetti, un'esponenziale semplice e, dunque, linearizzabile.

- ▶ In contesti in cui il valore dell'**asintoto** è **noto** (fissato dall'esterno) è possibile ricavare l'esponenziale semplice e stimare i parametri con **metodo OLS**.

## Curva Logistica

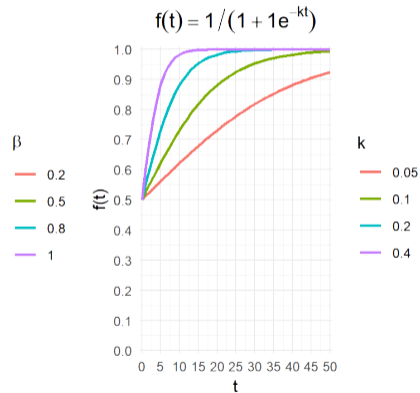
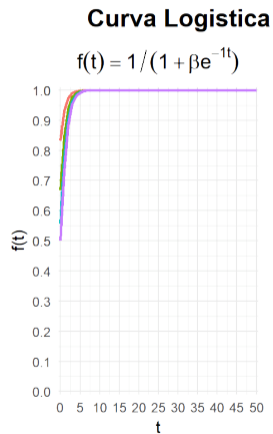
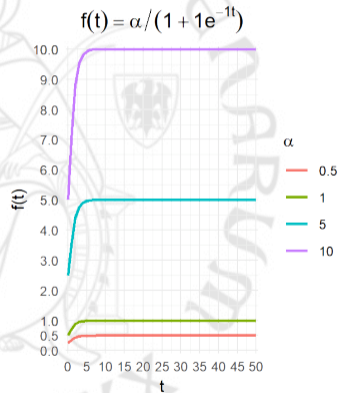
Un secondo esempio di modello con trend non lineare nei parametri e il modello con **curva di crescita logistica**:

$$f(t) = \frac{\alpha}{1 + \beta e^{-kt}},$$

dove:

- ▶  $\alpha > 0$  determina la **scala** della funzione di crescita;
- ▶  $\beta > 0$  determina, insieme ad  $\alpha$ , l'**intersezione** della curva **con l'asse verticale**;
- ▶  $k > 0$  determina l'**inclinazione** della funzione.

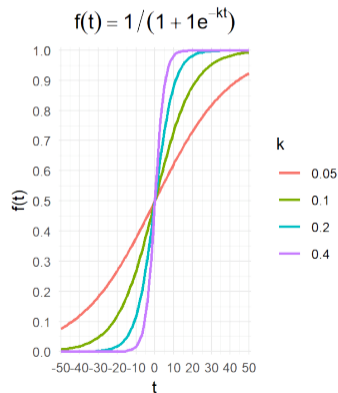
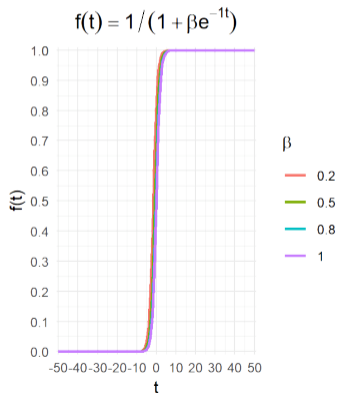
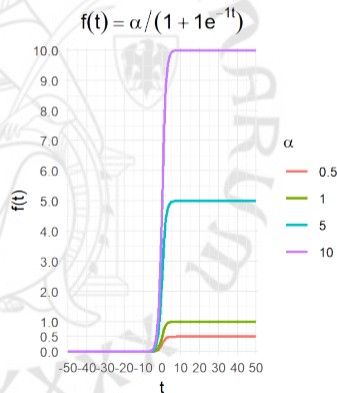
# Curva Logistica



## Curva Logistica Completa

Di seguito, vengono presentati gli stessi grafici con gli stessi parametri visti in precedenza nel caso in cui fosse possibile prendere in considerazione valori di  $t < 0$ .

### Curva Logistica Completa



## Curva Logistica come Esponenziale Modificata

Il reciproco del modello con curva di crescita logistica può essere ricondotto al modello con curva di crescita esponenziale modificata:

$$\begin{aligned}\frac{1}{f(t)} &= \frac{1 + \beta e^{-kt}}{\alpha} = \\ &= \frac{1}{\alpha} + \frac{\beta}{\alpha} e^{-kt} = \\ &= \alpha^* + \beta^* e^{-kt} = f^*(t)\end{aligned}$$

Il modello con trend  $f^*(t)$  è, a tutti gli effetti, un modello con curva di crescita esponenziale modificata nota a meno dei parametri  $\theta = (\alpha^*, \beta^*, k)$ .

## Curva di Gompertz

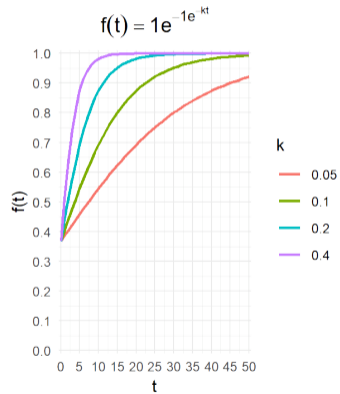
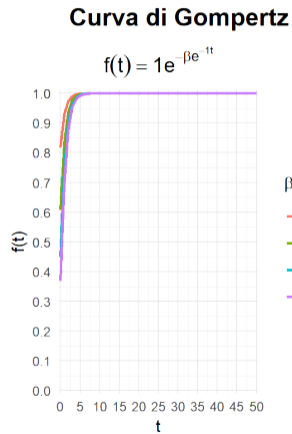
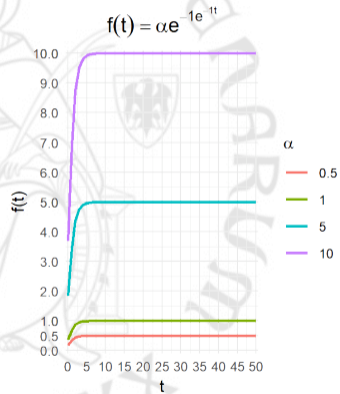
Un terzo esempio di modello con trend non lineare nei parametri e il modello con **curva di crescita di Gompertz**:

$$f(t) = \alpha \cdot e^{-\beta e^{-kt}},$$

dove i parametri  $\theta = (\alpha, \beta, k)$  hanno la stessa interpretazione di quelli nel modello con curva di crescita logistica.

La forma di  $f(t)$  è simile a quella della logistica, sebbene **non** sia **simmetrica** rispetto al punto di flesso  $t = 0$ .

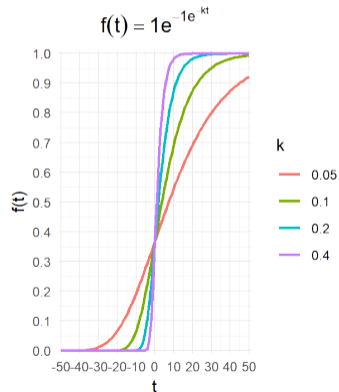
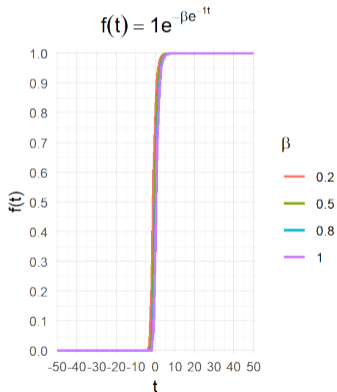
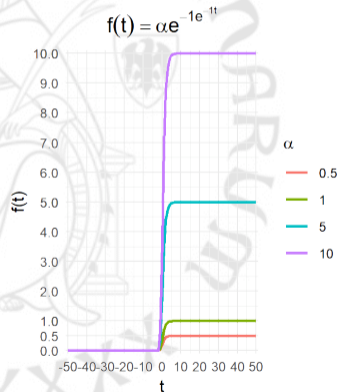
# Curva di Gompertz



## Curva di Gompertz Completa

Di seguito, vengono presentati gli stessi grafici con gli stessi parametri visti in precedenza nel caso in cui fosse possibile prendere in considerazione valori di  $t < 0$ .

### Curva di Gompertz completa



# Curva di Gompertz come Esponenziale Modificata

La **trasformazione logaritmica** della curva di Gompertz è un'**esponenziale modificata**:

$$\begin{aligned}\log \{f(t)\} &= \log \left\{ \alpha \cdot e^{-\beta e^{-kt}} \right\} = \\ &= \log \alpha + \log \left\{ e^{-\beta e^{-kt}} \right\} = \\ &= \log \alpha - \beta e^{-kt}\end{aligned}$$

Ponendo:

$$\beta_0 = \log \alpha, \quad \beta_1 = -\beta, \quad \beta_2^* = -k$$

si ottiene per  $\log \{f(t)\}$  la formulazione alternativa dell'esponenziale modificata.

# Contenuti

- 
- 1 Serie Storiche
  - 2 Modelli di Composizione
  - 3 Trend
  - 4 Componente Stagionale

## Componente Stagionale

La **componente stagionale** di una serie storica si riferisce ai movimenti del fenomeno nel corso dell'anno che tendono a ripetersi in maniera pressoché analoga nel medesimo periodo.

- ▶ Si caratterizza per una **periodicità**.
- ▶ È rappresentabile attraverso una **funzione del tempo** da stimare.

Si consideri una serie storica  $\{y_t\}_{t=1}^n$  per la quale si ipotizza un modello stocastico:

$$Y_t = g(t) + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma_\varepsilon^2) \quad (4)$$

Si supponga, inoltre, che la parte sistematica  $g(t)$  sia composta **solo** dalla **stagionalità**:

$$Y_t = S_t + \varepsilon_t, \quad t = 1, \dots, n.$$

## Come rappresentare $g(t)$ ?

La componente stagionale  $g(t)$  viene rappresentata mediante una **funzione periodica**.

### **Funzione periodica**

Funzioni il cui valore all'istante  $t$  si riproduce esattamente ad intervalli costanti, la cui lunghezza  $s$  costituisce il periodo:

$$g(t) = g(t + s) = g(t + 2s) = \dots = g(t + ks), \quad k \in \mathbb{N}.$$

Casi particolari nelle serie storiche:

- ▶  $s = 4$  è una **serie trimestrale**;
- ▶  $s = 12$  è una **serie mensile**.

# Funzioni trigonometriche

Una scelta comune per modellare la componente stagionale  $S_t$  è quella di definire  $g(t)$  mediante **funzioni trigonometriche**:

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

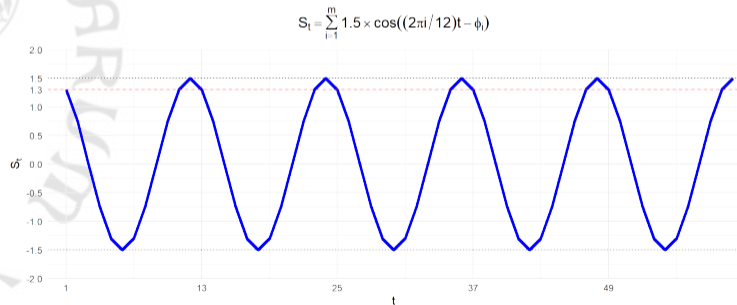
In questo contesto, la componente stagionale è data dalla **somma** di  $m$  **armoniche** con:

- ▶ periodo  $\frac{s}{i}$ ;
- ▶ frequenza angolare  $\omega_i = \frac{2\pi i}{s}$ ;
- ▶ ampiezza  $A_i$ ;
- ▶ angolo di fase  $\phi_i$ .

## Esempio: stagionalità mensile

Qualora si consideri una stagionalità mensile ( $s = 12$ ), si ha che ogni armonica completa il ciclo in 12 periodi.

Per esempio, la prima armonica ( $i = 1$ ) completerà il ciclo ad ogni istante temporale  $t = 12k + 1$ .



## Integrazione con Trend Polinomiale

La componente stagionale può essere integrata con il trend ottenendo il seguente modello stocastico:

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

dove  $f(t) = T_t$  è la componente deterministica dovuta al **trend** e  $g(t) = S_t$  è la componente deterministica dovuta alla **stagionalità**.

Supponendo l'esistenza di un **trend polinomiale** di grado  $q$  e una **componente stagionale trigonometrica** si ha:

$$y_t = \alpha_0 + \alpha_1 t + \cdots + \alpha_q t^q + \sum_{i=1}^m A_i \cos\left(\frac{2\pi i}{s} t - \phi_i\right) + \varepsilon_t, \quad t = 1, \dots, n.$$

Sfruttando l'**identità trigonometrica**, si può riscrivere il modello integrato come:

$$y_t = \alpha_0 + \alpha_1 t + \cdots + \alpha_q t^q + \sum_{i=1}^m \left( \beta_{i1} \cos(\omega_i t) + \beta_{i2} \sin(\omega_i t) \right) + \varepsilon_t, \quad t = 1, \dots, n,$$

con  $\beta_{i1} = A_i \cos(\phi_i)$  e  $\beta_{i2} = A_i \sin(\phi_i)$  che può essere stimato tramite **metodo OLS**.

## Software

[1] gretl: <https://cran.r-project.org/manuals.html>.

[2] The R Project for Statistical Computing: <https://www.r-project.org/>.