

1 – Introduzione

In molte situazioni emerge la necessità di pianificare le azioni future, ovvero di ricorrere a previsioni volte a risolvere problemi di natura decisionale.

decisioni di investimento, decisioni sulla gestione delle scorte, decisioni sullo scheduling della produzione

In ambito economico le previsioni rispondono a finalità diverse:

- in campo microeconomico (aziendale) diventano necessarie per la pianificazione e controllo con aspetti differenziati a seconda dell'area coinvolta.

organizzazione della produzione^a, delle scorte, del personale, acquisizione delle risorse, determinazione dell'ammontare delle risorse necessarie

^a *production scheduling*

Analisi delle Serie storiche

- in sede macroeconomica sono indispensabili per la pianificazione economica e per gli eventuali interventi di politica economica e sociale

fatturato, contabilità nazionale

Problema 1. Come effettuare una previsione?

Partiamo da alcune importanti considerazioni:

- il futuro è diverso dal passato;
- il futuro è più eterogeneo del passato;
- la previsione con certezza si discosterà dalla realtà.

Una possibile soluzione al problema (ma non l'unica) è misurare il fenomeno, che ci interessa, in un periodo che va da un tempo passato fino ad oggi; comprendere se il fenomeno in tale periodo ha un

andamento regolare. Se riusciamo a comprendere la legge o la procedura che ci permette di ottenere tutti o la maggior parte dei dati rilevati, allora, a meno di fattori endogeni aleatori, potremo effettuare una previsione sui valori futuri.

In tali termini la previsione dipenderà da due fattori:

- il numero dei dati in nostro possesso
- l'orizzonte temporale a cui è riferita^b.

Distinguiamo allora due metodi:

- **Metodo qualitativo o congetturale** imperniato sulla scarsità dei dati e adeguate conoscenze di tipo qualitativo che portano in particolare ad adottare

^b cioè se si tratta di una previsione *a breve, a medio o a lungo termine*.

Analisi delle Serie storiche

a breve periodo: i metodi:

- *Esplorativi* in cui si richiede l'opinione di esperti sull'evoluzione futura del fenomeno;

a lungo periodo: i metodi

- *Normativi* in cui, fissati una serie di vincoli temporali e/o di risorse, si eseguono previsioni su come realizzare un determinato obiettivo.

- **Metodo quantitativo o statistico** fondato su una consistenza di dati tale da procedere con:

a breve periodo: i metodi:

- *Endogeni o Extrapolativi:*
 - * *Analisi delle serie storiche*
analisi classica, analisi moderna

Analisi delle Serie storiche

– *Esogeni*:

- * metodi parametrici

 - regressione multipla, lineare e non lineare

- * metodi non parametrici

 - alberi di regressione

a lungo termine: i metodi:

- Dhelphi

- Brainstorming

In Tab. 1 è riportato uno schema riassuntivo delle tecniche di previsione dalla spiegazione del passato.

La tecnica prescelta in tale sede è l'analisi delle serie storiche.

Analisi delle Serie storiche

	<i>Metodi qualitativi (o congetturali)</i>	<i>Metodi quantitativi (o statistico)</i>
<i>a breve termine</i>	esplorativo	endogeni o estrapolativi (analisi delle serie storiche) esogeni (parametrici e non parametrici)
<i>a lungo termine</i>	normativo	Delphi Brainstorming

Tabella 1: I metodi di previsione del futuro conoscendo il passato

2 – Le serie storiche

Dato un fenomeno quantitativo X classificato rispetto alla modalità qualitativa *tempo*, la successione ordinata nel tempo dei suoi valori

$$x_1, \quad x_2, \quad \dots, \quad x_t, \quad \dots,$$

è detta **serie storica** o **serie temporale** di X .

Le serie storiche possono essere classificate in vari modi:

per l'unità di tempo si distinguono in:

- **serie posizionali** se t è un'*istante* come ore, giorni, settimane, mesi, ...

Il traffico telefonico giornaliero, il NIC trimestrale, il PIL annuale, il censimento della popolazione in un decennio, ...

Analisi delle Serie storiche

- **serie di flusso** se t è un *intervallo* come il tempo di esecuzione di un'operazione, ...
la produzione di un'azienda, il finanziamento di un istituto di credito ad un'azienda,...

per la natura del fenomeno si distinguono in:

- **serie economiche** = andamento temporale dei fenomeni economici rilevanti di un sistema economico;
quotazione giornaliera dei titoli in borsa, rilevazione delle forze lavoro, vendite mensili di un dato prodotto, dati trimestrali del PIL di un paese,...
- **serie storiche demografiche** = evoluzione della popolazione rispetto alle caratteristiche strutturali;
densità della popolazione straniera, tassi annuali di natalità - mortalità - flusso migratorio, ,...

Analisi delle Serie storiche

- **serie storiche fisiche** = successione temporale di fenomeni naturali soggetti a mutamenti od alterazioni;
livello di inquinamento atmosferico, livello giornaliero delle precipitazioni atmosferiche in una data area geografica,...
- **serie storiche binarie** = sequenze temporali di fenomeni dicotomici di vario genere.
andamento di un dato titolo di borsa (+ o -), andamento del PIL (crescita o decriscita),...

per il numero dei fenomeni si distinguono in:

- **univariate** se riportano l'andamento temporale di un solo fenomeno;
- **multivariate** se riportano contestualmente le evoluzioni temporali di almeno due fenomeni.

per il numero di osservazioni si distinguono in:

- **discrete** se il numero delle rilevazioni è discreto;
la vendita di un prodotto nei mesi dell'anno, gli aumenti salariali negli anni, ...
- **continue** quando le rilevazioni coprono tutti gli istanti del periodo di osservazione del fenomeno.
l'andamento istantaneo di un titolo di borsa, il battito cardiaco rilevato dall'holder,...

Indicheremo con:

- $[1, T]$ l'arco temporale della serie;
- x_t con $t \in [1, T]$ la serie storica;
- x_t il valore di X al tempo t

e ci occuperemo di *serie storiche posizionali, economiche, univariate e discrete*.

3 – Analisi delle serie storiche

L'analisi delle serie storiche è volta a scoprire le regolarità insite nell'andamento del fenomeno a fini descrittivi o di previsione.

In dettaglio occorre, sulla base del comportamento passato di X nel periodo temporale $[1, T]$, individuare le opportune funzioni analitiche che concorreranno nel costituire il **modello generatore** della serie^a che descrive il fenomeno. Inoltre per mezzo del modello generatore

^a Il modello deve essere tale che sostituendo in esso $1 \leq t \leq T$ troviamo esattamente il valore rilevato x_t

saremo in grado di effettuare previsioni a breve termine^b.

Problema 2. Quali sono le modalità che ci permettono di individuare il modello generatore della serie?

Esistono diversi approcci^c per determinare il modello generatore:

- *analisi grafica* basata sullo studio dell'andamento grafico della serie storica;
- *analisi classica* che considera il modello decompositivo ovvero formato da componenti tra cui le deterministiche sono le più rilevanti;

^b Le ragioni di un orizzonte temporale breve sono dovute a: il futuro è più eterogeneo del passato ed il modello generatore è funzione del passato.

^c Oltre a quelle elencate ricordiamo: l'analisi spettrale, l'analisi armonica, l'analisi periodale e la cimanalisi.

Analisi delle Serie storiche

- *analisi moderna* volta a rintracciare il processo stocastico che ha generato la serie storica.

Prima di scegliere l'approccio è conveniente “preparare la serie storica” ovvero effettuare gli opportuni aggiustamenti secondo un preciso ordine:

1. verificare che tutte le misurazioni del fenomeno X siano espresse nella medesima unità di misura;
2. controllare che l'arco temporale sia ricoperto in modo uniforme ovvero i dati risultino equispaziati dall'inizio alla fine;
3. accertare che il numero delle misurazioni non sia affetto da variazioni di calendario che determinerebbero sottoserie di lunghezza diversa per periodi temporali di uguale estensione.

Inoltre è consigliabile effettuare, per prima, l'analisi grafica della serie storica, in quanto il risultato ottenuto non pregiudica o altera, bensì supporta e conferma la successiva analisi quantitativa (classica *aut* moderna) che si vuole realizzare.

In tale sede tratteremo esclusivamente l'analisi grafica e l'analisi classica.

4 – Analisi grafica

La serie storica x_t con $t \in [1, T]$ può essere intesa come una 2-d. semplice $\{(t, x_t)\}_{1 \leq t \leq T}$ la cui forma grafica, detta **timeplot**, è la spezzata ottenuta unendo tutti i punti $P_t \equiv (t, x_t)$ del diagramma

cartesiano tOx_t ^d. L'**analisi grafica** di x_t con $t \in [1, T]$ consiste nel rilevare nella forma della spezzata corrispondente quelle regolarità significative che permettono di descrivere il fenomeno o di prevedere valori futuri a breve termine di X .

Esistono forme di immediata interpretazione grafica:

- lineare orizzontale = fenomeno costante;
- lineare obliqua verso l'alto = fenomeno che cresce in modo costante;
- lineare obliqua verso il basso = fenomeno che decresce in modo costante;

^d Sulle ascisse si rappresentano le unità di tempo t e lungo le ordinate i valori x_t di X .

Analisi delle Serie storiche

- ad “u” verso l’alto = fenomeno che prima decresce in modo quadratico e poi cresce in modo quadratico;
- ad “u” verso il basso = fenomeno che prima cresce in modo quadratico e poi decresce in modo quadratico;
- ad “s” rovesciata = fenomeno periodico;
- ad “l” = fenomeno che velocemente decresce;
- ad “l” ruotata = fenomeno che cresce velocemente;
- ...

È difficile che la forma grafica complessiva della serie storica suggerisca un’immediata interpretazione del fenomeno. Talvolta escludendo i dati estremali oppure suddividendo la serie in sottoserie di periodi temporali di uguale estensione possiamo pervenire a

risultati soddisfacenti per la descrizione del fenomeno in esame.

Le serie storiche di dati di Tab. 2 e di Tab. 3 sono rispettivamente graficate in Fig. 1 ed in Fig. 2. La serie storica del NIC presenta un andamento crescente e regolarità a cadenze trimestrali.

Analisi delle Serie storiche

 mese 	 indice 	 mese 	 indice 	 mese 	 indice 	 mese 	 indice
		gennaio -2009	97,2	gennaio-2010	99,8	gennaio-2011	94,7
febbraio-2008	104,7	febbraio	96,3	febbraio	97,3	febbraio	96,1
marzo	104,6	marzo	90,1	marzo	96,2		
aprile	105,4	aprile	92,5	aprile	96,7		
maggio	101,6	maggio	93,8	maggio	98,4		
giugno	105,1	giugno	94,2	giugno	97,1		
luglio	102,6	luglio	95,6	luglio	97,4		
agosto	101,4	agosto	97,5	agosto	97,6		
settembre	100,6	settembre	95,8	settembre	95,9		
ottobre	101,0	ottobre	96,0	ottobre	94,9		
novembre	100,4	novembre	96,9	novembre	95,6		
dicembre	99,0	dicembre	97,1	dicembre	96,8		

Tabella 2: Serie storica del PIL in Italia da feb'08 a feb'11 (*Fonte Istat*)

 mese 	 indice 	 mese 	 indice 	 mese 	 indice 	 mese 	 indice
gennaio	138,5	aprile	139,5	luglio	140,1	ottobre	140,4
febbraio	138,6	maggio	139,6	agosto	140,4	novembre	140,4
marzo	139,0	giugno	139,6	settembre	140,1	dicembre	140,9

Tabella 3: Serie storica del NIC da gen'10 a dic'10 (*Fonte Istat*)

Analisi delle Serie storiche

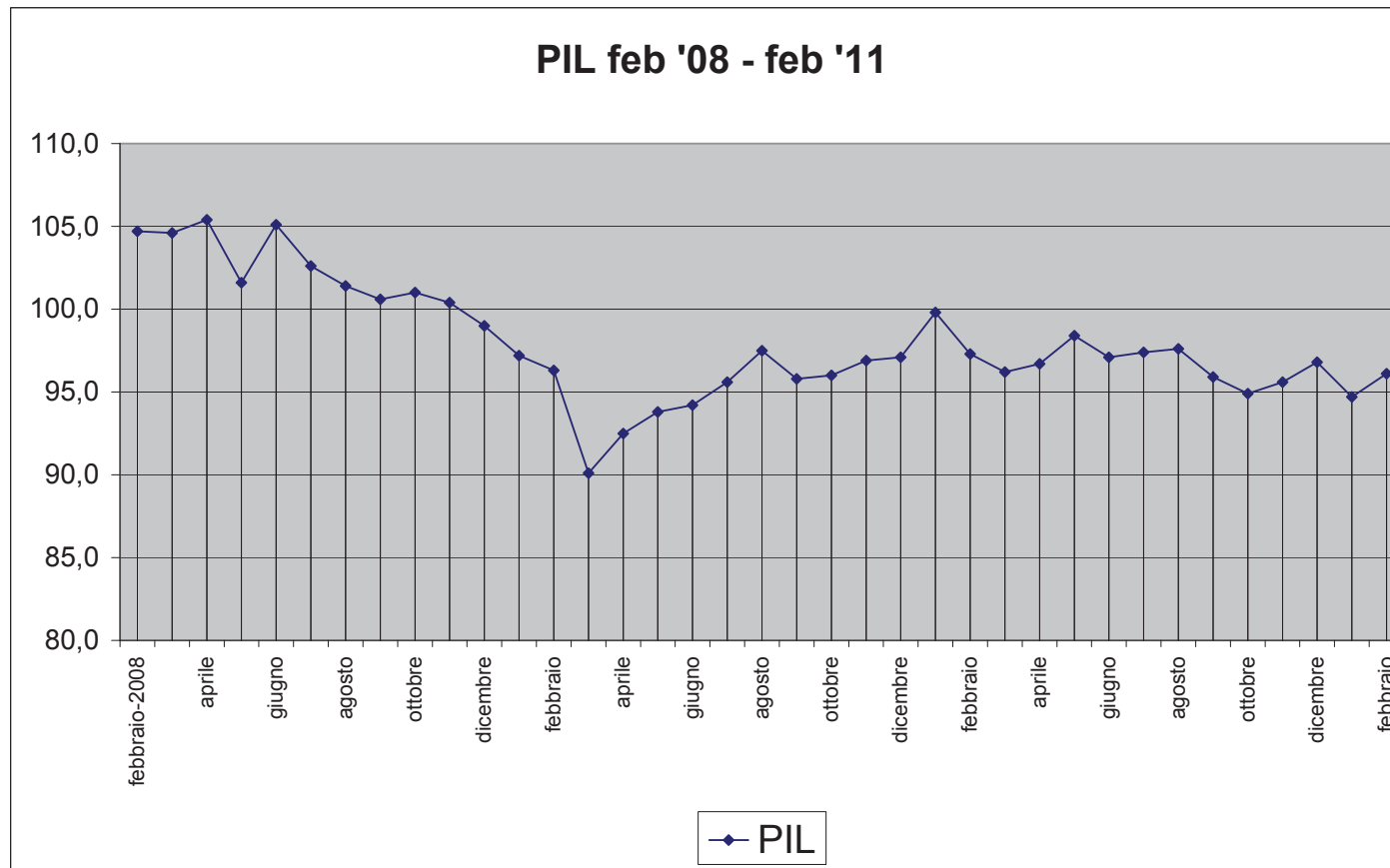


Figura 1: Rappresentazione grafica del PIL (*Fonte Istat*)

Analisi delle Serie storiche

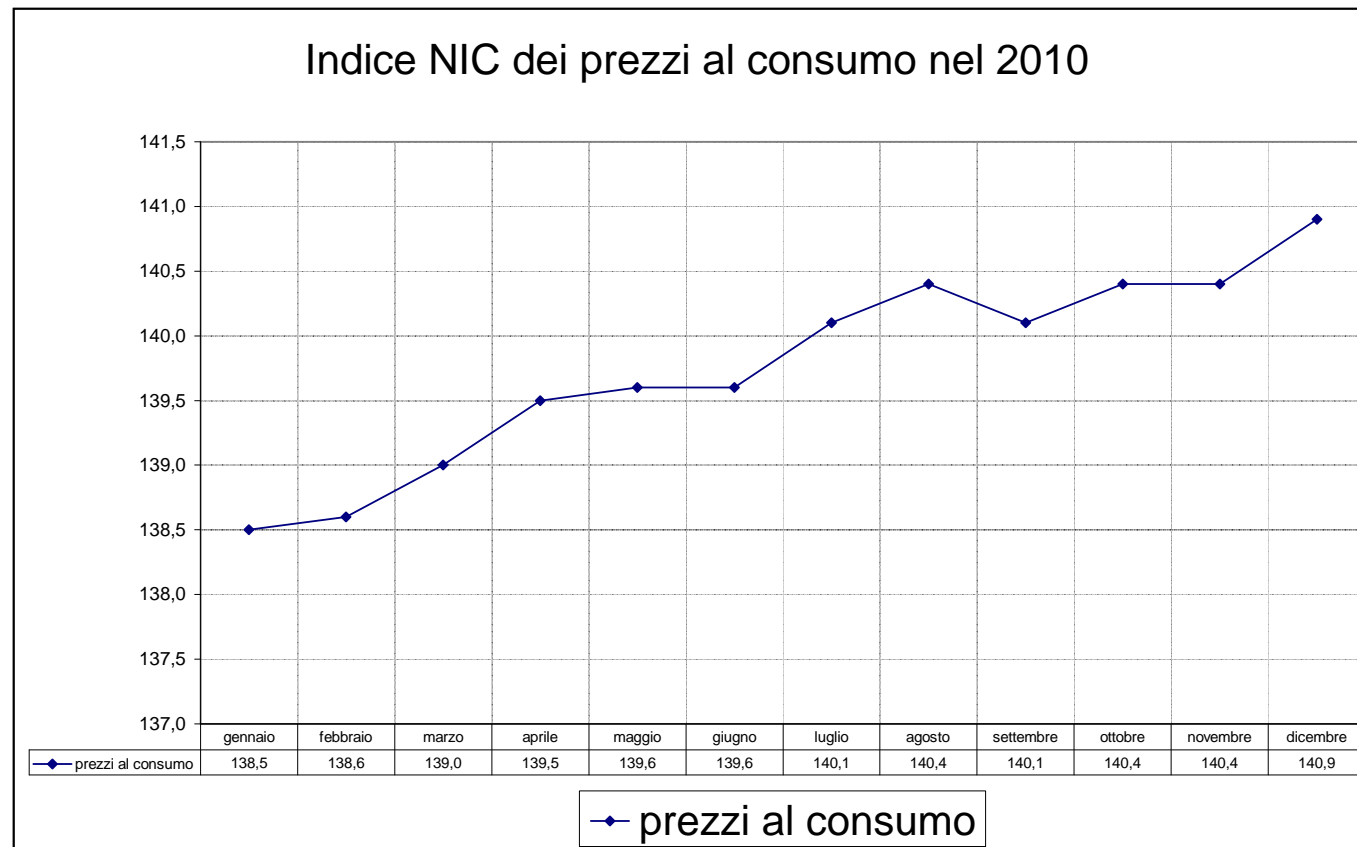


Figura 2: Rappresentazione grafica del NIC (*Fonte Istat*)

Proposizione 1. *Nelle serie storiche esiste dipendenza tra le osservazioni successive.*

$$x_t = F(x_{t-1}) \quad t = 2, 3, \dots, T$$

In altre parole, data la 2-d. semplice $\{(x_{t-1}, x_t)\}_{2 \leq t \leq T}$ i cui dati sono le coppie dei valori contigui della serie data, allora nel diagramma cartesiano $x_{t-1} O x_t$ i punti di coordinate tale coppie si addensano attorno ad una retta. E pertanto il coefficiente $r_P(x_{t-1}; x_t)$ di correlazione del Pearson di tale distribuzione bivariata semplice, ha un valore prossimo ad 1 che dimostra l'esistenza di una forte correlazione diretta o inversa tra i dati contigui.

In Fig. 3 ed in Fig. 4 viene verificata la Proposizione 1 dei fenomeni descritti in Tab. 2 ed in Tab. 3.

Analisi delle Serie storiche

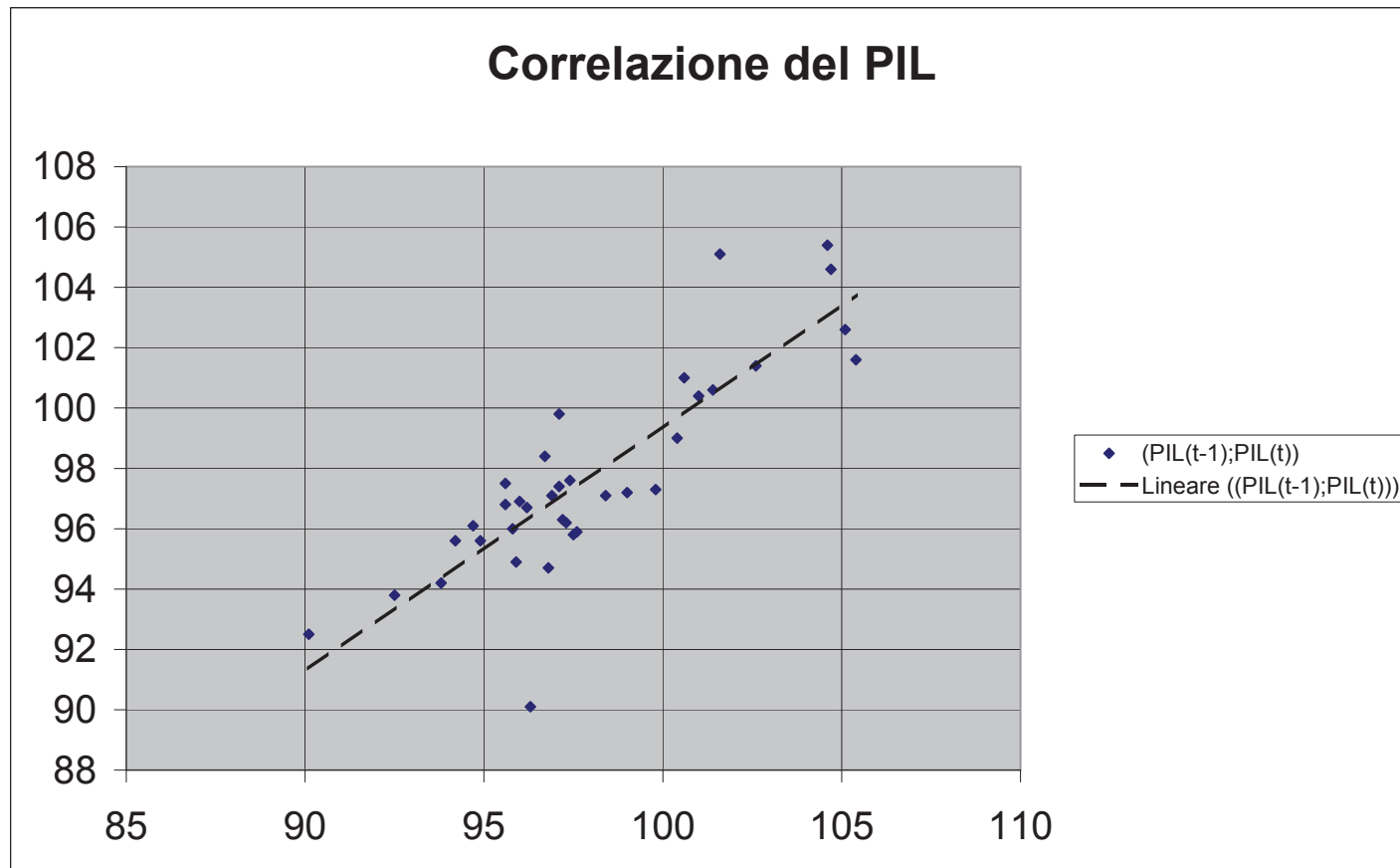


Figura 3: $r_P(PIL(t-1); PIL(t)) = 0.85$

Analisi delle Serie storiche

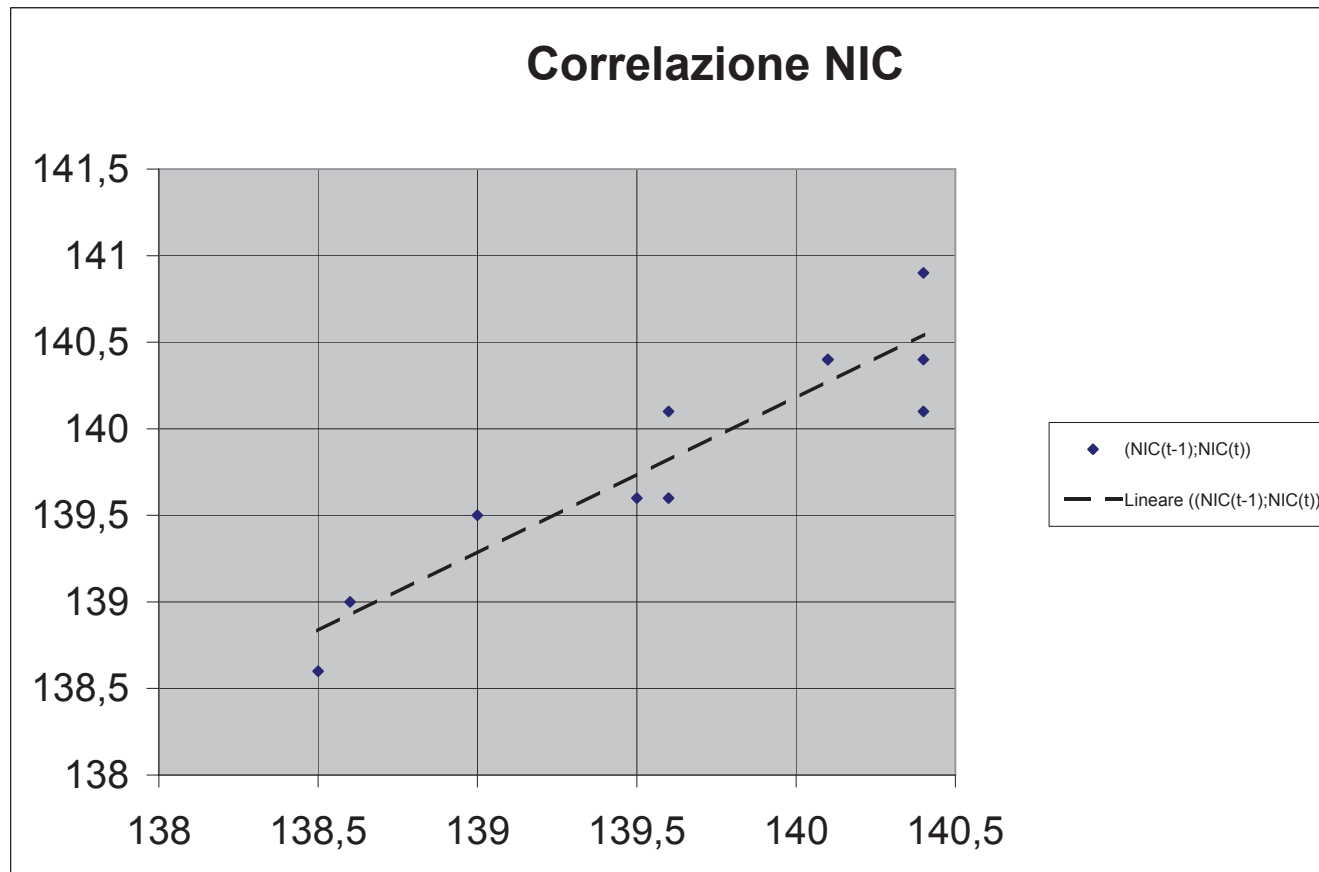


Figura 4: $r_P(NIC(t-1); NIC(t)) = 0.93$

5 – Analisi classica

Nell'approccio classico^a il modello generatore è un **modello decompositivo** o **modello per componenti** perchè in esso si rilevano i fattori detti componenti che influiscono sull'andamento temporale del fenomeno indagato.

5.1 – Le componenti della serie storica

Le componenti del modello decompositivo sono:

- F_t = *componente di fondo o trend*.

È la tendenza generale a crescere, a decrescere o rimanere costante in un arco di tempo; è la componente regolare di

^a L'analisi classica risale agli 20 del XX secolo e rappresenta la prima procedura formalizzata di interpretazione di un fenomeno attraverso le serie storiche.

maggior interesse che può assumere l'aspetto di una funzione matematica esplicita di tipo lineare, quadratico, esponenziale, logistico, etc...

- $C_t = \text{componente ciclica}$

È atta a descrivere l'andamento oscillante a grandi periodi che si ripetono intorno al trend; è una componente più indeterminata di F_t e si considera *congiunturale* perchè di norma è abbinata al trend.

prosperità, crescita, depressione, ripresa

- $S_t = \text{componente stagionale o stagionalità}$

È presente per quei fenomeni osservati in intervalli di tempo inferiore all'anno che subiscono variazioni legate all'alternanza delle stagioni; è una componente esterna.

Analisi delle Serie storiche

la vendita di panettoni a dicembre, il consumo di gelato in estate, ...

Se la serie è osservata con periodicità s ($=4$ trimestrale, $=3$ quadrimestrale, $=2$ semestrale, $=12$ mensile, $=52$ settimanale) l'effetto stagionale gode delle proprietà:

$$S_t = S_{t+s} \quad (1)$$

$$\sum_{i=1}^s S_t = 0 \quad (2)$$

cioè si ripete dopo il periodo s e la stagionalità ha un effetto limitato a sottoperiodi dell'anno.

- $E_t =$ *componente erratica o componente accidentale*

Giustifica la presenza nella serie storica di valori anomali legati a particolari eventi, ad innovazioni, o ristrutturazioni commerciali,

... Tale componente è considerata “residua” in quanto ottenuta come differenza tra il dato reale x_t ed il dato teorico \bar{x}_t calcolato dalla parte deterministica della serie.

In definitiva il modello generatore per componenti è una funzione delle quattro componenti:

$$x_t = \phi(F_t, C_t, S_t, E_t) \quad (3)$$

oppure è una funzione delle tre componenti FC_t, S_t, E_t :

$$x_t = \phi_1(FC_t, S_t, E_t) \quad (4)$$

con FC_t , componente *trend-ciclo*, ottenuta dalla commistione della F_t e della C_t .

Nota 1. Analizzeremo solo i modelli generatori del tipo (4).

5.2 – I modelli generatori per componenti

Le componenti FC_t , S_t e E_t possono aggregarsi in diversi modi.

- **il modello generatore additivo per componenti:**

La funzione ϕ_1 è l'operatore somma; le componenti sono indipendenti tra loro e hanno la stessa unità di misura della x_t (vedi Fig. 5).

$$x_t = FC_t + S_t + E_t \quad (5)$$

- **il modello generatore moltiplicativo per componenti:**

La funzione ϕ_1 è l'operatore prodotto; le componenti sono dipendenti tra loro e mentre FC_t ha misura uguale a x_t , le altre

$(S_t$ e E_t) sono adimensionali ossia numeri puri (vedi Fig. 6).

$$x_t = FC_t \cdot S_t \cdot E_t \quad (6)$$

- **il modello generatore misto per componenti:**

Componendo la (5) e la (6) si ottengono vari **modelli misti**^b. Il più significativo è il *modello misto con E_t additiva* (vedi Fig. 7 e Fig. 8).

$$x_t = (FC_t \cdot S_t) + E_t \quad (7)$$

Osservazione 1. Se al modello moltiplicativo (6) applichiamo una

^b La varietà e la validità dipende dal numero delle componenti in esame.

trasformazione logaritmica otteniamo un modello additivo.

$$\begin{aligned}\log(x_t) &= \log(FC_t \cdot S_t \cdot E_t) = \\ &= \log(FC_t) + \log(S_t) + \log(E_t)\end{aligned}$$

Tale trasformazione è utile perché:

1. non altera le caratteristiche della serie x_t (vedi Fig. 9);
2. riduce le fluttuazioni di S_t e E_t (vedi Fig. 10);
3. permette di ridurre l'analisi al solo modello additivo.

Analisi delle Serie storiche

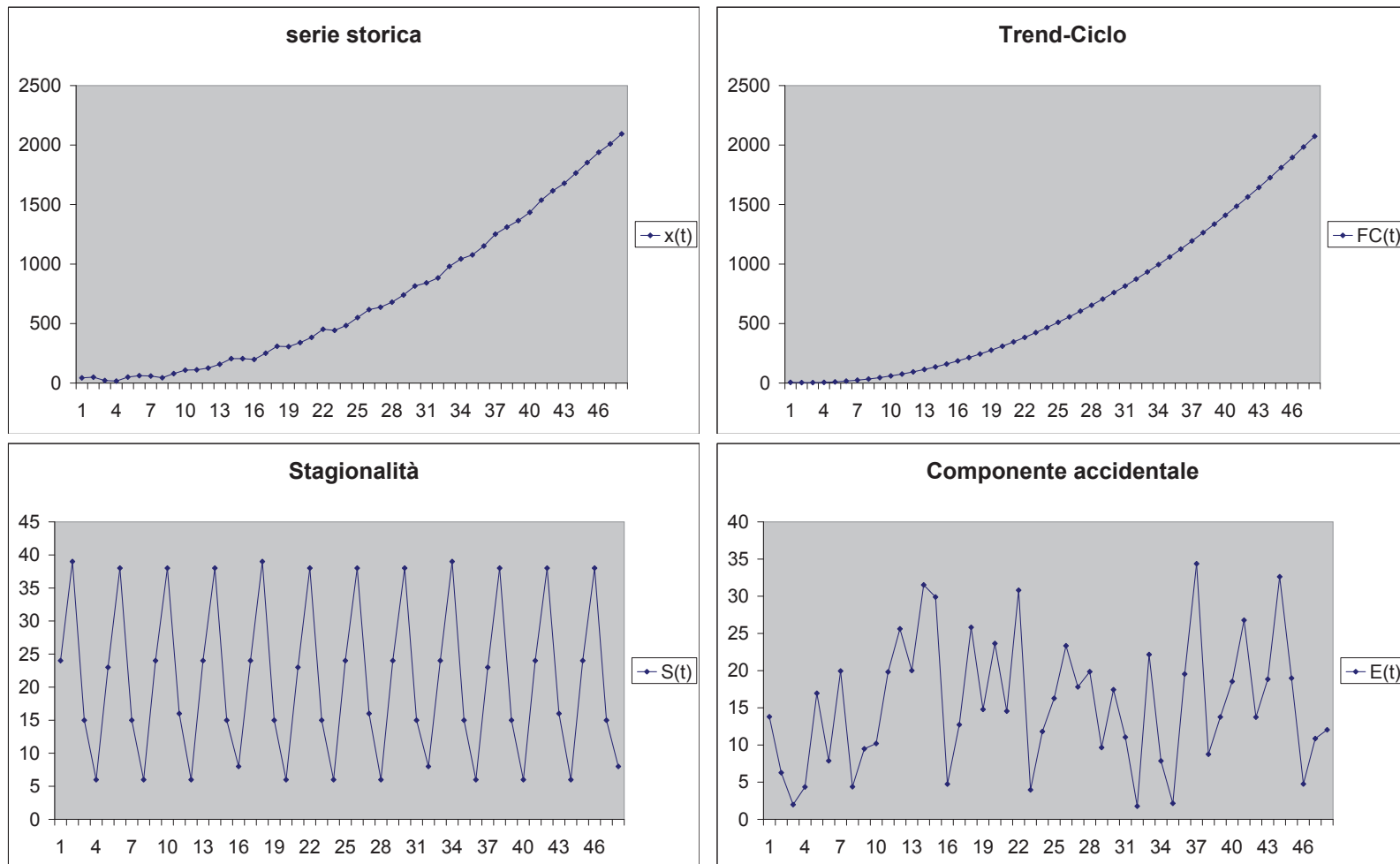


Figura 5: Esempio di serie storica con modello additivo

Analisi delle Serie storiche

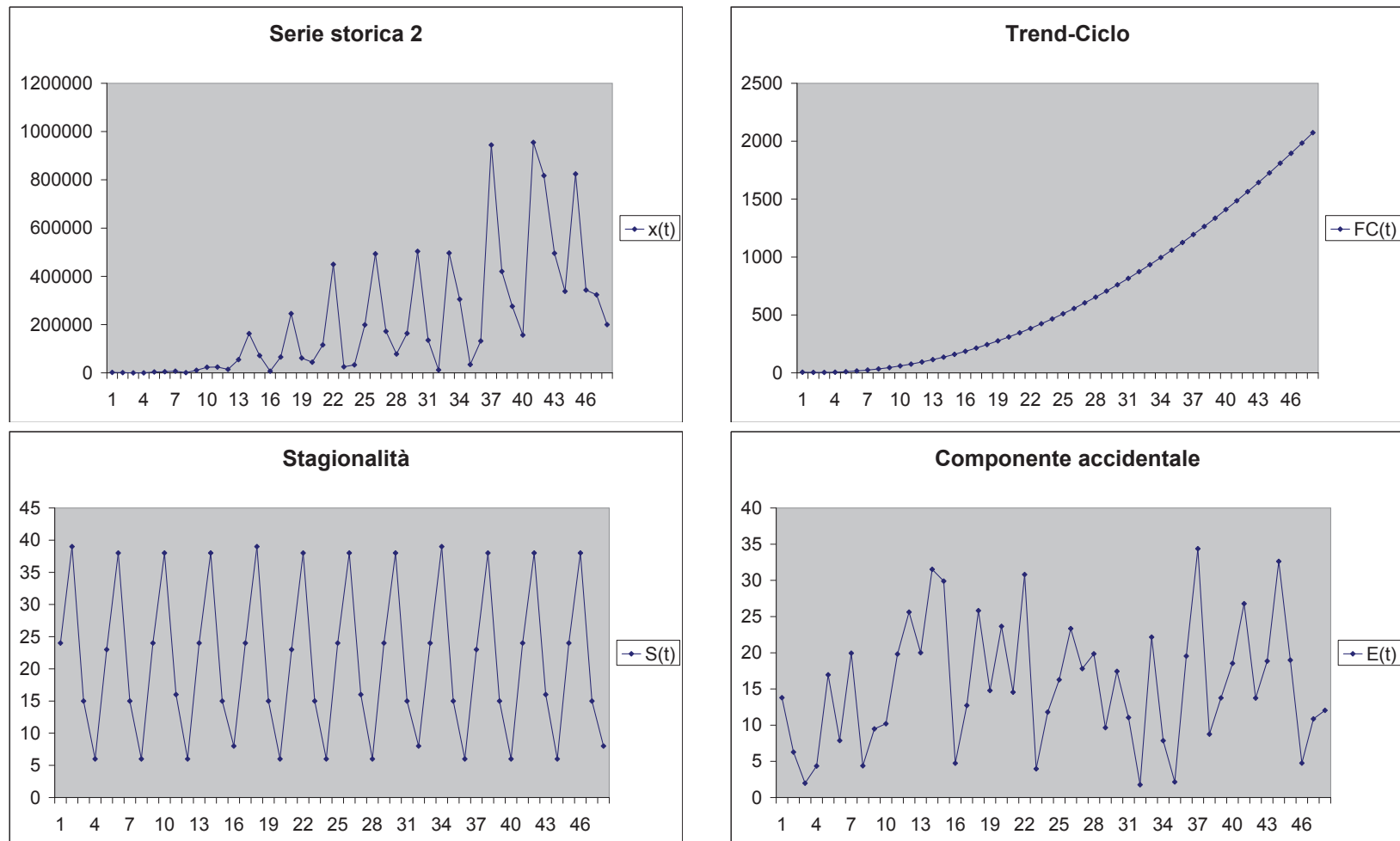


Figura 6: Esempio di serie storica con modello moltiplicativo

Analisi delle Serie storiche

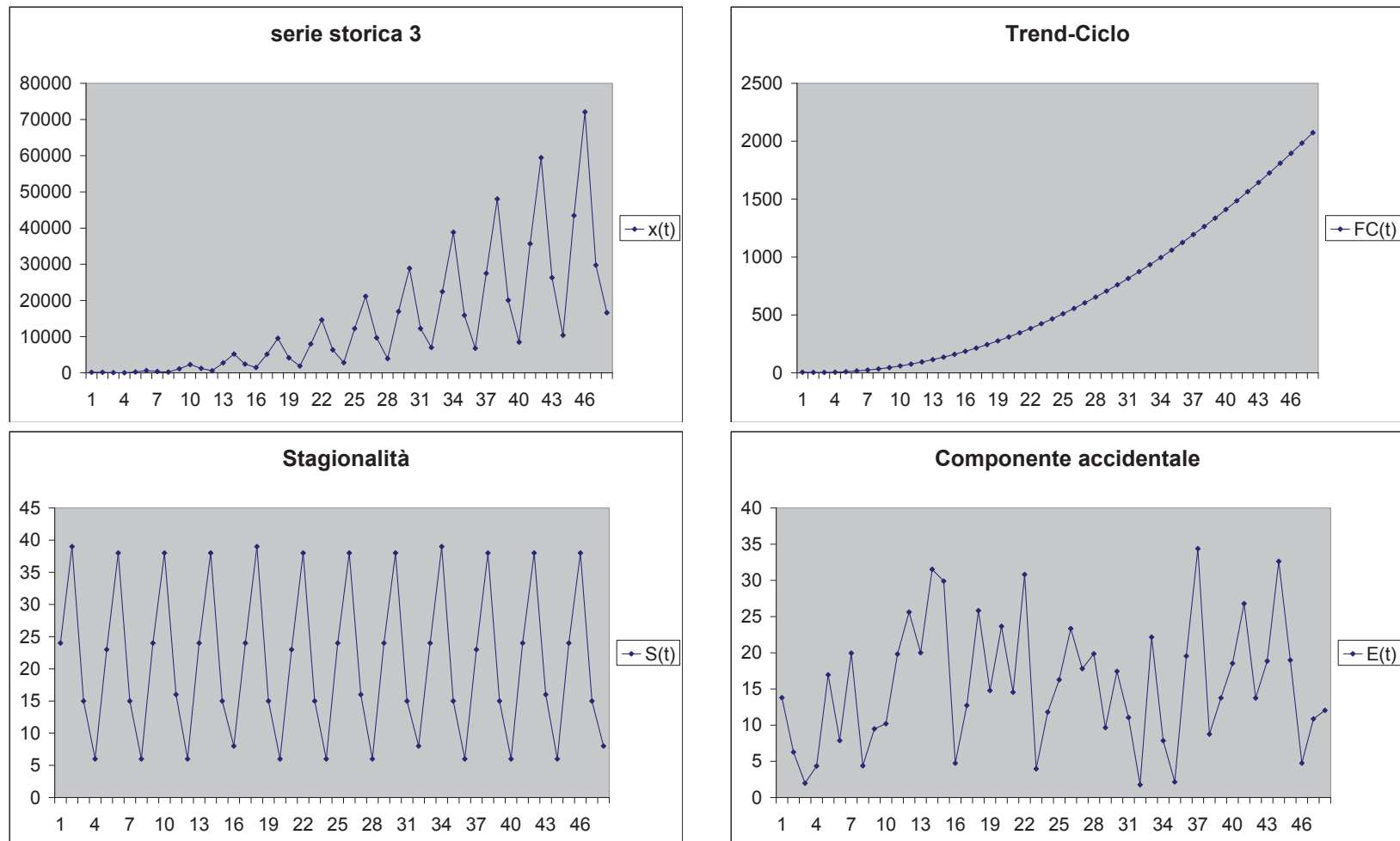


Figura 7: Esempio di serie storica con modello misto

Analisi delle Serie storiche

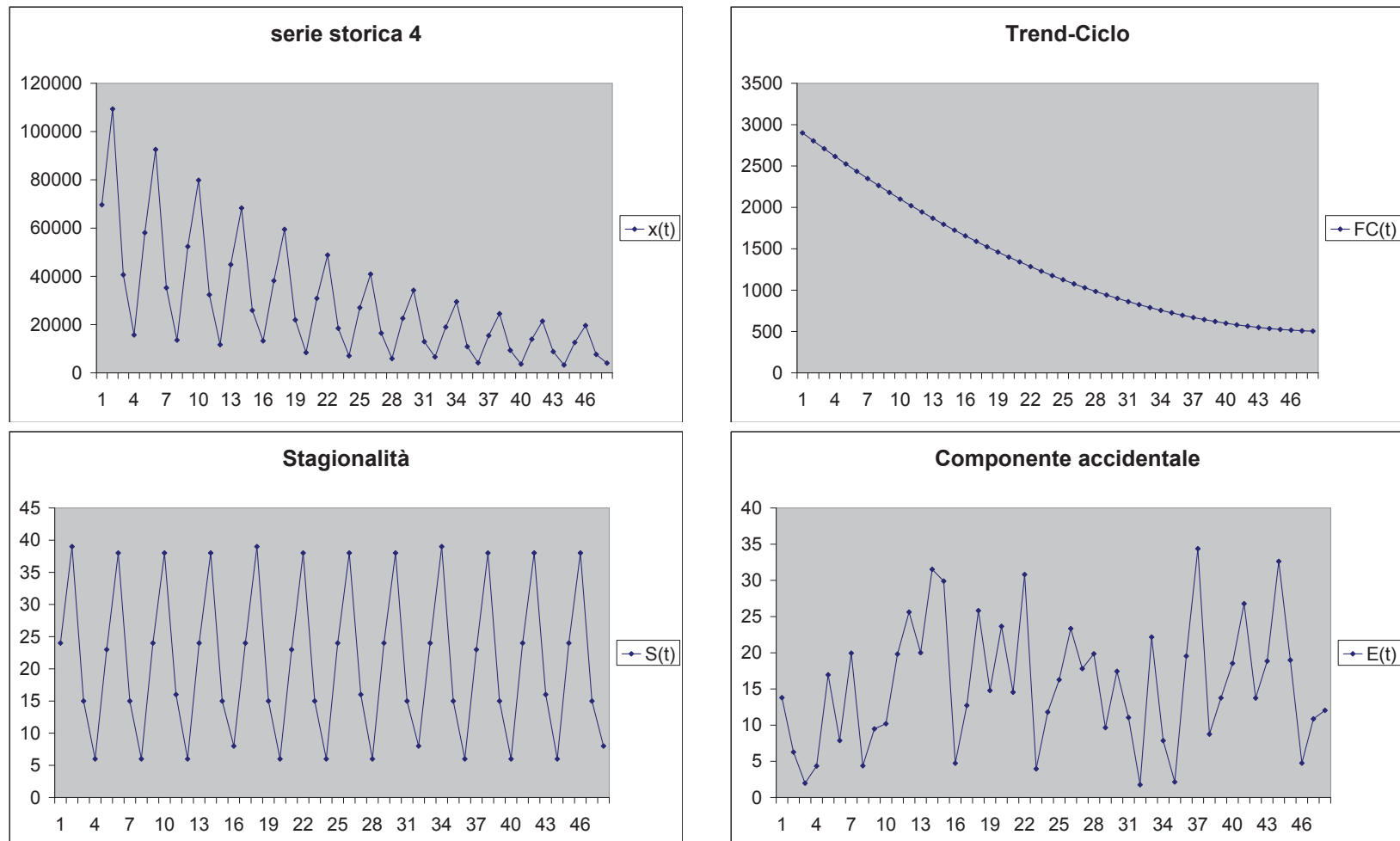


Figura 8: Esempio 2 di serie storica con modello misto

Analisi delle Serie storiche

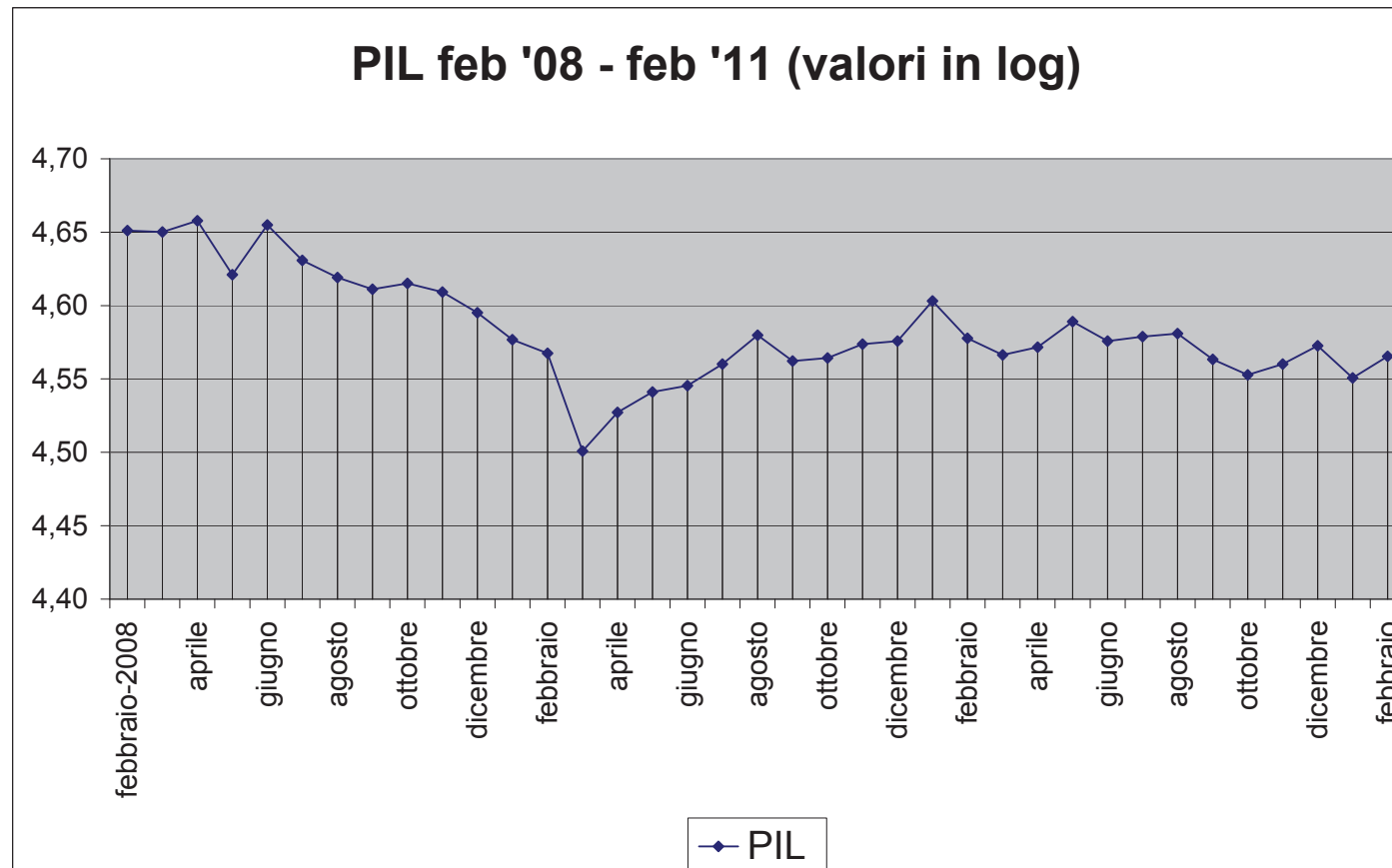


Figura 9: Serie storica del valore logaritmico del PIL

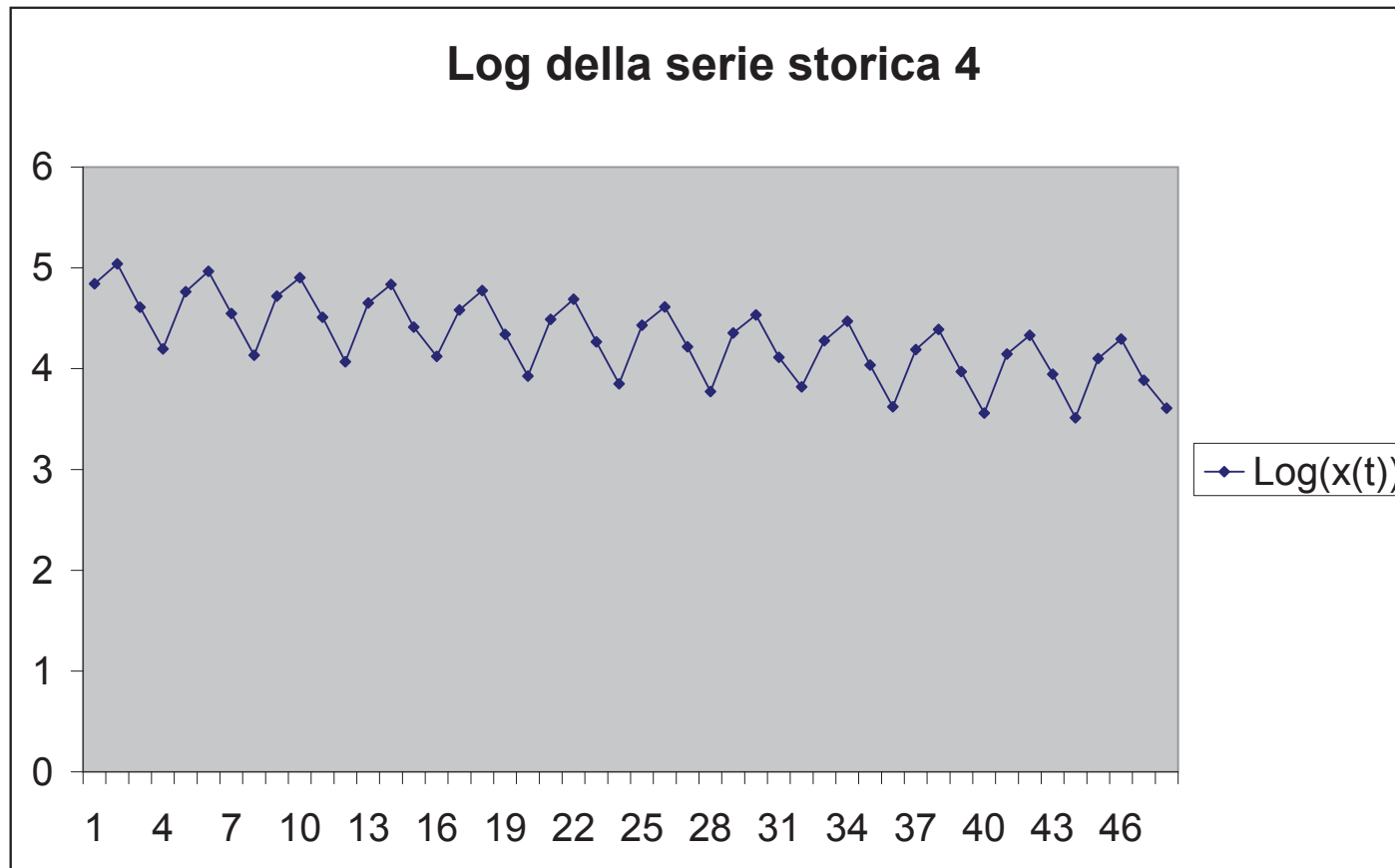


Figura 10: Serie storica del logaritmo della serie storica 4

6 – La Statistica nell'approccio classico

La Statistica ha un ruolo fondamentale nell'analisi delle serie storiche. Nella fase descrittiva - realizzazione del modello generatore - mette a disposizione le procedure:

- interpolazione statistica
- medie mobili

che permettono di individuare le componenti sistematiche ed erratica e nella fase di previsione:

- le distribuzioni di probabilità
- la teoria campionaria
- gli intervalli di confidenza

Di quelli già noti forniremo alcune precisazioni mentre di quelli non noti porremo le basi per la loro principale caratterizzazione.

6.1 – Interpolazione statistica

L'interpolazione statistica è quella tecnica che ci permette di individuare la curva che passa per dei punti assegnati o vicino ad essi. Nel caso delle serie storiche l'uso che se ne fa riguarda l'individuazione della componente trend-ciclo e pertanto è equivalente a determinare il polinomio lineare

$$\bar{x}_t = \alpha_0 + \alpha_1 t + \dots + \alpha_n t^n$$

tale che valga la condizione^a

$$\inf_{(\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{R}^n} \sum_{t=1}^T (x_t - \bar{x}_t)^2$$

del tutto equivalente alla risoluzione del sistema

$$\left\{ \begin{array}{l} \alpha_0 T + \alpha_1 \sum_t t + \alpha_2 \sum_t t^2 + \dots + \alpha_n \sum_t t^n = \sum_t x_t \\ \alpha_0 \sum_t t + \alpha_1 \sum_t t^2 + \dots + \alpha_{n-1} \sum_t t^n + \alpha_n \sum_t t^{n+1} = \sum_t t x_t \\ \dots \\ \alpha_0 \sum_t t^n + \alpha_1 \sum_t t^{n+1} + \dots + \alpha_n \sum_t t^{2n} = \sum_t t^n x_t \end{array} \right.$$

Nelle serie storiche gli x_t sono equispaziati e pertanto se si effettua la

^a La determinazione dei parametri α_i avviene mediante il **metodo dei minimi quadrati**.

sostituzione

$$s_t = 2t - (T + 1)$$

si annullano le somme delle potenze dispari ed il sistema diventa

$$\left\{ \begin{array}{l} \beta_0 T + \quad + \beta_2 \sum_t s_t^2 + \dots + \beta_n \sum_t s_t^n = \sum_t x_t \\ \quad + \beta_1 \sum_t s_t^2 + \dots + \beta_{n-1} \sum_t s_t^n + \quad = \sum_t s_t x_t \\ \dots \\ \beta_0 \sum_t s_t^n + \quad + \dots + \beta_n \sum_t s_t^{2n} = \sum_t s_t^n x_t \end{array} \right.$$

Nonostante la semplificazione effettuata è conveniente ricorrere alle procedure di linearizzazione^b per $n > 3$ individuando altre funzioni lineari.

^b Si procede ad un cambiamento di variabili passando ai logaritmi. Vedi § 2.2. del Capitolo 7 *La regressione*.

6.2 – Medie mobili

È una tecnica utilizzata quando l'andamento della serie storica è molto irregolare.

Definizione 1. Data una serie storica x_t di peso w_t (con $t \in [1, T]$) e fissato un $k \in \mathbb{N}$ tale che $1 \leq k \leq T$ definiamo **media mobile** di ordine k una somma ponderata di k valori consecutivi della serie. Per la serie x_t esistono $T - k + 1$ medie mobili.

$$MM(k) = \begin{cases} \sum_{t=1}^k w_t x_t, & 1^a \text{ media mobile;} \\ \sum_{t=2}^{k+1} w_t x_t, & 2^a \text{ media mobile;} \\ \dots, & \dots; \\ \sum_{t=T-k}^T w_t x_t, & T - k + 1\text{-esima media mobile.} \end{cases} \quad (8)$$

Definizione 2. Una **media mobile** di ordine k è **semplice** se i coefficienti di ponderazione della serie sono tutti uguali. Una media mobile semplice di ordine k è allora la media aritmetica di k valori consecutivi della serie.

$$MMS(k) = \begin{cases} \frac{1}{k} \sum_{t=1}^k x_t, & 1^a \text{ m.mobile semplice;} \\ \frac{1}{k} \sum_{t=2}^{k+1} x_t, & 2^a \text{ m.mobile semplice;} \\ \dots, & \dots; \\ \frac{1}{k} \sum_{t=T-k}^T x_t, & T - k + 1\text{-esima m.mobile semplice.} \end{cases} \quad (9)$$

Definizione 3. Una **media mobile** semplice di ordine k è **centrata** nella posizione i se cade nel periodo i di osservazione della serie originaria.

$$MMC_i(k) = \begin{cases} \frac{1}{k} \sum_{t=i-\left(\frac{k-1}{2}\right)}^{i+\left(\frac{k-1}{2}\right)} x_t, & \text{per } k \text{ dispari;} \\ \frac{1}{2} \left(\frac{1}{k} \sum_{t=i-\frac{k}{2}}^{t+\frac{k}{2}-1} x_t + \sum_{t=i-\frac{k}{2}+1}^{t+\frac{k}{2}} x_t \right), & \text{per } k \text{ pari.} \end{cases} \quad (10)$$

Osservazione 2. Tutte le medie mobili semplici di ordine dispari k sono centrate in $t + \frac{k+1}{2}$ con $t = 0, 1, \dots, T - k$.

★★★**Esercizio 1.** Calcolare le medie mobili di ordine 3 e di ordine 4 della serie storica pesata riportata in Tab. 4.

t	1981	1982	1983	1984	1985	1986	1987
x_t	3	4	8	5	6	3	12
w_t	0.15	0.10	0.20	0.15	0.23	0.07	0.10

Tabella 4: Serie storica pesata

Soluzione: Nella Tab. 5 sono riportati i risultati.



Analisi delle Serie storiche

t	x_t	w_t	$x_t \times w_t$	MM(3)	MM(4)
1981	3	0,15	0,45		
1982	4	0,10	0,40	2,45	3,2
1983	8	0,20	1,60	2,75	4,13
1984	5	0,15	0,75	3,73	3,94
1985	6	0,23	1,38	2,34	3,54
1986	3	0,07	0,21	2,79	
1987	12	0,10	1,20		

Tabella 5: Medie mobili di ordine 3 e di ordine 4

Analisi delle Serie storiche

★★★**Esercizio 2.** Con i dati di Tab. 6 calcolare le medie mobili centrate di ordine 3 e di ordine 4.

t	1981	1982	1983	1984	1985	1986	1987
x_t	6	4	8	3	4	11	12

Tabella 6: Serie storica

Soluzione: Ricordando che se k è dispari le m.mobili semplici sono anche centrate e per k pari occorre operare la media delle due m.mobili semplici a ridosso del periodo, in ossequio alla formula (10) otteniamo i risultati riportati in Tab. 7. Ad. es. 6 è la $MMC(3)$ centrata in $t = 1982$; 7 è la $MMC(4)$ centrata in $t = 1985$; ... \square

Analisi delle Serie storiche

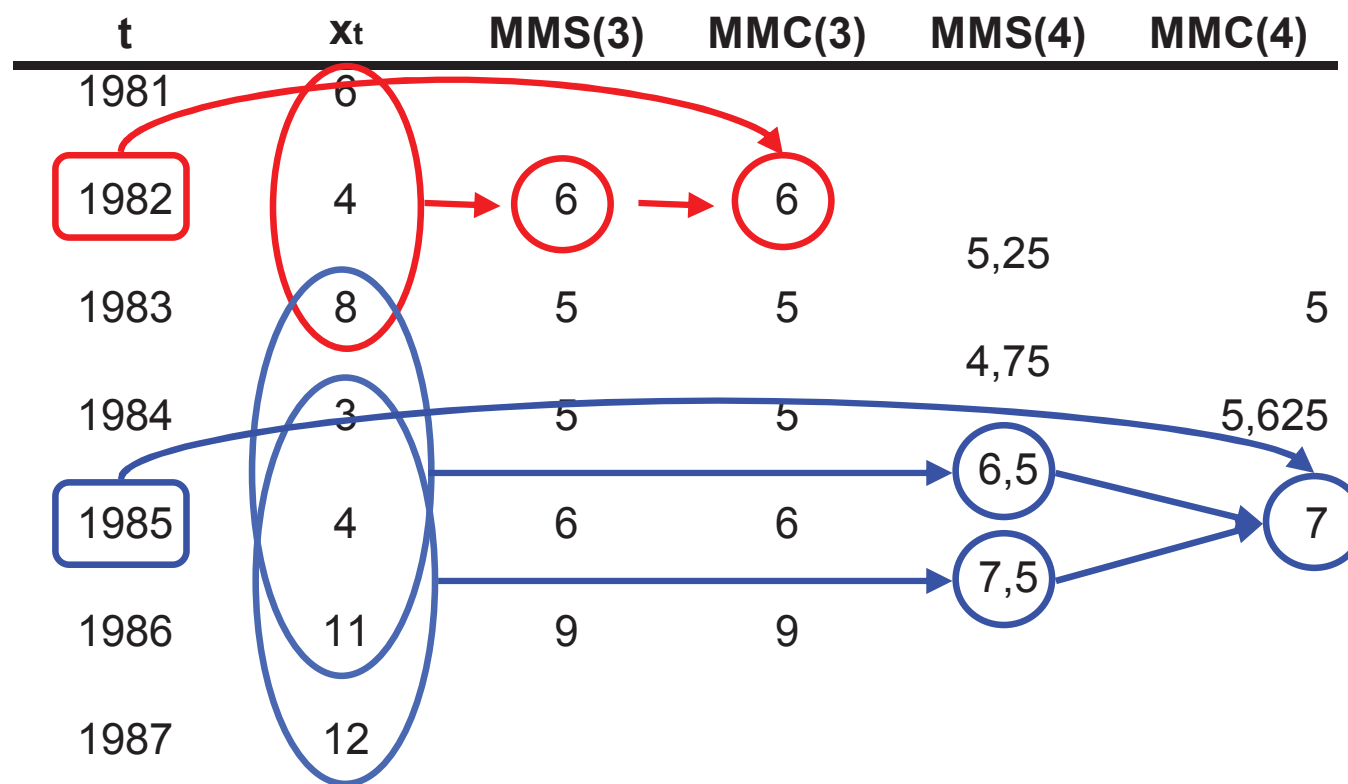


Tabella 7: Medie mobili centrate di ordine 3 e di ordine 4

6.3 – Teoria dei campioni

La teoria dei campioni trova la sua applicazione nella valutazione della componente accidentale E_t . Se intendiamo tale componente come una distribuzione della media campionaria degli errori \bar{X} di una popolazione $X \sim N(\mu, \sigma)$ dovrà soddisfare le seguenti condizioni^a

$$\mu_{\bar{X}} = \mu \quad \text{e} \quad \sigma_{\bar{X}}^2 = \frac{\sigma^2}{T}$$

Inoltre E_t è calcolata come residuo $x_t - \bar{x}_t$ per cui

$$M(E_t) = 0 \quad Var(E_t) = \sigma^2 \quad r_P(E_{t-1}, E_t) = 0 \quad \text{b} \quad (11)$$

^a Vedi Capitolo 12 con particolare riguardo al § 5.1 del

^b L'ultima condizione afferma che gli errori sono tra loro indipendenti.

7 – Processo di scomposizione

Attraverso opportuni metodi statistici, si scompone la serie fino alla completa specificazione delle sue componenti. Nel processo di scomposizione è privilegiata la determinazione del pattern sistematico (vedi Tab. 8) ossia di quella parte del modello generatore costituito dal trend-ciclo e dalla stagionalità che segue un andamento regolare.

MODELLO	additivo	moltiplicativo	misto
PATTERN SISTEMATICO	$FC_t + S_t$	$FC_t \cdot S_t$	$FC_t \cdot S_t$

Tabella 8: Pattern sistematico del modello generatore per componenti

La componente accidentale E_t calcolata a saldo deve soddisfare le condizioni in (11).

Il processo di scomposizione si articola in 8 fasi:

1. *scelta del modello*
2. *calcolo del trend-ciclo di prima approssimazione*
3. *calcolo della componente stagionale ed accidentale congiunte*
4. *stima della componente stagionale*
5. *destagionalizzazione della serie*
6. *stima del trend-ciclo*
7. *stima del pattern sistematica*
8. *stima del modello generatore*

Laddove si esclude la componente stagionale dal modello generatore perché la cadenza temporale della serie è maggiore od uguale all'anno allora dalla sequenza sono eliminate le fasi 1 - 2 - 3 - 4 - 5 - 7^c.

Comunque di questo particolare caso discuteremo successivamente e in una sezione distinta.

7.1 – Scelta del modello

Parliamo di scelta del modello perchè l'approccio classico tende a descrivere la serie storica e non ad interpretarla. Precisiamo ancora che possiamo osservare gli x_t ma non le singole componenti contenute in essi; è chiaro allora che il modello generatore che realizzeremo sarà

^c La fase 1 perde di significato perchè tutti i modelli generatori (additivo, moltiplicativo e misto) hanno lo stesso pattern sistematico costituito soltanto da FC_t .

uno delle tante interpretazioni di x_t con $t \in [1, T]$ e che più di altre esalterà alcune caratteristiche e meno di altre nasconderà altre peculiarità^d.

In base all'Osservazione 1 la scelta cadrà sul modello additivo o sul modello misto. Nelle fasi successive saranno indicate le formule rispettivamente a sinistra per il modello additivo e a destra per il modello moltiplicativo.

7.2 – Calcolo del trend-ciclo di 1^a approssimazione

In questa fase il trend-ciclo, indicato con \overline{FC}_t viene approssimato con la media mobile centrata $MMC_t(k)$ con $k = 2, 3, 4, 6, 12$ a seconda se

^d In altre parole l'approccio classico nei confronti della serie è come “cucire un vestito addosso ad una persona senza sapere la taglia, l'altezza, ...”.

Analisi delle Serie storiche

la cadenza temporale è semestrale, quadrimestrale, trimestrale, bimestrale e mensile.

$$MMC_t(k) = \overline{FC}_t \qquad MMC_t(k) = \overline{FC}_t \qquad (12)$$

Ricordiamo che questa prima stima porterà alla perdita di $\frac{k-1}{2}$ dati all'inizio della serie ed altrettanti $\frac{k-1}{2}$ alla fine per k dispari; altrimenti per k pari alla perdita dei primi $\frac{k}{2}$ termini e degli ultimi $\frac{k}{2}$ termini.

La perdita dei primi termini ha poca importanza mentre è senz'altro più rilevante la perdita degli ultimi che in ordine cronologico rappresentano i valori di X più recenti. Si può ovviare a tali perdite iniziali e finali rimpiazzando i dati mancanti con medie mobili centrate di ordine minore e decrescente fino al valore 2.

★☆☆ **Esercizio 3.** Completa la colonna delle $MMC_t(3)$ e $MMC_t(4)$ dei dati di Tab. 7

Soluzione: Per $k = 3$ la colonna delle $MMC_t(3)$ è priva del 1° termine che verrà rimpiazzato da

$$MMC_1(2) = \frac{x_1 + x_2}{2} = \frac{6 + 4}{2} = 5;$$

ed è deficitaria del 7°-termine che verrà sostituito con

$$MMC_7(2) = \frac{x_6 + x_7}{2} = \frac{11 + 12}{2} = 11.5.$$

Per quel che riguarda la colonna delle $MMC_t(4)$ poichè mancano 2 termini iniziali e 2 finali opereremo il rimpiazzo due volte con medie mobili centrate di ordine 3 per ottenere il 2° ed 6° valore e con medie

Analisi delle Serie storiche

mobili centrate di ordine 2 per ottenere il 1° ed 7° valore:

$$1 \mapsto MMC_1(2) = \frac{x_1 + x_2}{2} = \frac{6 + 4}{2} = 5;$$

$$2 \mapsto MMC_2(3) = \frac{6 + 4 + 8}{3} = 6;$$

... \mapsto ...;

$$t \mapsto MMC_t(4);$$

... \mapsto ...;

$$6 \mapsto MMC_6(3) = \frac{4 + 11 + 12}{3} = 9;$$

$$7 \mapsto MMC_7(2) = \frac{x_6 + x_7}{2} = \frac{11 + 12}{2} = 11.5.$$

□

7.3 – Calcolo di della componente $(S + E)_t$ o $(SE)_t$

Con il simbolo $(S + E)_t$ indichiamo la componente stagionale e la componente accidentale congiunte. Tale serie si ottiene sottraendo dalla serie originale la prima approssimazione della componente trend-ciclo

$$(S + E)_t = x_t - MMC_t \qquad (SE)_t = \frac{x_t}{MMC_t} \qquad (13)$$

7.4 – Stima della componente stagionale

Dalla componente $(S + E)_t$ congiunta di stagionalità ed accidentale eliminiamo la parte erratica, in gergo detta *noise* (rumore), ed otteniamo la stagionalità. Supposta costante l'oscillazione stagionale, in base alla proprietà (1) si ottiene in prima approssimazione che la

componente stagionale \bar{S}_t al tempo t è ottenuta come media aritmetica dei valori della congiunta stagionale-accidentale al tempo $t, t + k, t + 2k, \dots$

$$\bar{S}_t = \sum_j (S + E)_{t+jk}$$

$$\bar{S}_t = \sum_j (SE)_{t+jk}$$

e poi in ultima approssimazione tenendo conto della proprietà (2)^e che vuole che le oscillazioni si esauriscono all'interno dell'anno, allora la stima della stagionalità al tempo t si ottiene sottraendo ad ogni stima di prima approssimazione della stagionalità la media delle

^e Per il modello misto vale la proprietà $\frac{1}{k} \sum_{i=1}^k S_t = 1$.

stime stagionali di prima approssimazione^f

$$S_t = \bar{S}_t - \frac{1}{k} \sum_{t=1}^k \bar{S}_t \qquad S_t = \frac{\bar{S}_t}{\frac{1}{k} \sum_{t=1}^k \bar{S}_t} \qquad (14)$$

7.5 – Destagionalizzazione della serie

Nella fase precedente abbiamo calcolato la componente stagionale che si ripete con periodicità k per cui la serie destagionalizzata D_t si ottiene eliminando dai dati originali la stagionalità

$$D_t = x_t - S_t \qquad D_t = \frac{x_t}{S_t} \qquad (15)$$

^f Ricordiamo che per il modello additivo possiamo considerare la formula: $\sum_i (x_i - M(X)) = 0$ e per il modello misto la formula: $\sum_i \frac{x_i}{M(X)} = 1$.

7.6 – Stima del trend-ciclo

La stima del trend-ciclo FC_t può essere ottenuta mediante una media mobile centrata di ordine 3 sui dati di D_t

$$FC_t = MMC_t(3)_{[D_t]} \qquad FC_t = MMC_t(3)_{[D_t]} \qquad (16)$$

7.7 – Stima del pattern sistematico

La fase 3. ci ha permesso di calcolare la stagionalità S_t e la fase 5 di stimare il trend-ciclo FC_t . Siamo allora di ricostruire la parte deterministica della serie

$$\bar{x}_t = FC_t + S_t \qquad \bar{x}_t = FC_t \cdot S_t \qquad (17)$$

che per l'approccio classico è la parte più importante della serie

storica.

7.8 – Stima del modello generatore

È la fase conclusiva che attesta la bontà del modello generatore. Calcoliamo la componente accidentale E_t come residuo r_t della serie originale privata del pattern sistematico

$$E_t \stackrel{def}{=} r_t = x_t - \bar{x}_t \qquad E_t \stackrel{def}{=} r_t = x_t - \bar{x}_t \qquad (18)$$

L'analisi dei residui può avvenire in diversi modi:

- *andamento del timeplot della serie dei residui*

Si rappresenta il timeplot della serie dei residui e si osserva se l'andamento è sistematico. In caso di risposta positiva possiamo allora valutare il modello inefficace a descrivere la serie storica

altrimenti esprimiamo un giudizio positivo.

- *normal probability plot della serie dei residui*

Si confronta la serie accidentale ottenuta con quella accidentale di tipo normale. Se i comportamenti sono simili^g allora la bontà del modello è garantita.

Esistono ulteriori procedure per la stima del modello generatore, che non fanno uso dei residui:

- *l'analisi dei dati destagionalizzati rispetto al tempo*

Si rappresenta il timeplot della serie D_t e si osserva se sono presenti dei picchi. Se la risposta è positiva allora la destagionalizzazione non è riuscita e quindi il modello generatore

^g vedi la Regola dei tre sigma per la dispersione dei dati della serie degli errori.

non è corretto altrimenti la valutazione è buona.

- *calcolo degli indici MSE, MSEP*

Indice MSE = Mean Square Error

$$MSE = \frac{\sum_{t=1}^n (x_t - \bar{x}_t)^2}{n} \quad (19)$$

Indice $MSEP$ = Mean Square Error Percentual

$$MSEP = \frac{\sum_{t=1}^n \left(\frac{x_t - \bar{x}_t}{x_t} \right)^2}{n} \quad (20)$$

essendo n il numero complessivo delle osservazioni ottenute di \bar{x}_t

★★★**Esercizio 1.** Data la serie storica del PIL in Italia con periodo temporale febbraio 2008 - febbraio 2011 i cui dati sono riportati in Tab. 9 ricostruire il modello generatore per componenti di

Analisi delle Serie storiche

tipo additivo.

mese	indice	mese	indice	mese	indice	mese	indice
		gennaio -2009	97,2	gennaio-2010	99,8	gennaio-2011	94,7
febbraio-2008	104,7	febbraio	96,3	febbraio	97,3	febbraio	96,1
marzo	104,6	marzo	90,1	marzo	96,2		
aprile	105,4	aprile	92,5	aprile	96,7		
maggio	101,6	maggio	93,8	maggio	98,4		
giugno	105,1	giugno	94,2	giugno	97,1		
luglio	102,6	luglio	95,6	luglio	97,4		
agosto	101,4	agosto	97,5	agosto	97,6		
settembre	100,6	settembre	95,8	settembre	95,9		
ottobre	101,0	ottobre	96,0	ottobre	94,9		
novembre	100,4	novembre	96,9	novembre	95,6		
dicembre	99,0	dicembre	97,1	dicembre	96,8		

Tabella 9: Serie storica del PIL in Italia da feb'08 a feb'11 (*Fonte Istat*)

Soluzione: È richiesto di ricostruire un modello generatore per componenti di tipo additivo per cui partendo dal calcolo del

Analisi delle Serie storiche

trend-ciclo di prima approssimazione, rileviamo che la serie storica ha una cadenza mensile e pertanto l'ordine delle medie mobili centrate è $k = 12$. Calcoliamo le medie mobili centrate di ordine 12 e perdiamo 6 dati iniziali e 6 dati finali che essendo pari determinerà una procedura per la costruzione delle $MM S_t(12)$ e poi la centratura. Calcoliamo poi la componente stagionale ed accidentale congiunta per differenza tra i dati originali e le medie mobili centrate. Da qui si prosegue per la determinazione della componente stagionale prima come media dei dati della componente congiunta $(S + E)_t$ nei periodi k e poi corretta, cioè sottraendo ad ogni dato la medie dei valori congiunti del periodo k . Si procede a destagionalizzare e conseguentemente a calcolare il trend-ciclo e il pattern sistematico. Infine si calcolano i residui come differenza tra i dati originali e i dati

del pattern sistematico. I calcoli sono riportati in Tab. 10. In Fig. 11 è riportato il timeplot della serie residua che evidenzia oscillazioni non sistematiche e considerevoli variazioni solo nel periodo iniziale avvalorando così la validità del modello realizzato.

Inoltre nelle Fig. 12 e 13 sono riportati i dati destagionalizzati e il pattern sistematico.



Analisi delle Serie storiche

t	mese	indice	MMS(12)	MMC(12)	(S+E) _t	S _t	S _t	D _t	FC _t	FC _t +S _t	E _t = r _t
1	1 febbraio-2008	104,7				-0,1	0,0	104,7			
2	2 marzo	104,6				-3,6	-3,5	108,1	106,7	103,2	1,4
3	3 aprile	105,4				-1,9	-1,8	107,2	105,7	103,8	1,6
4	4 maggio	101,6				-0,2	-0,1	101,7	104,8	104,7	-3,1
5	5 giugno	105,1				-0,5	-0,4	105,5	103,0	102,6	2,5
6	6 luglio	102,6				0,6	0,7	101,9	102,6	103,3	-0,7
7	7 agosto	101,4	102,0	101,6	-0,2	1,0	1,1	100,3	100,9	102,0	-0,6
8	8 settembre	100,6	101,3	100,7	-0,1	0,0	0,1	100,5	100,4	100,5	0,1
9	9 ottobre	101,0	100,1	99,5	1,5	0,6	0,7	100,3	100,0	100,8	0,2
10	10 novembre	100,4	99,0	98,7	1,7	1,0	1,1	99,3	99,3	100,4	0,0
11	11 dicembre	99,0	98,3	97,9	1,1	0,7	0,7	98,3	97,8	98,5	0,5
12	12 gennaio -2009	97,2	97,4	97,1	0,1	1,4	1,5	95,7	96,8	98,2	-1,0
13	1 febbraio	96,3	96,8	96,7	-0,4	-0,1	0,0	96,3	95,2	95,2	1,1
14	2 marzo	90,1	96,5	96,3	-6,2	-3,6	-3,5	93,6	94,8	91,3	-1,2
15	3 aprile	92,5	96,1	95,9	-3,4	-1,9	-1,8	94,3	94,0	92,1	0,4
16	4 maggio	93,8	95,7	95,6	-1,8	-0,2	-0,1	93,9	94,3	94,2	-0,4
17	5 giugno	94,2	95,4	95,3	-1,1	-0,5	-0,4	94,6	94,5	94,1	0,1
18	6 luglio	95,6	95,3	95,4	0,2	0,6	0,7	94,9	95,3	96,0	-0,4
19	7 agosto	97,5	95,5	95,5	2,0	1,0	1,1	96,4	95,7	96,7	0,8
20	8 settembre	95,8	95,6	95,8	0,0	0,0	0,1	95,7	95,8	95,9	-0,1
21	9 ottobre	96,0	96,1	96,2	-0,2	0,6	0,7	95,3	95,6	96,3	-0,3
22	10 novembre	96,9	96,4	96,6	0,3	1,0	1,1	95,8	95,8	96,9	0,0
23	11 dicembre	97,1	96,8	96,9	0,2	0,7	0,7	96,4	96,8	97,6	-0,5
24	12 gennaio-2010	99,8	97,0	97,1	2,7	1,4	1,5	98,3	97,3	98,8	1,0
25	1 febbraio	97,3	97,2	97,2	0,1	-0,1	0,0	97,3	98,5	98,4	-1,1
26	2 marzo	96,2	97,2	97,2	-1,0	-3,6	-3,5	99,7	98,5	95,0	1,2
27	3 aprile	96,7	97,2	97,2	-0,5	-1,9	-1,8	98,5	98,9	97,1	-0,4
28	4 maggio	98,4	97,1	97,1	1,3	-0,2	-0,1	98,5	98,2	98,1	0,3
29	5 giugno	97,1	97,0	97,0	0,1	-0,5	-0,4	97,5	97,6	97,2	0,0
30	6 luglio	97,4	97,0	96,8	0,6	0,6	0,7	96,7	96,9	97,6	-0,2
31	7 agosto	97,6	96,6	96,5	1,1	1,0	1,1	96,5	96,4	97,4	0,2
32	8 settembre	95,9	96,5			0,0	0,1	95,8	95,5	95,6	0,3
33	9 ottobre	94,9				0,6	0,7	94,2	94,8	95,6	-0,7
34	10 novembre	95,6				1,0	1,1	94,5	94,9	96,0	-0,4
35	11 dicembre	96,8				0,7	0,7	96,1	94,6	95,3	1,5
36	12 gennaio-2011	94,7				1,4	1,5	93,2	95,1	96,6	-1,9
37	1 febbraio	96,1				-0,1	0,0	96,1			
						-0,1	0,0				

Tabella 10: Calcoli dell'Esercizio 1

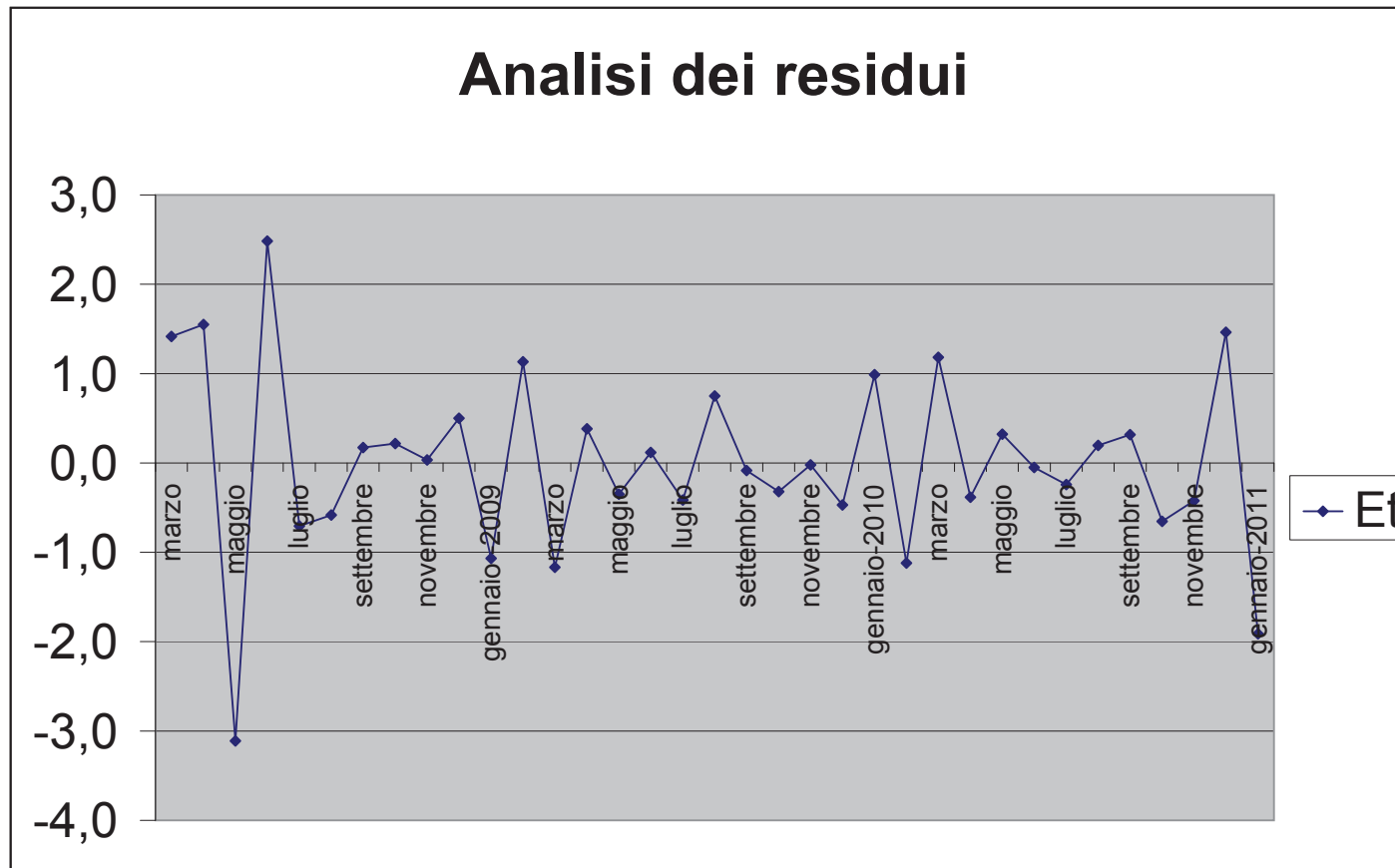


Figura 11: Serie dei residui dell'Esercizio 1

Analisi delle Serie storiche

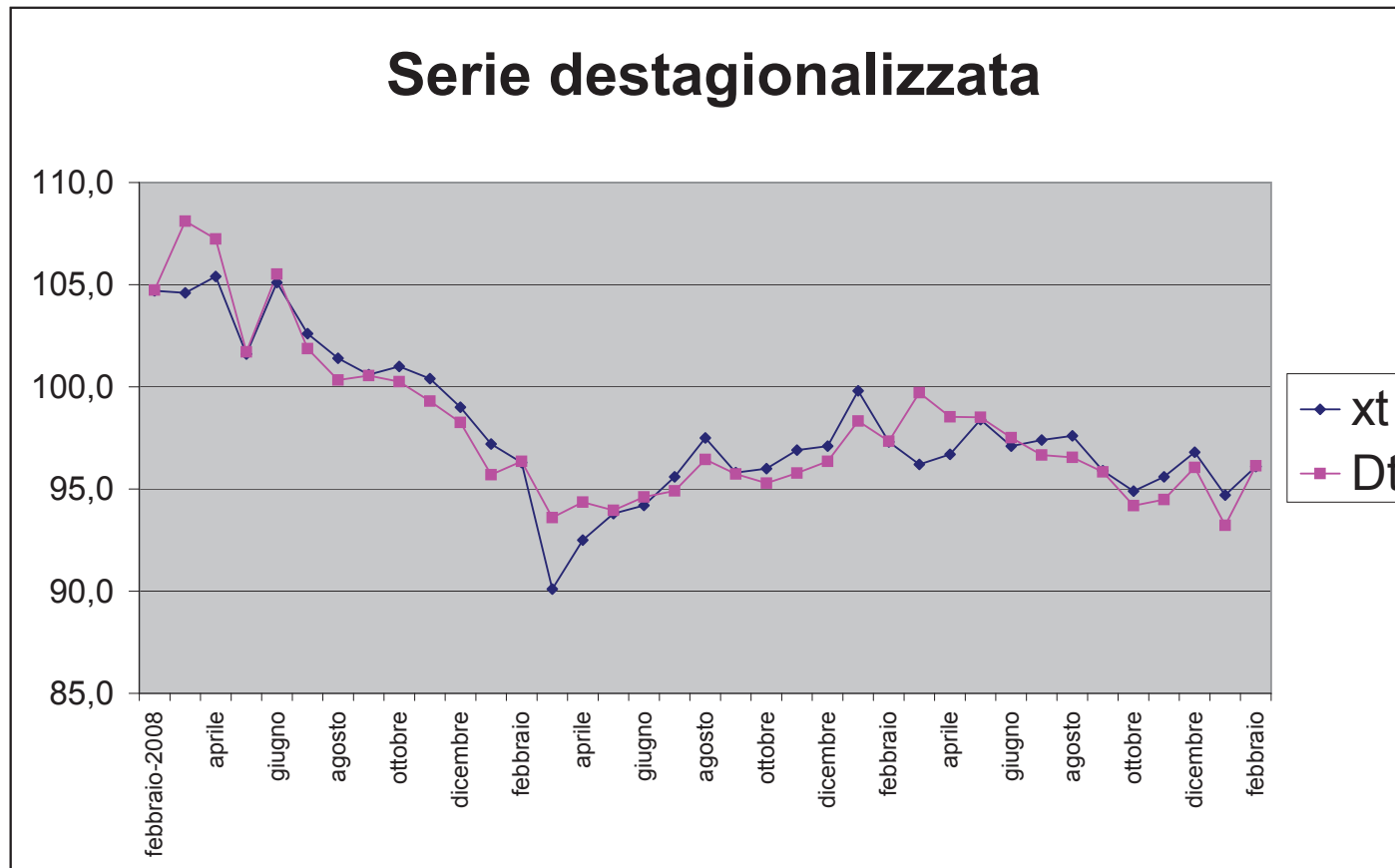


Figura 12: Serie destagionalizzata dell'Esercizio 1

Analisi delle Serie storiche

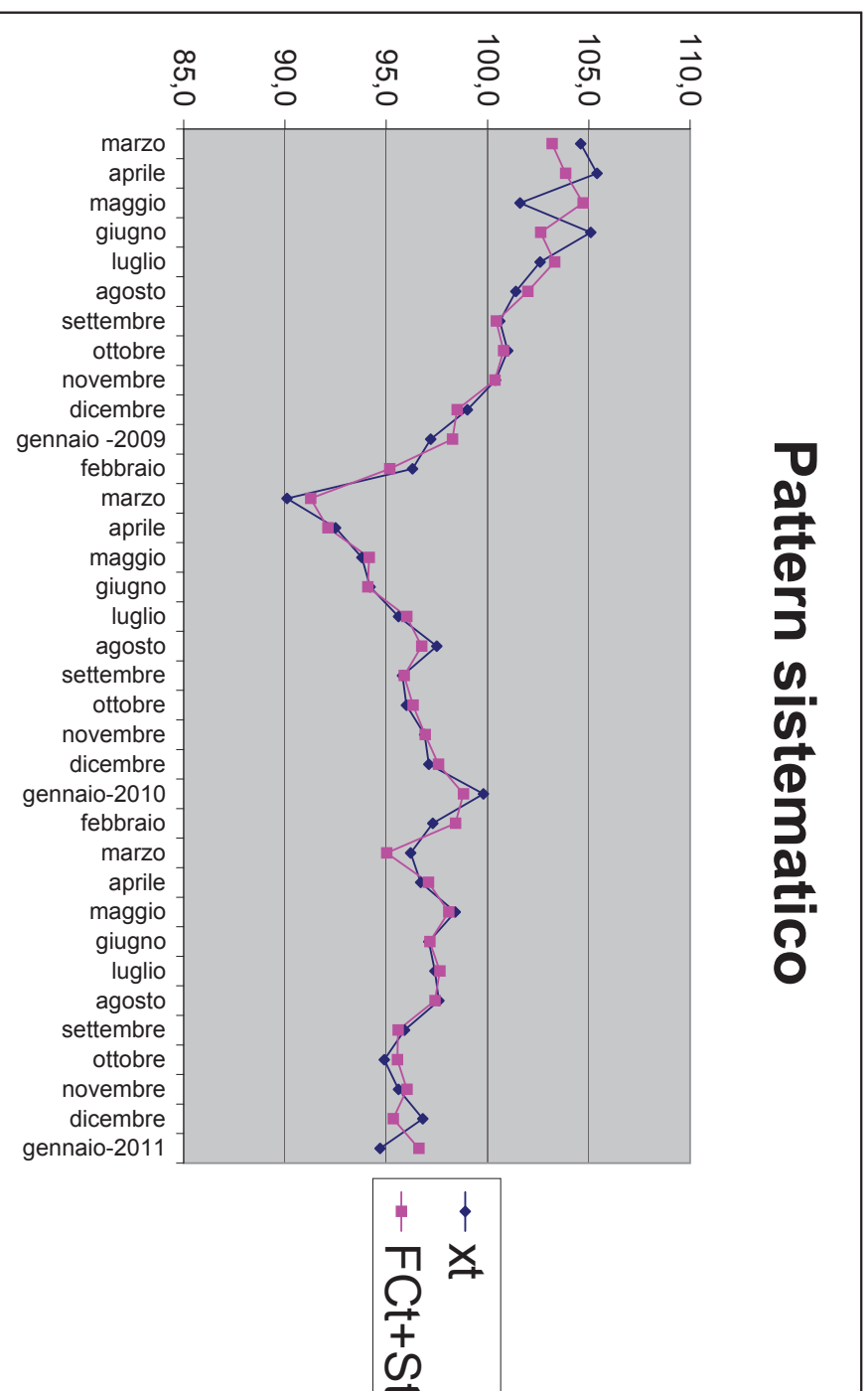


Figura 13: Pattern sistematico dell'Esercizio 1

★★★☆☆ **Esercizio 2.** Con i dati dell'Esercizio 1 ricostruire il modello generatore per componenti di tipo misto.

Soluzione: Occorre procedere come nell'Esercizio 1 utilizzando delle (12), (13), (14) (15), (16), (17) e (18) le formule di sinistra. I calcoli sono riportati in Tab. 11. In Fig. 14 è riportato il timeplot della serie residua che evidenzia oscillazioni non sistematiche e considerevoli variazioni solo nel periodo iniziale avvalorando così la validità del modello realizzato.

Inoltre nelle Fig. 15 e 16 sono riportati i dati destagionalizzati e il pattern sistematico. □

Se mettiamo a confronto i residui ottenuti con il modello additivo e con il modello misto (vedi Fig. 17) osserviamo che quelli del misto

subiscono variazioni contenute con picchi nel periodo iniziale (che possiamo accettare), e le serie destagionalizzate (vedi Fig. 18) differiscono in modo costante per tutto il periodo di osservazione.

★☆☆ **Esercizio 4.** Valutare gli indici MSE e MSEP delle analisi dell'Esercizio 1 e dell'Esercizio 2.

★★★☆☆ **Esercizio 3.** Condurre l'analisi classica con modello generatore additivo e misto della serie storica di dati in Tab. 2.

Analisi delle Serie storiche

t	mese	indice	MMS(12)	MMC(12)	(SE) _t	S _t	St	D _t	FC _t	FC _t -S _t	Et = r _t
1	1 febbraio-2008	104,7				1,00	1,00	104,8			
2	2 marzo	104,6				0,94	0,94	111,5	108,4	109,3	-4,7
3	3 aprile	105,4				0,96	0,97	109,0	107,9	108,8	-3,4
4	4 maggio	101,6				0,98	0,98	103,2	106,1	107,0	-5,4
5	5 giugno	105,1				0,99	0,99	106,0	103,8	104,7	0,4
6	6 luglio	102,6				1,00	1,01	102,0	102,7	103,7	-1,1
7	7 agosto	101,4	102,0								
8	8 settembre	100,6	101,3	101,6	1,00	1,01	1,01	100,1	100,8	101,8	-0,4
9	9 ottobre	101,0	100,1	100,7	1,00	1,00	1,00	100,3	100,2	101,2	-0,6
10	10 novembre	100,4	99,0	99,5	1,01	1,01	1,01	100,1	99,8	100,8	0,2
11	11 dicembre	99,0	98,3	98,7	1,02	1,01	1,01	99,1	99,1	100,1	0,3
12	12 gennaio -2009	97,2	97,4	97,9	1,01	1,01	1,01	98,0	97,6	98,6	0,4
13	1 febbraio	96,3	97,1	97,1	1,00	1,01	1,02	95,6	96,6	97,6	-0,4
14	2 marzo	90,1	96,8	96,7	1,00	1,00	1,00	96,1	95,0	96,0	0,3
15	3 aprile	92,5	96,5	96,3	0,94	0,96	0,97	93,3	94,5	95,5	-5,4
16	4 maggio	93,8	96,1	95,9	0,96	0,98	0,98	94,1	93,7	94,7	-2,2
17	5 giugno	94,2	95,7	95,6	0,98	1,00	1,00	93,7	94,1	95,1	-1,3
18	6 luglio	95,6	95,4	95,3	0,99	0,99	1,00	94,4	94,3	95,3	-1,1
19	7 agosto	97,5	95,3	95,4	1,00	1,01	1,01	94,7	95,1	96,1	-0,5
20	8 settembre	95,8	95,5	95,5	1,02	1,01	1,01	96,2	95,5	96,5	1,0
21	9 ottobre	96,0	95,6	95,8	1,00	1,00	1,00	95,5	95,6	96,6	-0,8
22	10 novembre	96,9	96,1	96,2	1,00	1,01	1,01	95,1	95,4	96,4	-0,4
23	11 dicembre	97,1	96,4	96,6	1,00	1,01	1,01	95,6	95,6	96,6	0,3
24	12 gennaio-2010	99,8	96,8	96,9	1,00	1,01	1,01	96,2	96,6	97,6	-0,5
25	1 febbraio	97,3	97,0	97,1	1,03	1,01	1,02	98,1	97,1	98,2	1,6
26	2 marzo	96,2	97,2	97,2	1,00	1,00	1,00	97,1	98,3	99,3	-2,0
27	3 aprile	96,7	97,2	97,2	0,99	0,96	0,97	99,6	98,4	99,4	-3,2
28	4 maggio	98,4	97,2	97,2	1,00	0,98	0,98	98,4	98,8	99,8	-3,1
29	5 giugno	97,1	97,1	97,1	1,01	1,00	1,00	98,3	98,0	99,0	-0,6
30	6 luglio	97,4	97,0	97,0	1,00	0,99	1,00	97,3	97,4	98,4	-1,3
31	7 agosto	97,6	97,0	96,8	1,01	1,01	1,01	96,5	96,7	97,7	-0,3
32	8 settembre	95,9	96,6	96,5	1,01	1,01	1,01	96,3	96,2	97,2	0,4
33	9 ottobre	94,9	96,5			1,00	1,00	95,6	95,3	96,3	-0,4
34	10 novembre	95,6				1,01	1,01	94,0	94,7	95,7	-0,8
35	11 dicembre	96,8				1,01	1,01	94,3	94,7	95,8	-0,2
36	12 gennaio-2011	94,7				1,01	1,01	95,9	94,4	95,4	1,4
37	1 febbraio	96,1				1,01	1,02	93,1	95,0	96,0	-1,3
						1,00	1,00	95,9			
						1,00	1,00				

Tabella 11: Calcoli dell'Esercizio 1

Analisi delle Serie storiche

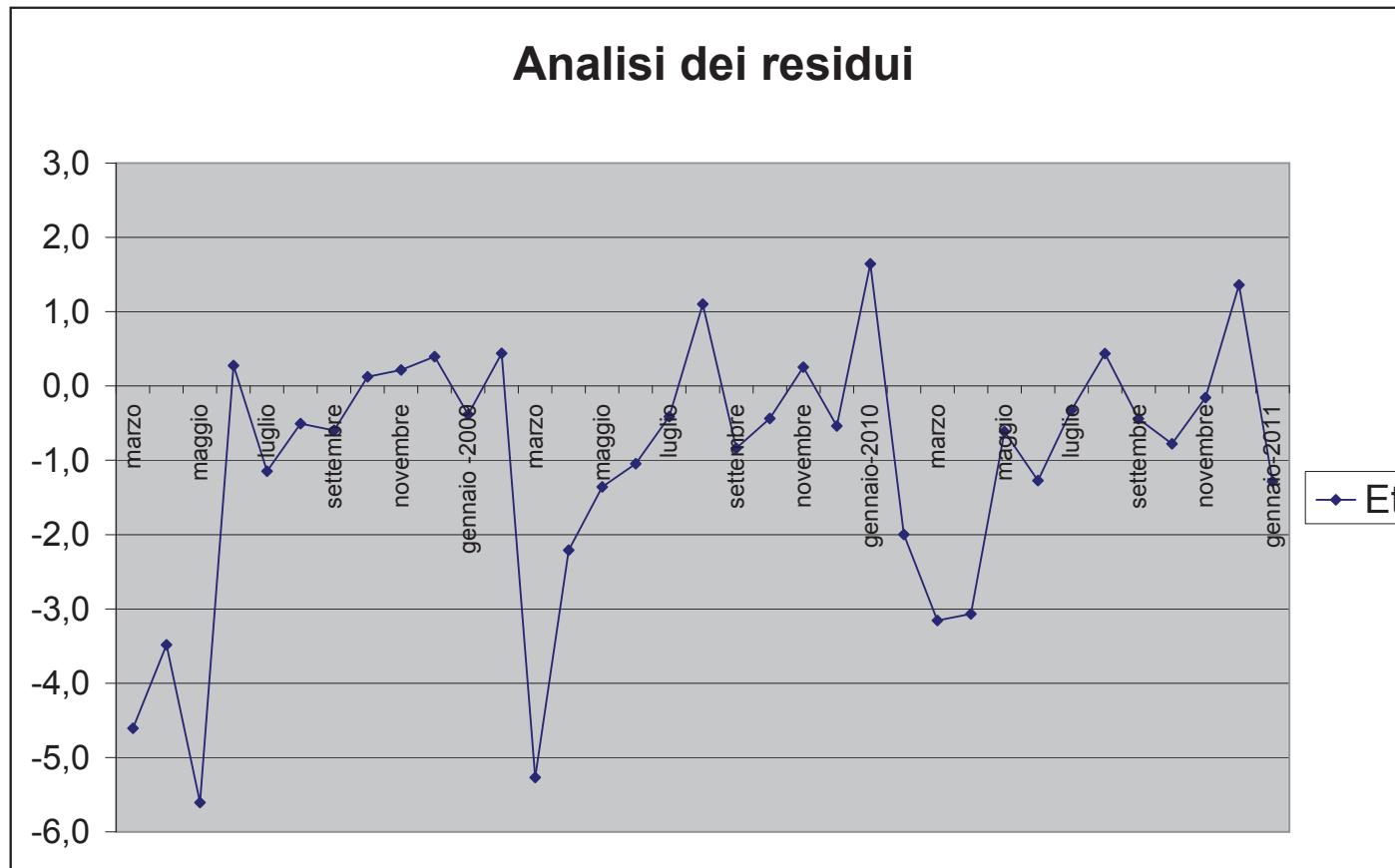


Figura 14: Serie dei residui dell'Esercizio 2

Analisi delle Serie storiche

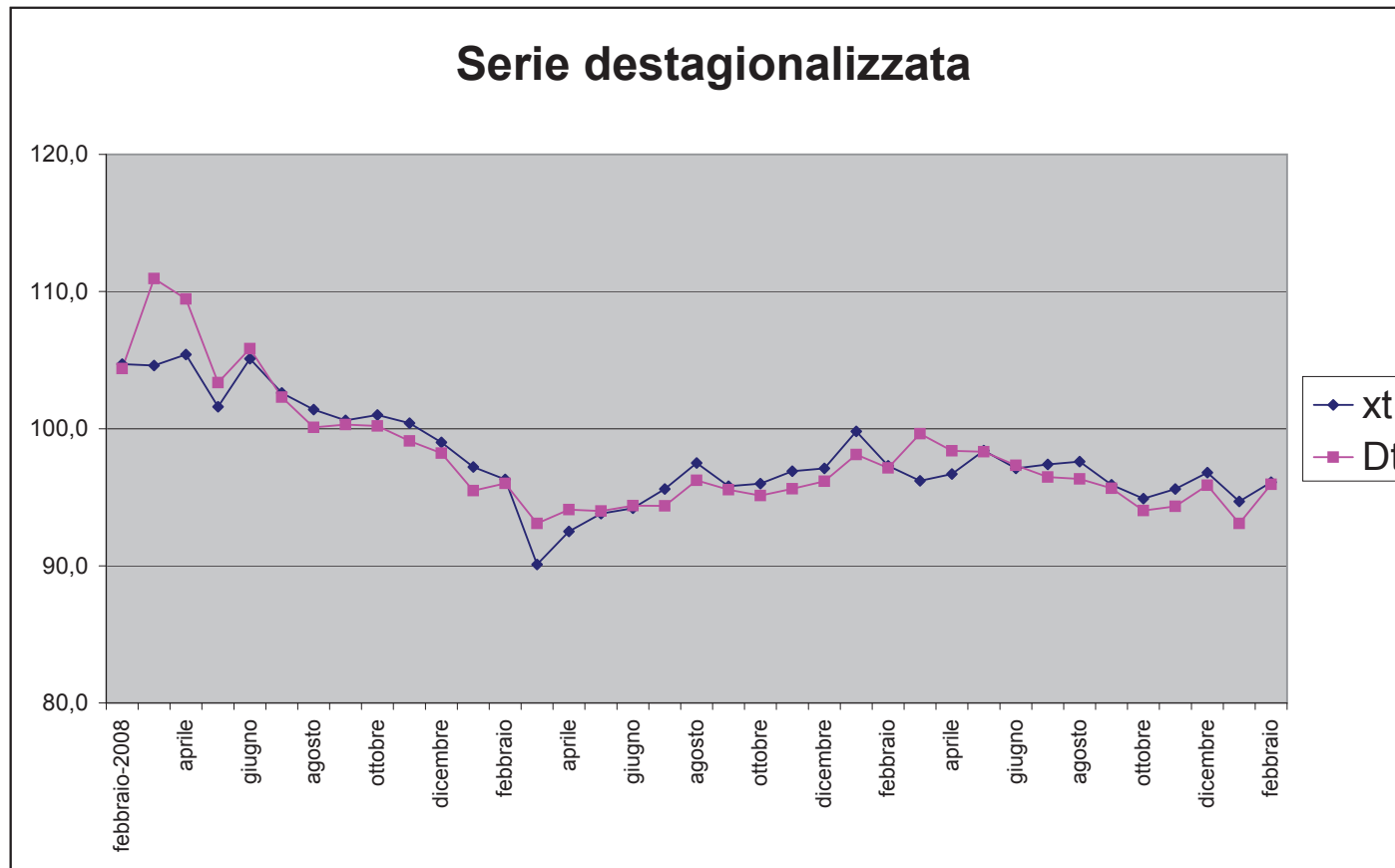


Figura 15: Serie destagionalizzata dell'Esercizio 2

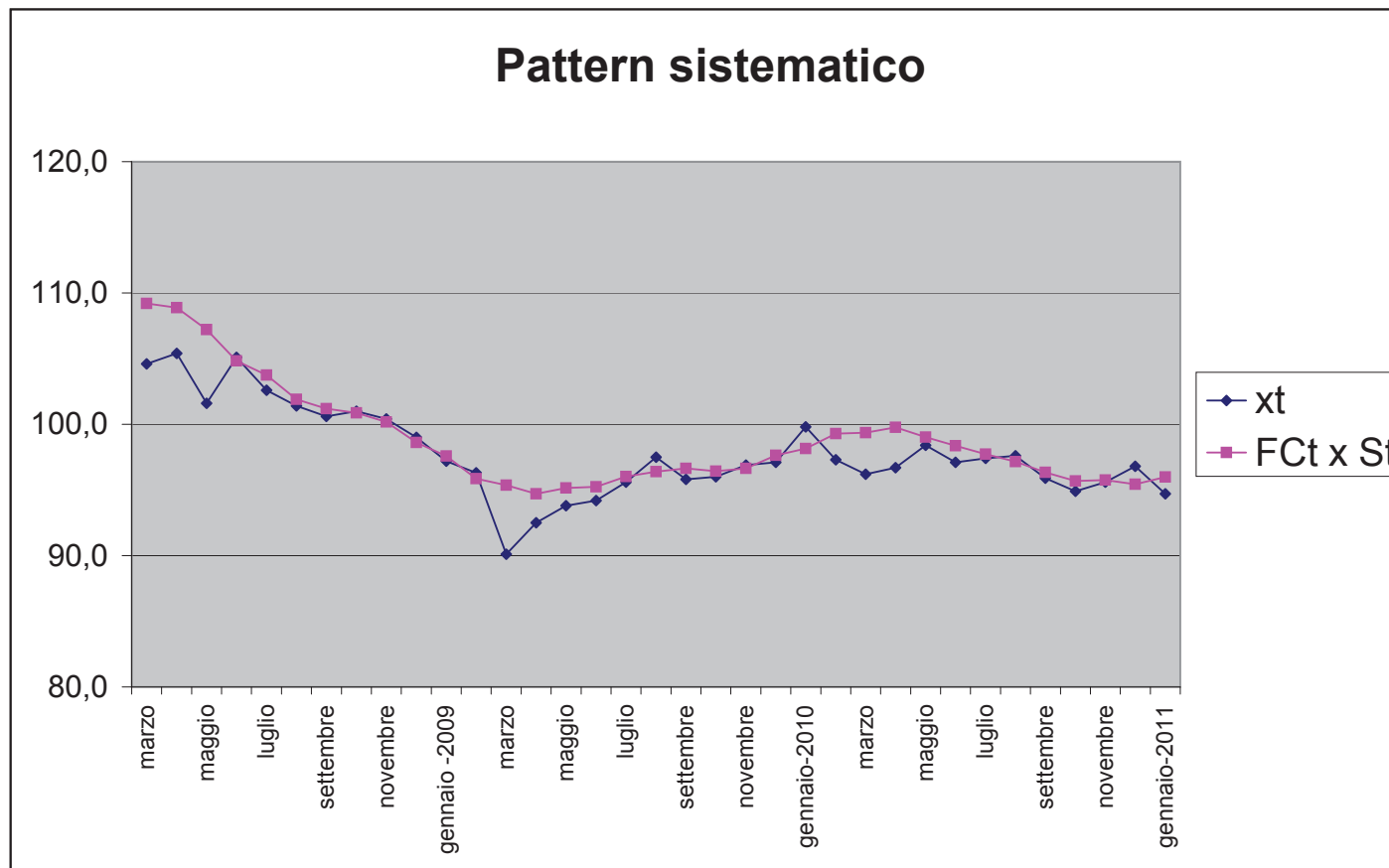


Figura 16: Pattern sistematico dell'Esercizio 2

Analisi delle Serie storiche

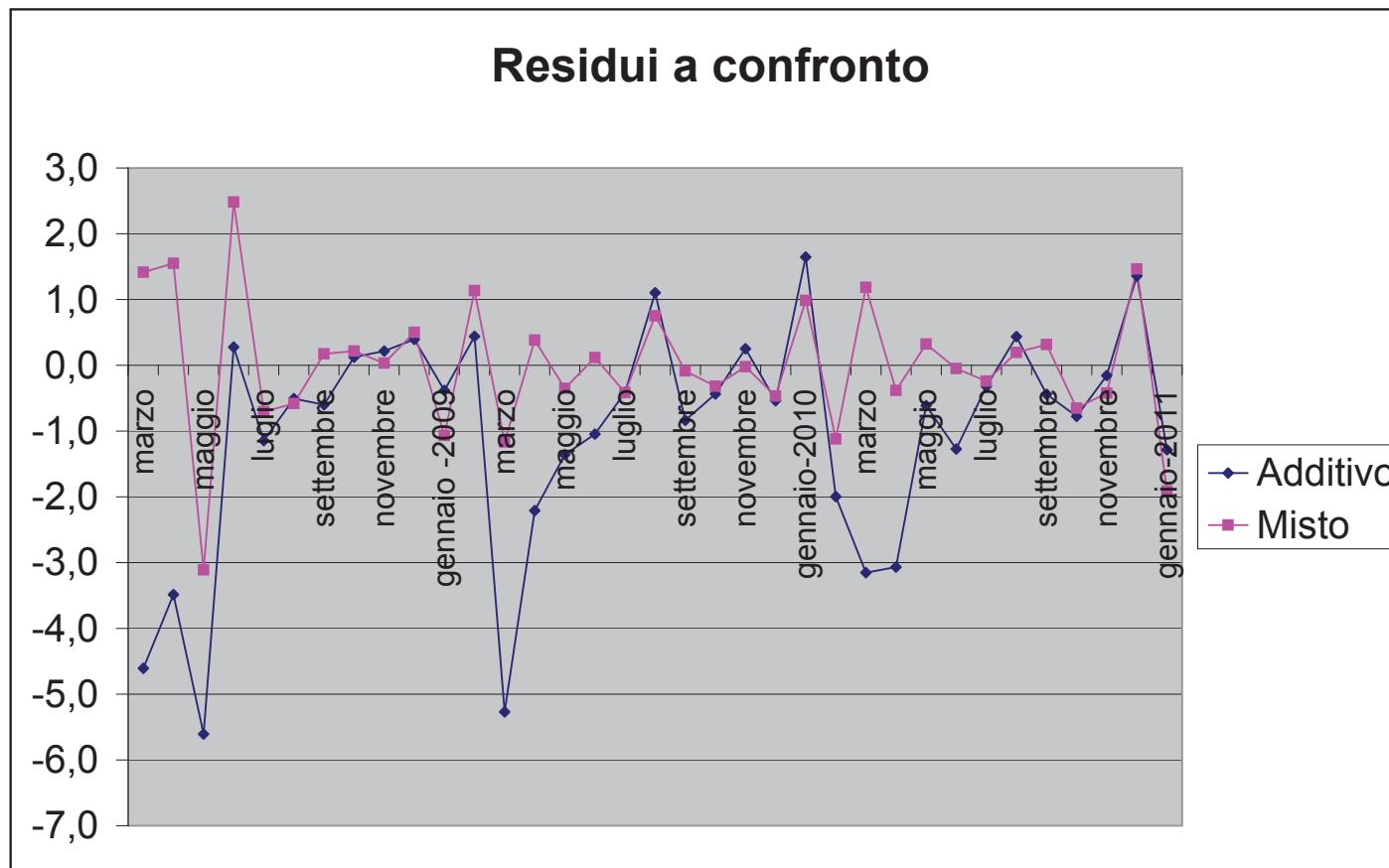


Figura 17: Residui dell'Esercizio 1 e dell'Esercizio 2

Analisi delle Serie storiche

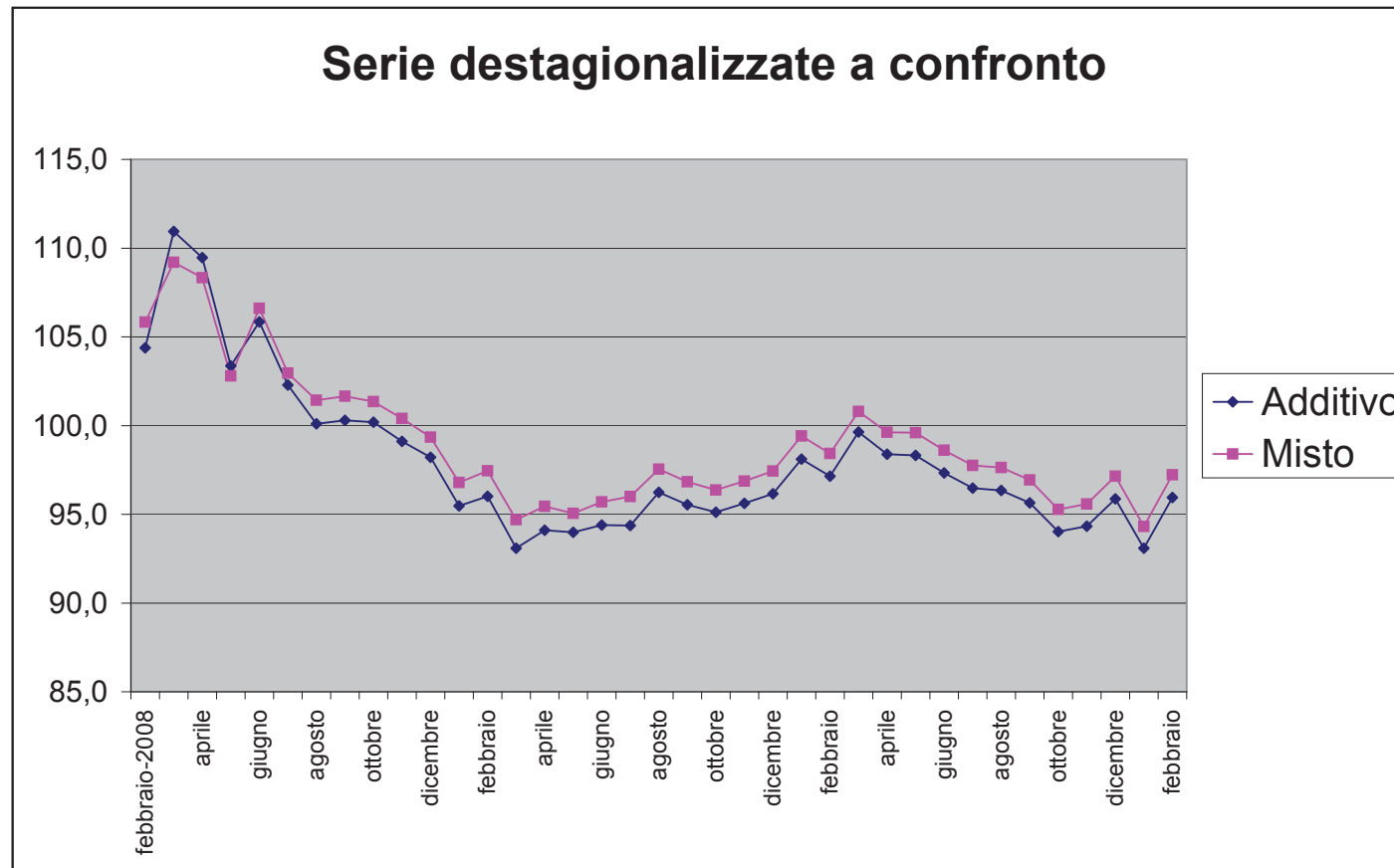


Figura 18: Serie destagionalizzate dell'Esercizio 1 e dell'Esercizio 2

8 – Il Processo di composizione senza stagionalità

Supponiamo che la serie x_t con $t \in [1, T]$ non sia soggette ad oscillazioni stagionali. Il modello generatore per componenti presenta come pattern sistematico il solo trend-ciclo FC_t a cui si aggiunge^a la componente accidentale E_t . Poichè l'approccio classico pone la massima attenzione sulla parte deterministica, in tal caso, allora l'analisi si concentrerà esclusivamente sulla stima di FC_t . Esistono due procedure per stimare il trend-ciclo e poi valutare il modello generatore:

- **metodo analitico** secondo cui costruiamo una funzione matematica del tempo valida per l'orizzonte temporale cui la

^a Ricordiamo che consideriamo soltanto il modello additivo e misto.

serie si riferisce.

- **metodo delle medie ponderate** secondo cui individuiamo un polinomio interpolante valido per un gruppo di osservazioni successive.

8.1 – Il metodo analitico

Il metodo determina il *trend-ciclo globale* e si articola in tre fasi sequenziali:

1. *scelta della forma della funzione*

Lo strumento utilizzato per individuare la funzione matematica del tempo è le *differenze successive*.

Data la serie storica x_t definiamo:

- *differenza prima* la serie storica i cui termini sono le differenze

dei termini consecutivi della serie storica data:

$$\Delta_1(x_t) = x_t - x_{t-1} \quad t = 2, 3, \dots, T$$

- *differenza seconda* la serie storica i cui termini sono le differenze dei termini consecutivi della differenza prima:

$$\begin{aligned}\Delta_2(x_t) &= \Delta_1(x_t) - \Delta_1(x_{t-1}) = \\ &= (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = \\ &= x_t - 2x_{t-1} + x_{t-2}\end{aligned}$$

- *differenza terza* la serie storica i cui termini sono le differenze

dei termini consecutivi della differenza seconda:

$$\begin{aligned}\Delta_3(x_t) &= \Delta_2(x_t) - \Delta_2(x_{t-1}) = \\ &= (\Delta_1(x_t) - \Delta_1(x_{t-1})) - (\Delta_1(x_{t-1}) - \Delta_1(x_{t-2})) = \\ &= [(x_t - x_{t-1}) - (x_{t-1} - x_{t-2})] - \\ &\quad - [(x_{t-1} - x_{t-2}) - (x_{t-2} - x_{t-3})] = \\ &= x_t - 2x_{t-1} + x_{t-2} - x_{t-1} + 2x_{t-2} - x_{t-3} = \\ &= x_t - 3x_{t-1} + 3x_{t-2} - x_{t-3}\end{aligned}$$

• ...

La scelta della funzione dipende dalla condizione imposta sul valore della differenza successiva come riportato in Tab. 12.

Analisi delle Serie storiche

<i>condizione</i>	<i>funzione</i>	<i>equazione</i>
$\Delta_1(x_t) = k$ ^b	lineare	$f(t) = \beta_0 + \beta_1 t$
$\Delta_2(x_t) = k$	quadratica	$f(t) = \beta_0 + \beta_1 t + \alpha_2 t^2$
${}_{t-1}I_t = k$	esponenziale semplice	$f(t) = \beta_0 e^{\beta_1 t}$
${}_{t-1}I_{\Delta,t} = k$	esponenziale modificata	$f(t) = \gamma + \beta_0 e^{\beta_1 t}$
$\Delta_1(\log(x_t)) = k$	curva di Gompertz	$f(t) = \frac{1}{\gamma + \beta_0 \beta_1 t}$

Tabella 12: Schema riassuntivo della funzione scelta con il metodo analitico

^b k è una costante

2. *stima dei parametri della funzione*

Nella funzione sono presenti dei parametri $\beta_0, \beta_1, \dots, \gamma$ che possono essere determinati:

- con il *metodo dei minimi quadrati* se f è lineare o quadratica;
- con il *metodo di linearizzazione* se f è esponenziale semplice;
- con le *tecniche di stima non lineari* se f è esponenziale modificata o di Gompertz.

3. *verifica della validità del modello scelto*

Dalla funzione f al variare di t otteniamo i dati teorici \bar{x}_t . Per verificare la bontà di adattamento della funzione basta procedere con l'analisi dei residui $r_t = x_t - \bar{x}_t$.

8.2 – Il metodo delle medie ponderate

Il metodo determina il *trend-ciclo locale* e si articola in tre fasi sequenziali:

1. *realizzazione di sottoserie di eguale lunghezza*

Data la serie storica x_t con $t \in [1, T]$ e preso $r = 2m + 1$ ovvero dispari^c costruiamo le sottoserie di lunghezza r formate da elementi successivi

$$x_{s-m}, x_{s-m+1}, \dots, x_s, \dots, x_{s+m-1}, x_{s+m}$$

2. *determinazione del polinomio interpolante*

Costruiamo il polinomio

$$\beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_p t^n$$

^c Henderson e Spenser proposero rispettivamente $m = 4, 6, 11$ ($r = 9, 13, 23$) e $m = 7, 10$ ($r = 15, 21$).

interpolante di grado $n < r - 1 = 2m$ che passi tra i termini del gruppo ovvero tale che sia minimo la somma dei quadrati delle differenze tra il valore della serie all'istante t ed il rispettivo valore teorico calcolato nel polinomio

$$\inf_{(\beta_0, \beta_1, \dots, \beta_n) \in \mathbb{R}^{n+1}} \sum_{t=s-m}^{s+m} (x_t - \beta_0 - \beta_1 t - \beta_2 t^2 - \dots - \beta_n t^n)^2 \quad \text{d}$$

3. *verifica della validità del modello scelto*

Si procede come nel metodo analitico.

★★☆ **Esercizio 4.** Stimare il FC_T con il metodo analitico della serie storica del PIL (in milioni di €) dal 1998 al 2007 (*Fonte: Istat*).

^d Per $t = 0$ (se $s = m$) il valore del polinomio interpolante è β_0 per cui $FC_0 = \alpha_0$ ed è pari ad una media ponderata con pesi che dipendono da r e da n .

Soluzione: La rappresentazione grafica della serie storica in oggetto in Fig. 19 evidenzia un andamento regolare da far ritenere la sola presenza del trend-ciclo. Inoltre visivamente appare chiaro che la funzione stima di FC_t è di tipo lineare. I calcoli per determinare i parametri della funzione $f(t)$ sono riportati in Tab. 13 unitamente all'analisi dei residui, che con un valore di \mathcal{R}^2 pari a 0.999973803 conferma l'ottimo adattamento dei dati teorici del modello $f(t) = \alpha_0 + \alpha_1 t$ ai dati reali x_t . Infine in Fig. 20 sono riportate contestualmente la serie data e la stima ottenuta. □

Analisi delle Serie storiche

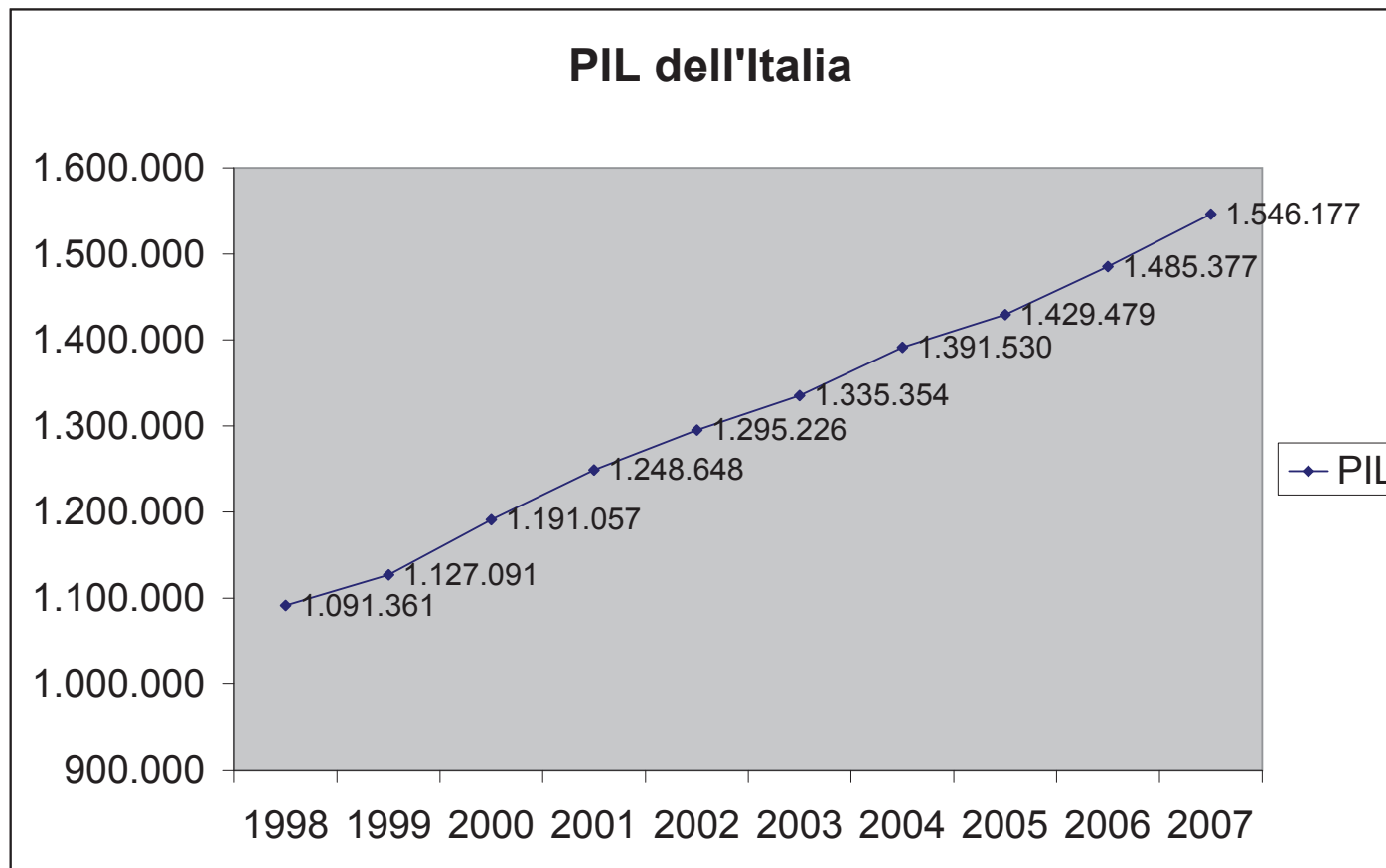


Figura 19: Serie storica del PIL in Italia dal 1998 al 2007 (*Fonte: Istat*)

Analisi delle Serie storiche

anno	x_t	$\text{anno} \cdot x_t$	anno^2	x_t teorico	residui	residui^2	x_t^2
1998	1.091.361	2180540259	3992004	1088796,592	2.565	6.578.705	1.191.069.904.038
1999	1.127.091	2253055079	3996001	1138870,718	- 11.780	138.759.765	1.270.334.313.886
2000	1.191.057	2382114640	4000000	1188944,845	2.112	4.462.552	1.418.617.539.478
2001	1.248.648	2498544854	4004001	1239018,971	9.629	92.720.189	1.559.122.084.851
2002	1.295.226	2593041888	4008004	1289093,097	6.133	37.609.048	1.677.609.661.268
2003	1.335.354	2674713505	4012009	1339167,223	- 3.814	14.542.789	1.783.169.562.325
2004	1.391.530	2788626444	4016016	1389241,349	2.289	5.238.664	1.936.356.190.364
2005	1.429.479	2866105920	4020025	1439315,475	- 9.836	96.751.082	2.043.410.960.517
2006	1.485.377	2979666949	4024036	1489389,601	- 4.012	16.098.215	2.206.345.850.266
2007	1.546.177	3103178025	4028049	1539463,727	6.714	45.073.293	2.390.664.526.233
2002,5	1.314.130	26319587563	40100145		0	457.834.303	17.476.700.593.226
		413111,5399	8,25				20.731.982.755

$\alpha_1 =$ 50074,12605

$\alpha_2 =$ 0,999973803

$\alpha_0 =$ -98959307,26

Tabella 13: Calcoli per la scelta del tipo di funzione approssimante

Analisi delle Serie storiche

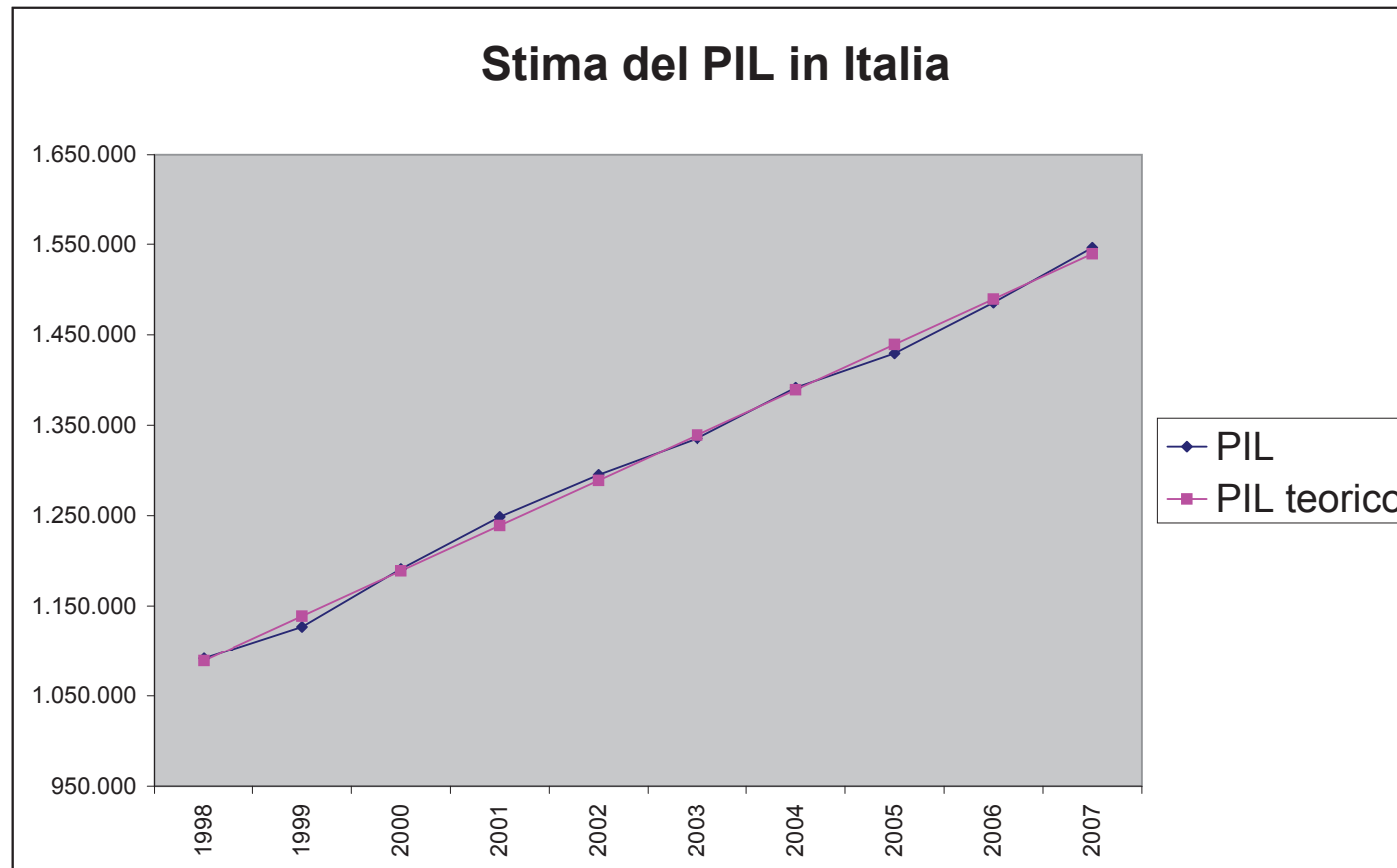


Figura 20: Stima della serie storica PIL in Italia dal 1998 al 2007 mediante la funzione lineare $f(t) = -98959307.26 + 50074.12605t$.

Osservazione 3. Come visto nell'esempio precedente, l'analisi grafica suggerisce la scelta del modello analitico che stimerà il trend globale senza ricorrere alle differenze successive (vedi Tab. 14).

Nell'esempio il modello era di tipo lineare e con il coefficiente di determinazione \mathcal{R}^2 abbiamo valutato la bontà di adattamento^e. In altri modelli analitici stimatori però i parametri sono più di due (ad es. sono tre nella quadratica, esponenziale modificata e nella curva di Gompertz) come possiamo allora valutare la bontà di adattamento del modello?

Poichè il fitting, ovvero la dissomiglianza tra reale e teorico, diminuisce all'aumentare del numero delle variabili esplicative (parametri) allora il coefficiente che tiene conto di tutti i parametri

^e Nel caso lineare vedi [Correlazione&Regressione.pdf](#)

considerati è l' \mathcal{R}^2 corretto definito dalla formula:

$$\mathcal{R}_{\text{corretto}}^2 = 1 - \frac{\frac{Dev(R)}{n-p}}{\frac{Dev(Y)}{n-1}}$$

dove con n e p indichiamo rispettivamente il numero delle osservazioni e il numero dei parametri della funzione $f(t)$.

Analisi delle Serie storiche

<i>andamento grafico</i>	<i>funzione analitica</i>
alterno rispetto ad un dato base	costante
con incrementi q. costanti	lineare
con incrementi di incrementi q. costanti	quadratico
molto accelerato	esponenziale
a forma di esse allungata	di Gompertz

Tabella 14: Modello analitico suggerito dal grafico

9 – La previsione nell'Analisi Classica

Il metodo di scomposizione della serie storica x_t con $t \in [1, T]$

nell'approccio classico perviene alla costruzione del pattern sistematico che abbiamo indicato con \bar{x}_t . Per effettuare una previsione a breve periodo, ad es. per $T + 1$, si dovrà costruire il pattern sistematico \bar{x}_t con $t \in [1, T + 1]$ e ciò oltre a non essere semplice perchè si dovrà prevedere il trend-ciclo e la stagionalità, non produce risultati soddisfacenti.

In dettaglio:

- *per la componente trend-ciclo*: il valore FC_{T+1} si otterrà estrapolando la retta di regressione i cui punti hanno coordinate (FC_{t-1}, FC_t) per $t = 2, 3, \dots, T$;
- *per la stagionalità*: è più semplice data la periodicità k basterà risolvere l'equazione $T + 1 = j \cdot t + k$ nell'incognita j al variare di t e scegliere $S_{T+1} = S_t$;

Se la previsione interessa l'intero pattern sistematico al periodo temporale $T + 1$ allora si avrà (a sinistra per il modello additivo ed a destra per il modello misto):

$$\bar{x}_{T+1} = FC_{T+1} + S_{T+1} \qquad \bar{x}_{T+1} = FC_{T+1} \cdot S_{T+1}$$

Nel caso in cui la serie storica è priva di stagionalità, si esegue solo la previsione del trend-ciclo con il metodo della retta di regressione su proposto.

★★★**Esercizio 5.** Effettuare la previsione della serie storica del PIL i cui dati sono riportati in Tab. 1 per il mese di maggio 2011.

10 – Limiti dell'Analisi Classica

L'approccio classico basato sul metodo di scomposizione del modello generatore ha i seguenti limiti:

Analisi delle Serie storiche

1. *è un'analisi descrittiva e non interpretativa*

Non è possibile osservare le singole componenti per cui:

- quali sono le componenti presenti e quali quelle assenti?
- come testare i singoli risultati ottenuti?

2. *le componenti identificate dipendono dal metodo scelto per identificarle*

Non è realizzabile alcuna verifica senza premettere delle ipotesi vincolanti per cui

- Come è verificabile una condizione non prevista?
- Se cambiando tecnica muta la componente quale risulta essere più adatta a descrivere la serie?

3. *esistono numerose scomposizioni ed ognuna è ugualmente valida*

La scomposizione della serie dipende dal modello (additivo, moltiplicativo o misto) per cui

- scelto un modello come effettuare un'analisi di una variazione di trend che il modello non contempla?