MODELLI DI SERIE STORICHE

APPROCCIO CLASSICO

Un modello stocastico generale per descrivere il processo generatore dei dati di una serie storica $y_1, y_2, \ldots, y_n$ relativa ad una variabile Y è dato da

$$Y_t = f(t) + u_t$$

Si assume che la serie osservata sia il risultato della composizione di:

a) una sequenza completamente deterministica, $\{f(t)\}$, che costituisce la parte sistematica della serie;
b) una sequenza di variabili casuali $\{u_t\}$, che rappresenta la parte stocastica della serie ed obbedisce ad una determinata legge di probabilità.

Le due sequenze non sono individualmente osservabili, ma vanno determinate sulla base del campione.

Se trattiamo il precedete modello secondo l'approccio *classico* alle serie storiche, si suppone che esista una legge di evoluzione temporale del fenomeno, rappresentata da $f(t)$. La componente casuale u_t, viene assunta a rappresentare l'insieme delle circostanze, ciascuna di entità trascurabile, che non si vogliono o non possiamo considerare in Y_t. I residui di Y_t vengono pertanto imputati al caso ed assimilati ad *errori accidentali*. Questo equivale ad ipotizzare che la componente stocastica u_t sia generata da un processo *white noise*, ovvero da una successione di variabili casuali indipendenti, identicamente distribuite, di media nulla e varianza costante.

In sintesi, nell'approccio classico l'attenzione viene concentrata su $f(t)$, essendo u_t considerato un processo a componenti incorrelate e dunque trascurabile.

Nell'approccio *moderno* si ipotizza invece che $f(t)$ manchi o sia già stata 'eliminata' (mediante stima o altri metodi). L'attenzione viene posta sulla componente stocastica u_t, che si ipotizza essere un processo a componenti correlate del tipo

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

che va trattato con opportune tecniche statistiche.

Obiettivi analisi delle serie storiche:

- descrizione
- spiegazione
- previsione
- filtraggio
- controllo

Modelli di composizione, componenti:

Trend: è la tendenza di fondo del fenomeno considerato, riferita ad un lungo periodo di tempo.
(Ex: PIL italiano, la crescita di lungo periodo del 3%)

Ciclo: è costituito dalle fluttuazioni attribuibili al succedersi nel fenomeno considerato di fasi ascendenti e di fasi discendenti, generalmente collegate con le fasi di espansione e di contrazione dell'intero sistema economico.
(Ex: PIL italiano, fasi di boom economico contrapposte a fasi di recessione)

Stagionalità: è costituita dai 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 (mese o trimestre). (Ex: PIL italiano, nel mese di agosto tutte le grandi fabbriche sono chiuse)

Componente accidentale: come nel modello di regressione, anche nei modelli di serie storiche non vi è mai una relazione perfetta tra la variabile sotto osservazione e le diverse componenti, la componente accidentale tiene conto di questo e del comportamento non perfettamente prevedibile degli agenti economici.

Tipi di composizione:

1) additiva
2) moltiplicativa
3) misto

1) ipotesi di indipendenza tra le componenti modello additivo:

$$\boxed{Z_t = T_t + C_t + S_t + a_t}$$

2) non indipendenza tra le componenti, modello moltiplicativo

$$\boxed{Z_t = T_t \times C_t \times S_t \times a_t}$$

Il caso 2) si riduce al caso 1) considerando i logaritmi, cioè:

$$\boxed{\log Z_t = \log T_t + \log C_t + \log S_t + \log a_t}$$

3) modello misto:

$$\boxed{Z_t = T_t \times S_t + C_t + a_t}$$

Pregi

-semplicità

-serie anche corte

-prima approssimazione

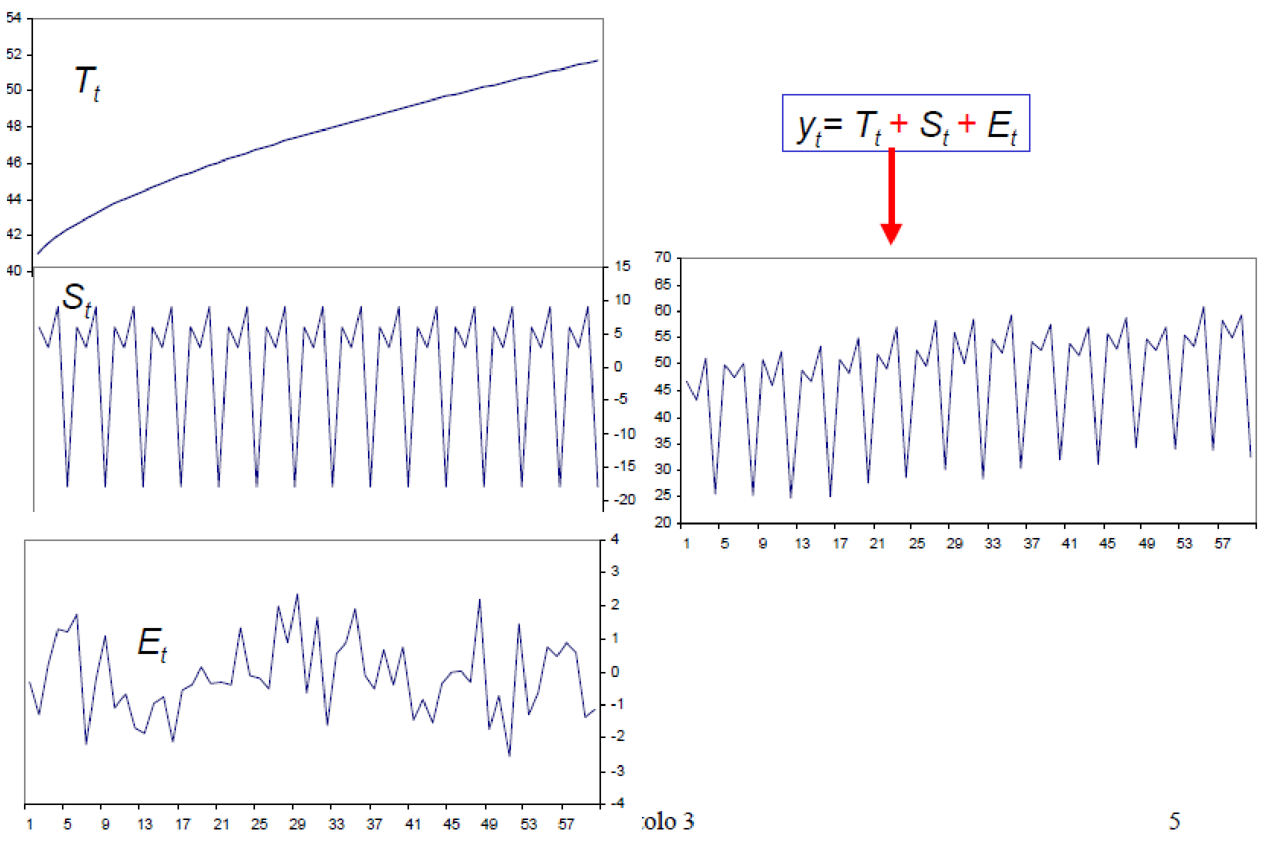
Difetti

-pluralità di soluzioni

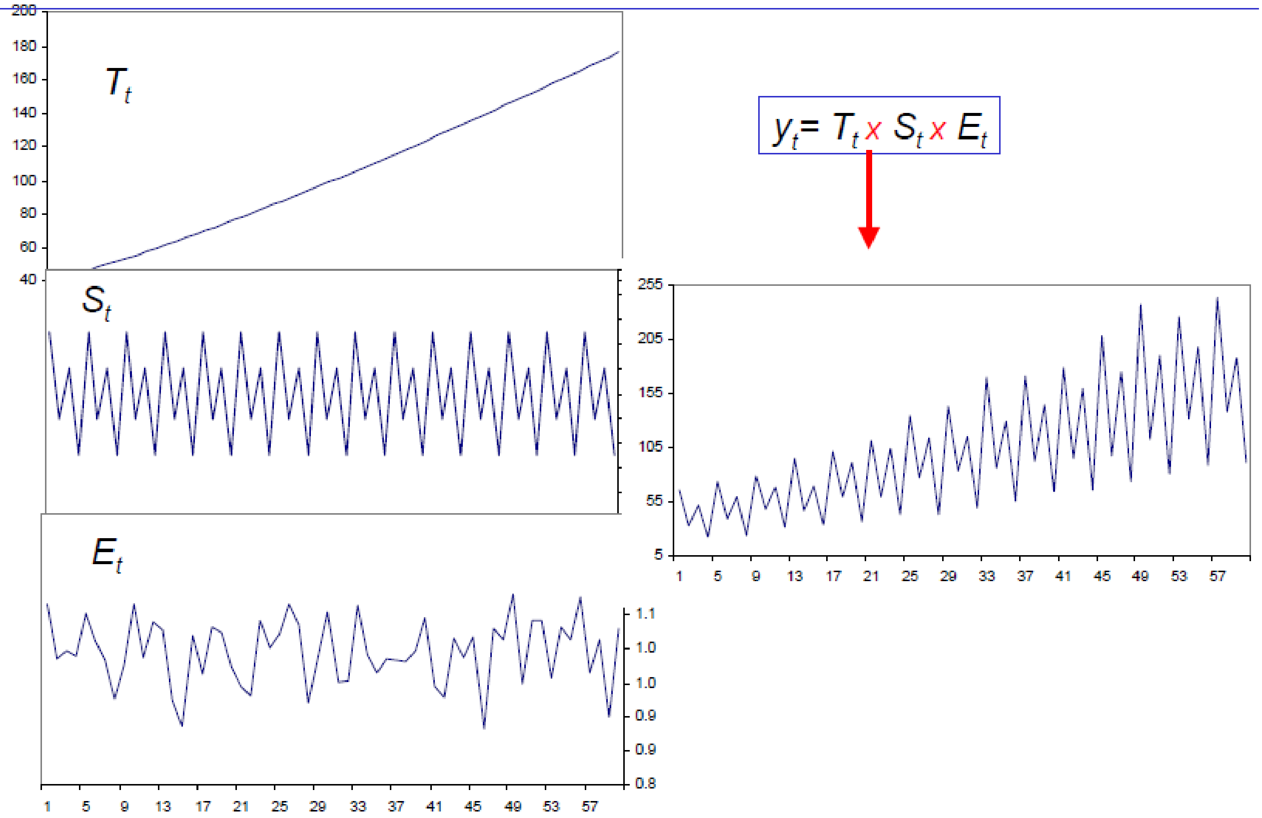
-assunzione modellistica troppo rigida

-visione settorizzata

Esempio di una serie (trimestrale) a componenti additive



Esempio di una serie (trimestrale) a componenti moltiplicative



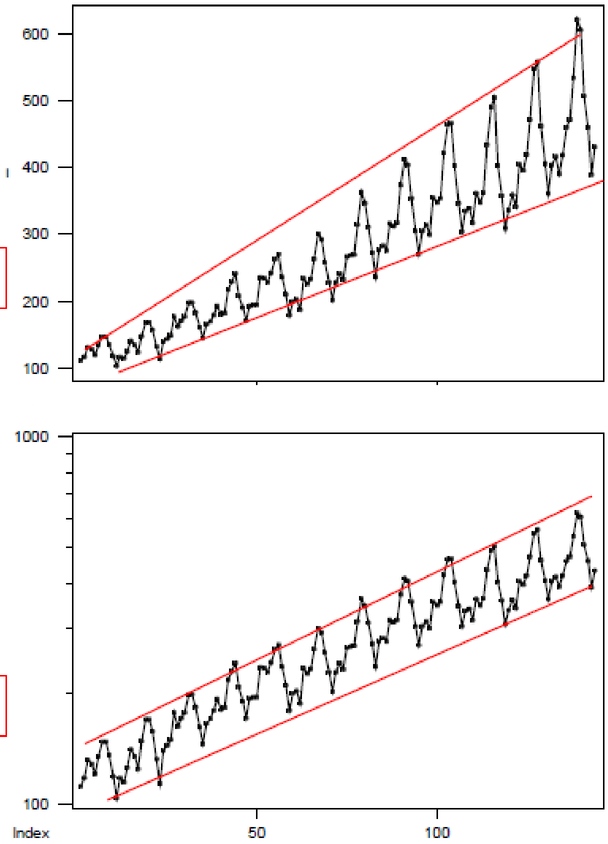
Che serie è?

Dati originari (mensili)
Passeggeri aerei 1949-60

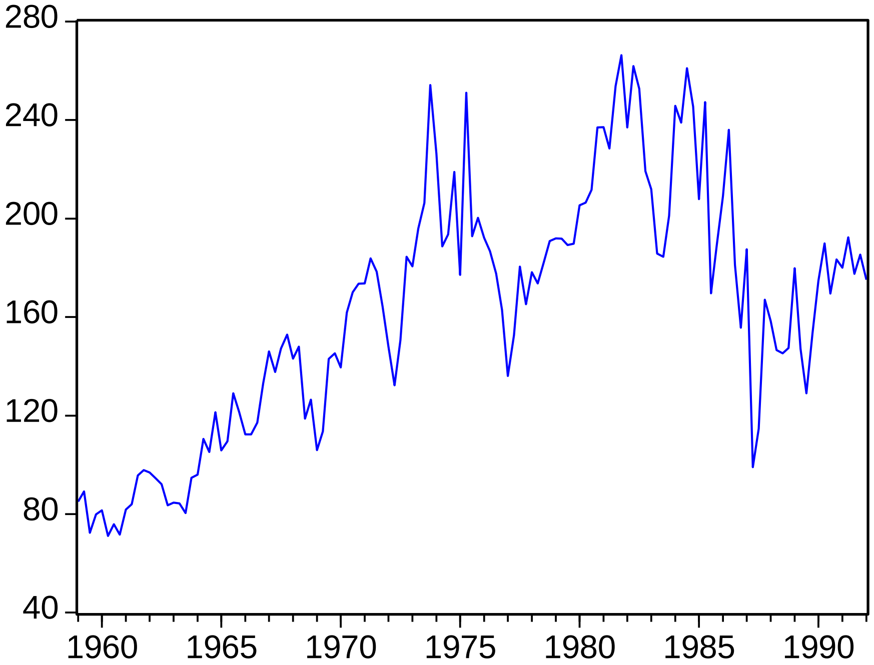
E' una serie moltiplicativa

Trasformazione logaritmica

Diventa una serie additiva



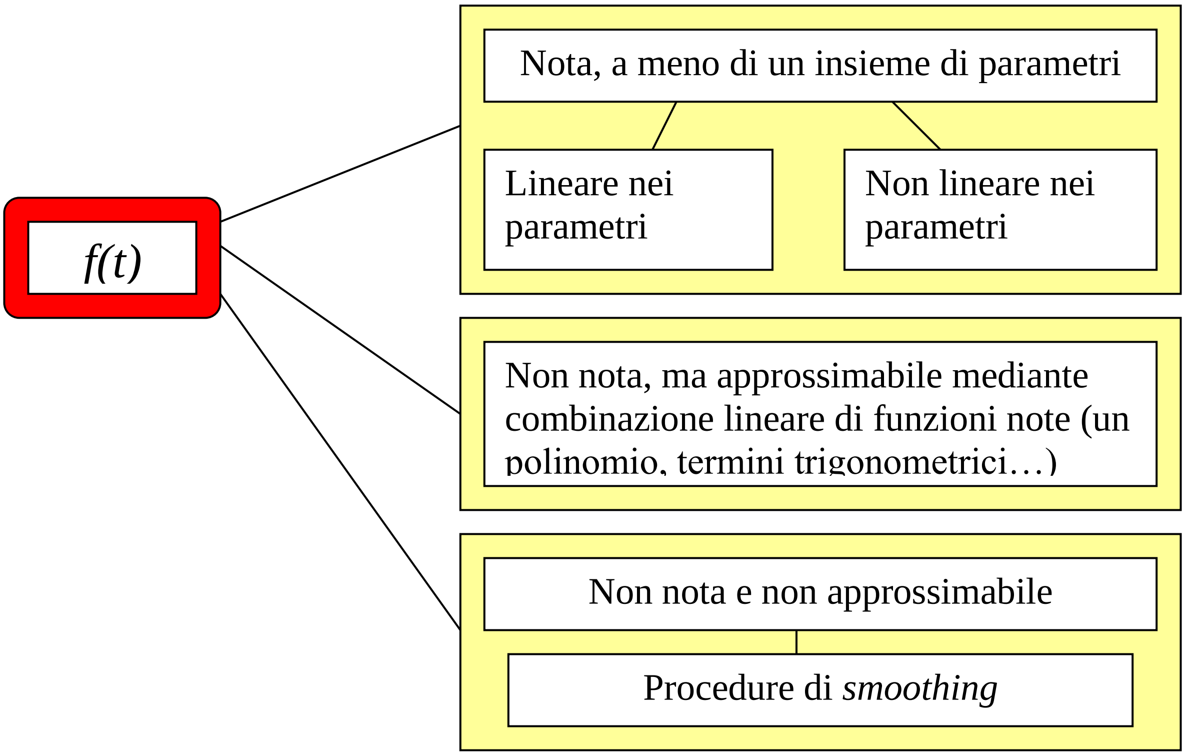
IL TREND



Il trend di una serie storica è la tendenza di fondo del fenomeno nel lungo periodo. Si caratterizza per un'evoluzione lenta e regolare nel tempo. In genere è rappresentabile mediante una qualche funzione del tempo da stimare.
Data una serie storica $\{y_t\}_{t=1}^{n}$ per la quale ipotizziamo un modello del tipo

$$Y_t = f(t) + \varepsilon_t \qquad \varepsilon_t \approx WN(0, \sigma_\varepsilon^2) \qquad (1.)$$

supponiamo per la nostra analisi che la parte sistematica *f(t)* della serie sia composta dal solo trend.
Come si rappresenta *f(t)*?


f(t)
Nota, a meno di un insieme di parametri
Lineare nei parametri
Non lineare nei parametri
Non nota, ma approssimabile mediante combinazione lineare di funzioni note (un polinomio, termini trigonometrici…)
Non nota e non approssimabile
Procedure di smoothing

1. Trend lineare o linearizzabile nei parametri

Tali trend sono stimabili attraverso le procedure derivate dal modello di regressione lineare.

- ***Trend polinomiale***

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

La (1.) diventa il modello di regressione lineare:

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

che può essere espressa in forma matriciale:

$$\mathbf{y} = \mathbf{P}\alpha + \varepsilon$$

con

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_q \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

e

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1^2 & \cdots & 1^q \\ 1 & 2 & 2^2 & \cdots & 2^q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n & n^2 & \cdots & n^q \end{bmatrix}$$

e stimata con il metodo dei minimi quadrati (OLS), dal quale si ottengono le stime del vettore α

$$\hat{\alpha} = (\mathbf{P}'\mathbf{P})^{-1} P' y$$

ATTENZIONE: il polinomio stimato può essere usato a fini interpolativi, ma deve essere usato con molta cautela per fini previsivi.

L'orine q del polinomio dipende dal comportamento di fondo della serie storica. Di solito si sceglie q abbastanza piccolo perché altrimenti si perdono gradi di libertà.
I più usati sono:

$$q = 0 \quad Y_t = \alpha_0 + \varepsilon_t \qquad \text{trend costante}$$

$$q = 1 \quad Y_t = \alpha_0 + \alpha_1 t + \varepsilon_t \qquad \text{trend lineare}$$

$$q = 2 \quad Y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \varepsilon_t \qquad \text{trend parabolico}$$

Come scegliere q?

<u>*Criterio delle differenze successive*</u>

Indichiamo B *l'operatore ritardo* che trasforma la serie y_t in

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

Per cui $\quad B^h\, y_t=y_{t-h}$

Definiamo $B^0=I$ *operatore identità (lascia immutata la serie)* $Iy_t=y_t$

Si chiama *operatore differenza prima all'indietro (I-B)* per il quale

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

- L'operazione differenza prima su di un polinomio ne riduce il grado

In generale se $f(t)$ è un polinomio di grado q, $(I-B)^q f(t)$ è costante.

<u>*Esempio*</u>

$$f(t) = \alpha_0 + \alpha_1 t$$

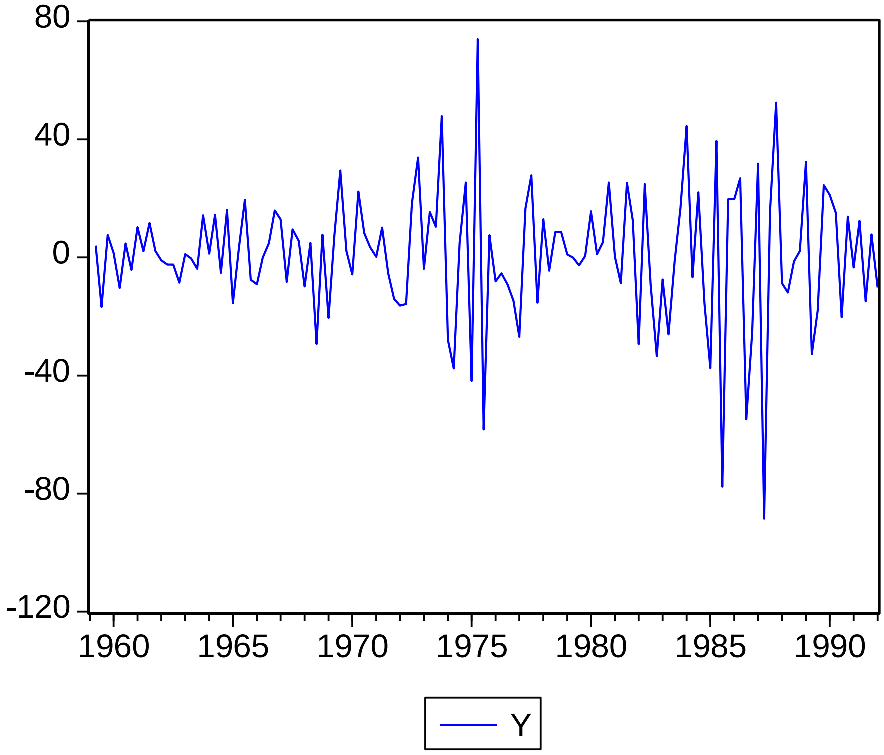
$$(I - B) f(t) = f(t) - f(t-1)$$

$$= \alpha_0 + \alpha_1 t - \alpha_0 - \alpha_1 (t-1) = \alpha_1$$

Esempio trend polinomiale di secondo grado (applicare le differenze, vedi anche libro p.25).

Tale risultato può essere utilizzato per individuare il trend polinomiale. Si calcolano le differenze successive della serie, arrestando l'operazione per un certo valore *r* per il quale la serie $(I\text{-}B)^r y_t$ sia approssimativamente costante.

ATTENZIONE: L'operatore *B* fa sentire i suoi effetti sulla componente accidentale ε_t, aumentando marcatamente la varianza della serie.



Esempio: tasso disoccupazione in Italia fig.2.1

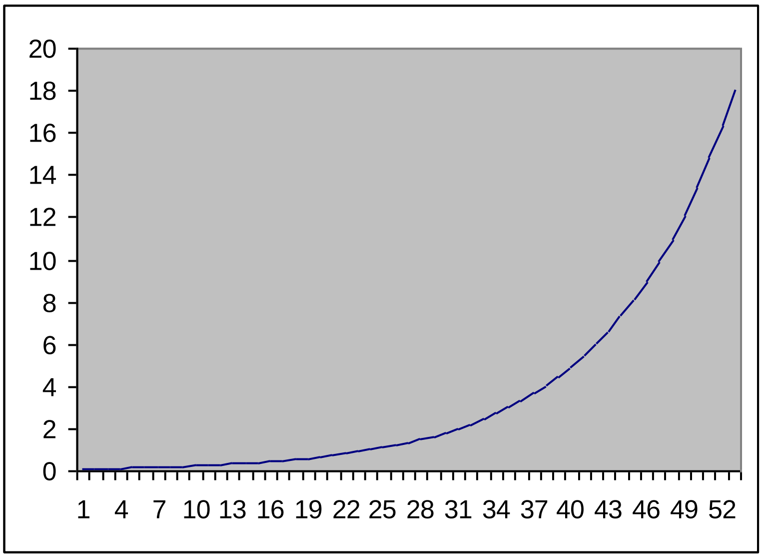
- *Trend esponenziale*
Adatto a fenomeni caratterizzati da una crescita nel tempo "esplosiva".

$$f(t) = \alpha_0 e^{\alpha_1 t} \quad (2.)$$

Con $\alpha_0>0$ e l'andamento dipende da α_1
La derivata della 2. rappresenta il *tasso di crescita* della curva al tempo t .

Il tasso di crescita relativo $\frac{df(t)}{dt} / f(t) = \alpha_1$ è costante.

Es. $f(t) = 0{,}1e^{0{,}1t}$



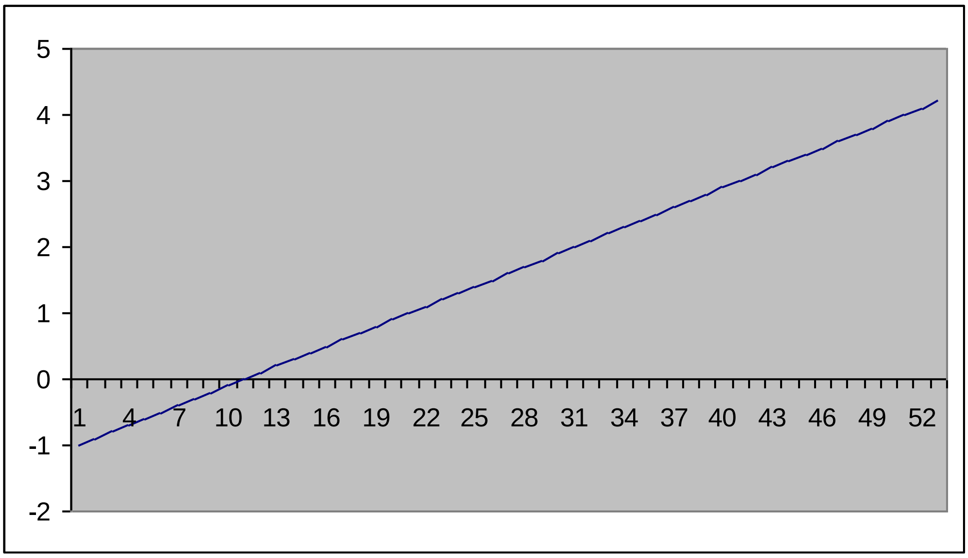
La trasformata logaritmica della 2. è la retta

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

Dunque

$$(I-B)\log f(t) = \alpha_1$$

Possiamo utilizzare queste caratteristiche per individuare una curva con trend esponenziale.



Sfruttando la trasformazione è possibile stimare i parametri α_0 e α_1
Per far questo bisogna però utilizzare un modello di tipo moltiplicativo:

$$y_t = f(t) \cdot \varepsilon_t = \alpha_0 e^{\alpha_1 t} \cdot \varepsilon_t$$

da cui

$$\log y_t = \log \alpha_0 + \alpha_1 t + \log \varepsilon_t$$
$$= \alpha_0^* + \alpha_1 t + \varepsilon_t^*$$

ATTENZIONE: la stima è distorta $E(y_t) \neq f(t)$, anche la previsione risulta distorta.

Se si utilizza la specificazione del modello additiva

$$y_t = \alpha_0 e^{\alpha_1 t} + \varepsilon_t$$

Non è lineare e non linearizzabile nei parametri. Utilizzando i minimi quadrati si ha un sistema non lineare nei parametri e la soluzione è raggiungibile solo per via iterativa.

Esempio: Numero di server Internet nel mondo fig.2.4

2. TREND NON LINEARE NEI PARAMETRI: CURVE DI CRESCITA

Sono trend non riconducibili a quelli polinomiali in t; sono legati a fenomeni caratterizzati da crescite molto accelerate. Le curve utilizzate non sono più lineari nei parametri.

- *Curva esponenziale modificata*

Tasso di crescita al tempo t direttamente proporzionale all'ammontare di crescita ancora da raggiungere.

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

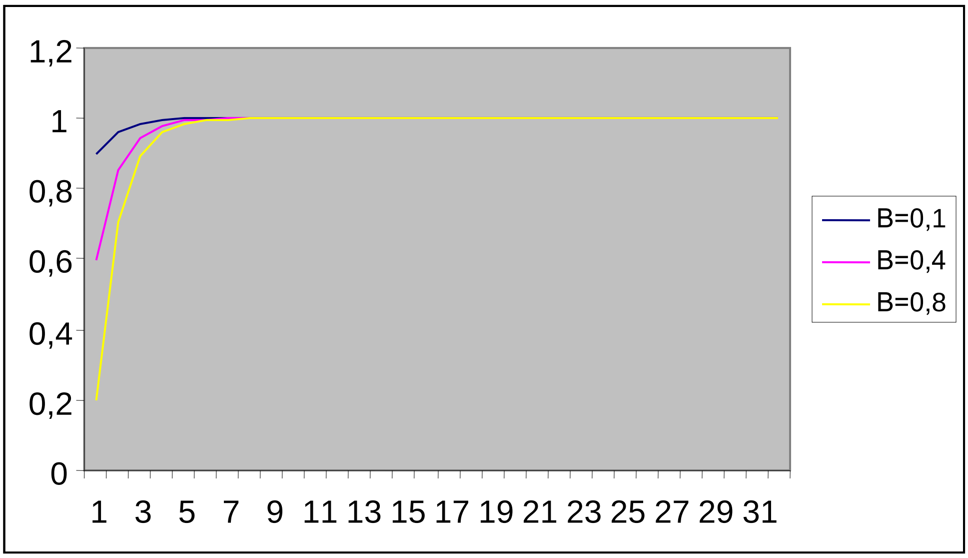
con β>0 determina l'intersezione della curva con l'asse verticale

α>0 denota il valore limite di crescita ed è anche fattore di scala della funzione

k costante di proporzionalità che controlla la scala lungo l'asse dei tempi

Esempio: curva esponenziale modificata per vari valori di β

$$f(t) = 1 - \beta e^{-t}$$



Usualmente la curva si trova nella forma

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

$$con\ \beta_0 = \alpha > 0 \quad \beta_1 = -\alpha\beta < 0 \quad e \quad \beta_2^* = -k < 0$$

Tale curva differisce da quella esponenziale semplice per la presenza della costante o *asintoto superiore* β_0.

Non è linearizzabile tramite trasformazione logaritmica, tuttavia la differenza $\beta_0 - f(t)$ è un'esponenziale linearizzabile e, conoscendo il valore dell'asintoto (spesso fissato dall'esterno), potremmo stimare i parametri come fatto in precedenza con la trasformazione logaritmica.

- La curva logistica

In questo caso tasso di crescita direttamente proporzionale al prodotto tra il livello raggiunto e l'ammontare di crescita ancora da raggiungere (con k>0 fattore di proporzionalità e α>0 valore limite della crescita).

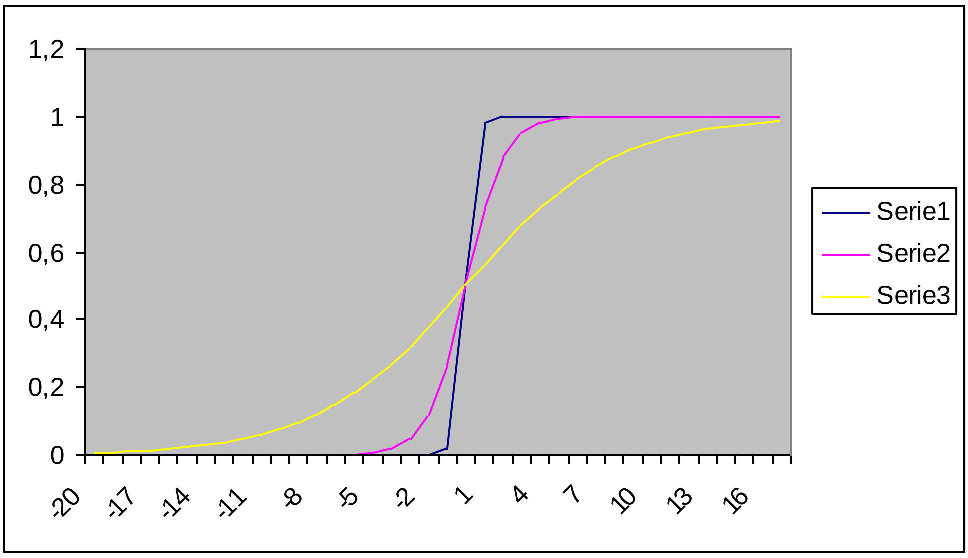
$$f(t) = \frac{\alpha}{1+\beta e^{-kt}} \quad \text{con } \beta>0$$

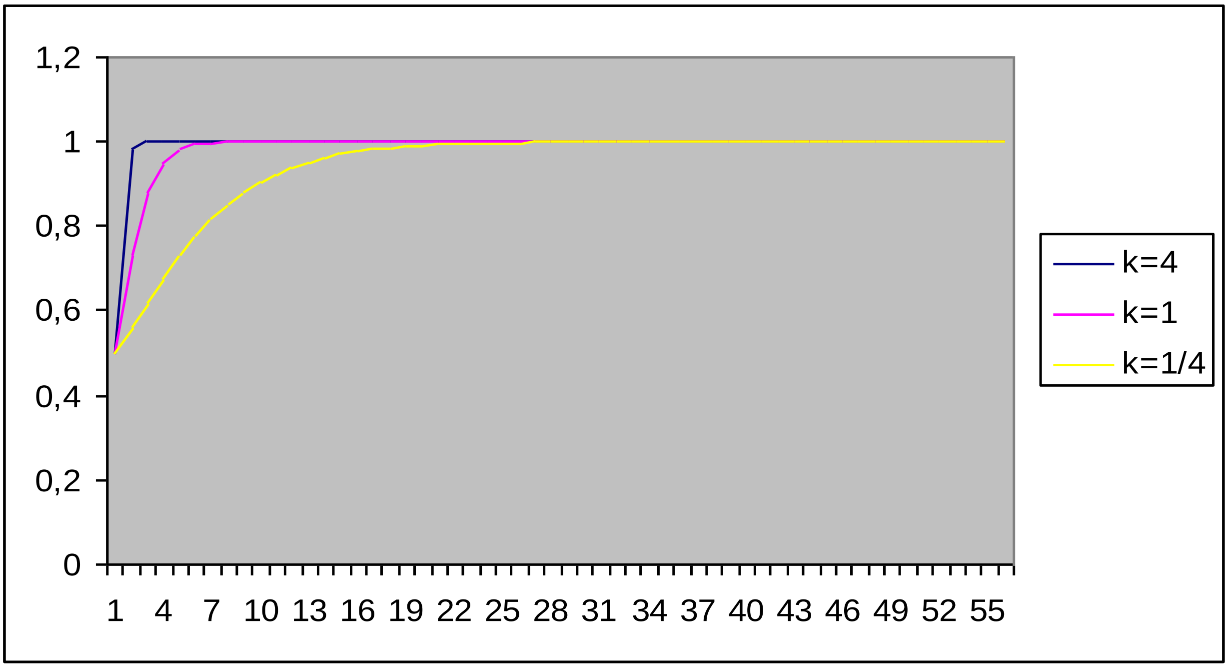
Da notare: α determina la scala della funzione;

k determina la scala lungo l'asse dei tempi cioè l'inclinazione della funzione;

β determina il punto d'incontro della curva con l'asse verticale.

Esempio: $f(t) = 1/(1+e^{-kt})$





Prendendo il reciproco 1/f(t) e ridefinendo le variabili si ottiene una esponenziale modificata.

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

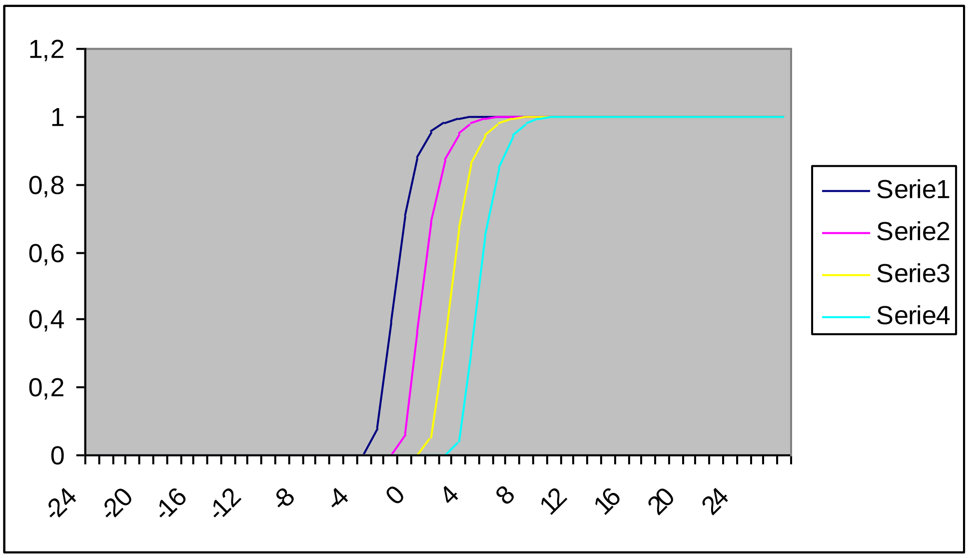
$$f^*(t) = \alpha^* + \beta^* e^{-kt}$$

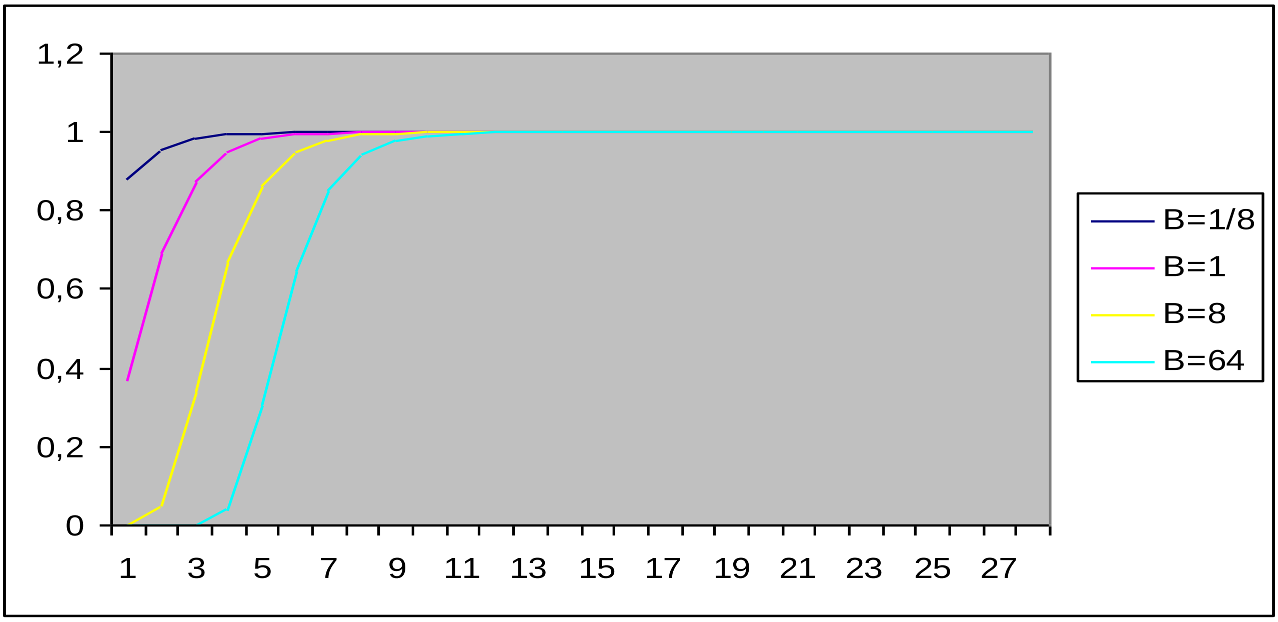
- ***La curva di Gompertz***

$$f(t)=\alpha\exp\left\{-\beta e^{-kt}\right\}\quad \text{con } \beta>0$$

La curva ha una forma ad S allungata simile alla logistica (senza la simmetria attorno al punto di flesso). α, β e K hanno la stessa interpretazione di quelli nella logistica e *logf(t)* è un'esponenziale modificata.

Esempio: α=1 K=1





LA COMPONENTE STAGIONALE

Anche questa componente, come il trend, può essere stimata con il modello di regressione. Si rappresenta mediante una funzione periodica *g(t)*.

Si dicono periodiche quelle funzioni il cui valore all'istante t si riproduce esattamente ad intervalli costanti, la cui lunghezza s costituisce il periodo

$g(t)=g(t+s)=g(t+2s)=g(t+3s)=\dots$

s=4 serie trimestrale   s=12 serie mensile

Supponiamo che il processo sia del tipo

$$Y_t=S_t+\varepsilon_t$$

Come si tratta tale componente nel modello di regressione?

<u>*L'uso di funzioni trigonometriche*</u>

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

La componente stagionale è data dalla somma di m armoniche con

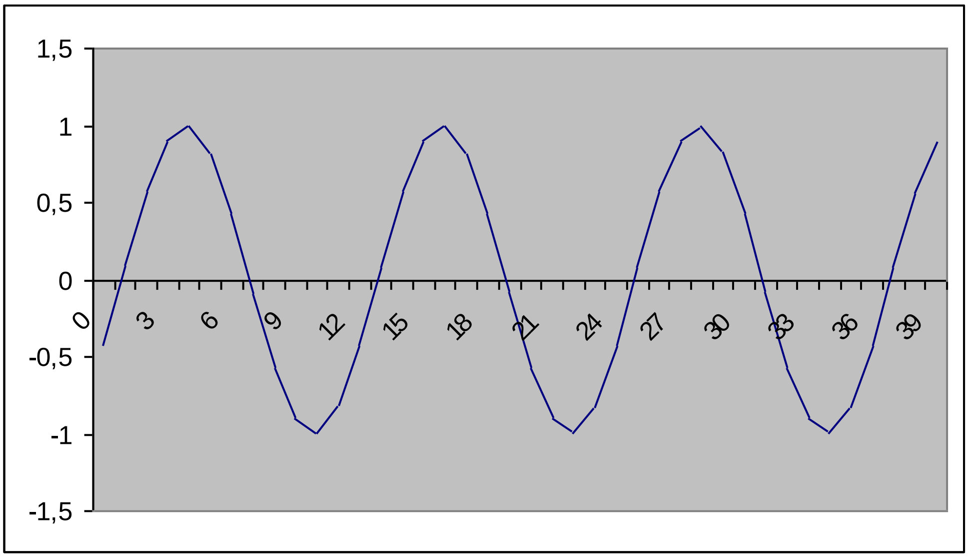
s/i periodo

$\omega_i = \dfrac{2\pi i}{s}$ frequenza angolare

A_i ampiezza

ϕ_i angolo di fase

Ad esempio: per dati mensili s=12, la prima armonica (i=1)



completa il ciclo in 12 periodi.

Combinando tale modello con un trend polinomiale di grado q si ha:

$$y_t = \alpha_0 + \alpha_1 t + \ldots + \alpha_q t^q + \sum_{i=1}^{m} A_i \cos(\omega_i t - \phi_i) + \varepsilon_t$$

Che sfruttando l'identità trigonometrica diventa

$$y_t = \alpha_0 + \alpha_1 t + \ldots + \alpha_q t^q + \sum_{i=1}^{m} \beta_{i1} \cos \omega_i t + \beta_{i2} sen \omega_i t + \varepsilon_t$$

con $\beta_{i1} = A_i \cdot \cos(\phi_i)$ e $\beta_{i2} = A_i \cdot \sin(\phi_i)$

che può essere comunque stimato con l'usuale metodo dei minimi quadrati.

ESERCIZI:

1. Trend polinomiale grado 1 e 2. Detrendizzare con il metodo delle differenze.
2. Data una serie di numeri, applicare le differenze.

MEDIE MOBILI

Quando siamo di fronte ad un fenomeno con un andamento molto irregolare che richiederebbe, dal punto di vista analitico, approssimazioni con polinomi di grado molto elevato, si può provare ad individuare la componente di fondo senza evidenziarne la legge sottostante.
Uno strumento utile a tal fine è la *media mobile*, utilizzata per stimare il trend, destagionalizzare ed eliminare o ridurre la componente erratica.

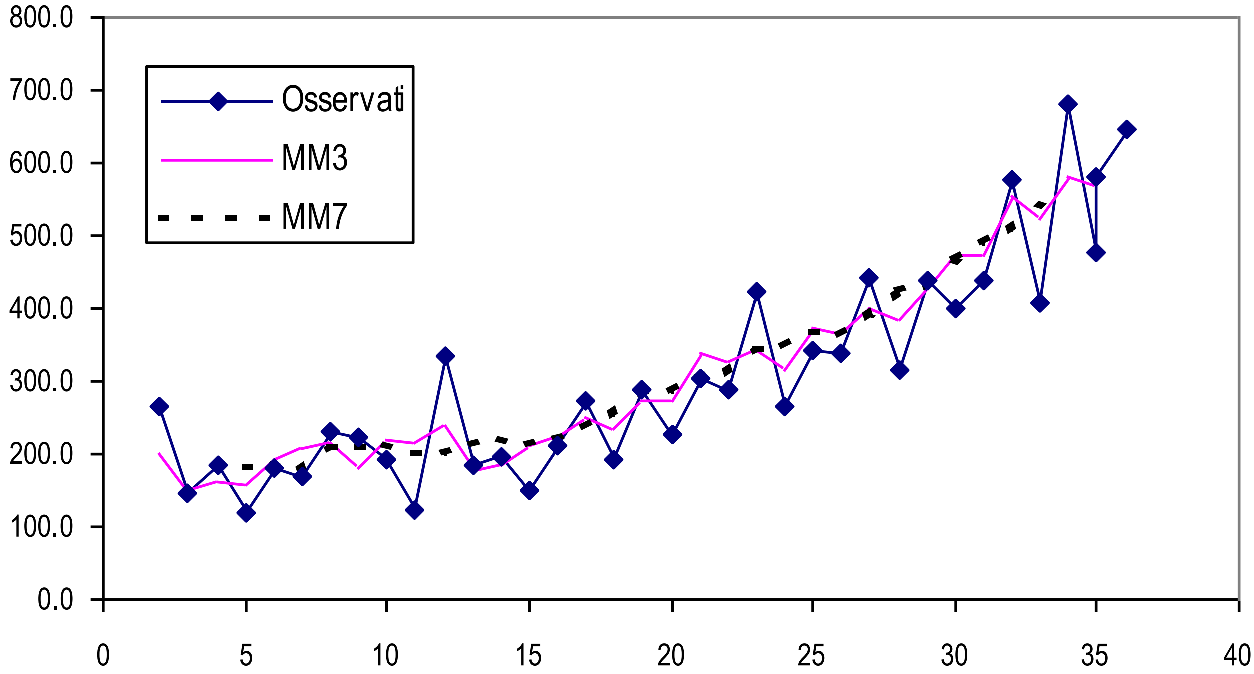
Consideriamo la seguente serie mensile

					Mese							
Anno	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	266.0	145.9	183.1	119.3	180.3	168.5	231.8	224.5	192.8	122.9	336.5	185.9
2	194.3	149.5	210.1	273.3	191.4	287.0	226.0	303.6	289.9	421.6	264.5	342.3
3	339.7	440.4	315.9	439.3	401.3	437.4	575.5	407.6	682.0	475.3	581.3	646.9

Media mobile a 3 e 7 termini

$$y_t\,(3)=(y_{t-1}+y_t+y_{t+1})/3, \qquad t=2,\ldots,n-1$$

$$y_t\,(7)=(y_{t-3}+y_{t-2}+y_{t-1}+y_t+y_{t+1}+y_{t+2}+y_{t+3})/7, \qquad t=4,\ldots,n-3$$



Data una serie storica per cui valga il modello additivo:

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

Dobbiamo trovare una trasformazione che, ad esempio, conservi il trend ed annulli le altre componenti. La media mobile consente tutto ciò.

La trasformazione della serie Y_t è una somma pesata dei valori della serie originale corrispondenti ad istanti temporali intorno a t.

$$y_t^* = \vartheta_{-m_1} y_{t-m_1} + \vartheta_{-m_1+1} y_{t-m_1+1} + \ldots + \vartheta_{m_2} y_{t+m_2}$$

cioè

$$y_t^* = \sum_{i=-m_1}^{m_2} \vartheta_i y_{t+i}$$

Il numero dei termini utilizzati (m_1+m_2+1) è detto *ordine* della media mobile.
Può essere scritta anche attraverso l'operatore ritardo B:

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

ottenendo una combinazione lineare finita con pesi ϑ_i di potenze successive di B.
Se i pesi sono tutti uguali, si ha la *media mobile semplice*

$$\vartheta_i = \frac{1}{m_1 + m_2 + 1}$$

$$\text{e se } m_1 = m_2 = m$$

$$y_t^* = \frac{1}{2m+1} \sum_{i=-m}^{m} y_{t+i}$$

tale media mobile è detta *centrata*.

ES. (3.1) pag. 85

T	1998	1999	2000	2001	2002	2003	2004
Y_t	6	7	14	3	13	5	15

Se m=1 $\quad y_{1999}^* = \frac{6+7+14}{3} = 9 \quad$ …..

- La media mobile gode della proprietà di <u>*composizione*</u>, cioè si possono applicare più medie mobili ad una serie, anche in sequenza e cambiando l'ordine della loro applicazione ottengo lo stesso risultato.
- Una media mobile è *simmetrica* se è centrata e se sono uguali i coefficienti aventi indice simmetrico rispetto a 0.

$$M = \left\{ [5]; \left[\frac{1}{5}, \frac{1}{5}, \mathbf{\frac{1}{5}} \right] \right\}.$$

- Una serie storica è *invariante* rispetto ad una media mobile M se $My_t=y_t$ per ogni t, cioè se la media mobile conserva y_t. Una media mobile conserva un polinomio di grado p se i suoi coefficienti sono tali che:

$$\sum_{i=-m}^{m} \vartheta_i = 1 \qquad e \qquad \sum_{i=-m}^{m} i^r \vartheta_i = 0 \quad per \quad r = 1,\dots,p$$

- Se si trasforma il termine di errore white noise mediante una media mobile,

$$\varepsilon_t^* = \sum_{i=-m}^{m} \vartheta_i \varepsilon_{t+i} \qquad E[\varepsilon_t^*] = \sum_{i=-m}^{m} \vartheta_i E[\varepsilon_{t+i}] = 0 \qquad \sigma^{*2} = Var[\varepsilon_t^*] = \sigma^2 \sum_{i=-m}^{m} \vartheta_i^2$$

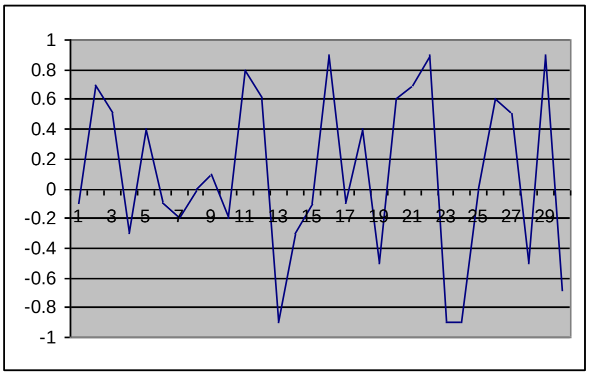
la serie trasformata, se $\sum_{i=-m}^{m} \vartheta_i^2 < 1$, ha varianza ridotta. La m.m. in questo caso ha un'azione spianante, cioè tende a ridurre le irregolarità di tipo casuale presenti nella serie. La quantità

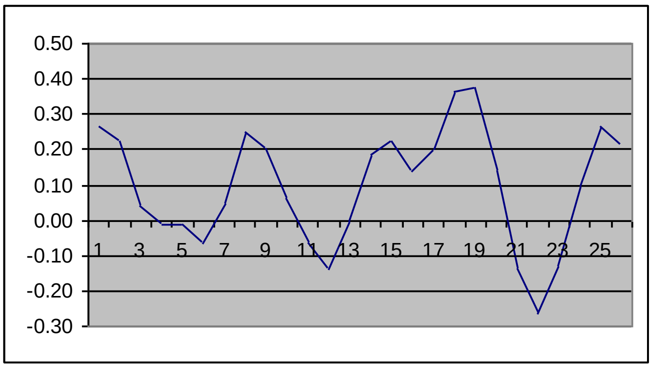
$$\frac{\sigma^{*2}}{\sigma^2} = \sum_{i=-m}^{m} \vartheta_i^2$$

è detta *rapporto di riduzione della varianza residua* e misura la capacità della m.m. di ridurre la perturbazione.

ATTENZIONE: le variabili così trasformate sono tra loro correlate. L'esistenza di correlazioni non nulle tra i valori successivi del modello ε_t^* introduce effetti di correlazione spuria (effetto di Slutsky-Yule), vedi esempio libro pag. 91.

ES. 3.2





Componendo più m.m. semplici si ottiene un'ampia famiglia di m.m. che forniscono buone approssimazioni e coefficienti di struttura semplice.
Ad es.
Componendo due medie mobili semplici di ordine 4 se ne ottiene una che è in grado di conservare trend lineari e di annullare stagionalità di periodo 4.
m.m. di Spencer: sono due m.m. di ordine 15 e 21.

M_{15} è la composizione di due m.m. semplici di ordine 4, di una m.m. di ordine 5 più una quarta m.m. M_4=[5]; ¼[-3, 3, **4**].
M_{21} si ottiene componendo 7 medie mobili.
M_{15} ed M_{21} annullano componenti stagionali di periodo 4 e 5 e conservano trend polinomiali di grado inferiore o pari a 3.

La stima di componenti di una serie storica utilizzando le medie mobili si basa sulle seguenti considerazioni:

- Esistono m.m. che preservano trend polinomiali fino ad un certo ordine: un trend polinomiale non è alterato dall'applicazione di queste m.m.;
- Se una m.m. di ordine *s* è applicata ad una serie contenente un'onda di periodo s, l'onda viene eliminata;
- Le m.m. con rapporto di riduzione della varianza minore di 1 riducono gli effetti della componente accidentale.

Stima della componente ciclica
Per ottenere una stima della componente ciclica si può utilizzare:

- Se i dati sono mensili m.m. centrate di 13 termini
- Se i dati sono trimestrali m.m. centrate di 5 termini

$$y_t^{**} = \frac{1}{8}(y_{t-2} + 2y_{t-1} + 2y_t + 2y_{t+1} + y_{t+2})$$

ATTENZIONE: si creano effetti di correlazione spuria (effetto di Slutsky-Yule) che introducono andamenti ciclici spuri.

Stima della componente stagionale
Lo studio della stagionalità presenta due problemi fondamentali:

a) la semplice stima di questa componente;
b) l'eliminazione, dopo la stima, di tale componente dall'andamento generale della serie storica (destagionalizzazione).

Per prima cosa si stimano le componenti trend-ciclo con m.m. centrate e si ricava la componente stagionale ed erratica, a seconda del modello di composizione: $y_t - y_t^{**}$ $\quad \frac{y_t}{y_t^{**}}$

Sviluppando il ragionamento per il modello moltiplicativo:

$$S\varepsilon_t \equiv IS_t = \frac{y_t}{y_t^{**}} \qquad t = m+1,\dots,n-m$$

Tali quantità sono dette *indici specifici* di stagionalità.
Prima di procedere a separare le due componenti stagionale ed erratica occorre verificare se quest'ultima è significativamente presente. Si ipotizza l'assenza di stagionalità e, se questa è vera, le medie degli IS per lo stesso mese calcolate in anni diversi non differiscono significativamente tra loro.
Se si rigetta tale ipotesi, otteniamo una serie di 12 valori detti *coefficienti grezzi* di stagionalità:

$$\hat{S}_j^* = \frac{1}{N}\sum_{T=1}^{N} IS_{T,j} \quad j = 1,...,12$$

Per il modello moltiplicativo deve valere che $\prod_{j=1}^{12}\hat{S}_j^* = 1$, ma questo non accade quasi mai, dunque si devono ricorreggere i coefficienti:

$$\hat{S}_j = \frac{\hat{S}_j^*}{\sqrt[12]{\prod_{j=1}^{12}\hat{S}_j^*}}$$

detti coefficienti *ideali* di stagionalità.
La serie destagionalizzata si ottiene dividendo y_t per $\hat{s}_j$.

- Le m.m. sono utilizzate anche per approssimare i polinomi locali. Tali polinomi, a differenza di quelli già introdotti nel trend, variano nei vari periodi temporali. I coefficienti di tali polinomi possono essere approssimati dalla quantità $\sum_{i=-m}^{m}\vartheta_i y_{t+1}$, che è una m.m. centrata di ordine 2m+1, simmetrica.
- Una m.m. che conserva i polinomi locali di grado p, conserva anche trend globali descritti da polinomi di grado p.
- Una m.m. che conserva i polinomi locali, conserva anche un trend globale costante poiché la somma dei coefficienti è sempre pari a 1.

Principi generali per la costruzione di una m.m.
Esistono tante m.m. che soddisfano criteri di invarianza di certi trend e di annullamento di componenti stagionali. *Come scegliere?*
Se si vuole tenere conto anche della componente accidentale, un criterio è minimizzare $\frac{\sigma^{*2}}{\sigma^2} = \sum_{i=-m}^{m}\vartheta_i^2$
Inoltre per verificare la bontà di una m.m. si analizza l'evoluzione dei coefficienti in funzione dell'indice.
Se considero la serie

$$y_t = \begin{cases} 0 & se \quad t \neq 0 \\ 1 & se \quad t = 0 \end{cases}$$

La sua trasformata secondo una media mobile di ordine 2m+1 è data da:

$$y_t^* = \begin{cases} 0 & se \quad t \leq -m-1 \\ \vartheta_{-t} & se \quad -m \leq t \leq m \\ 0 & se \quad t \geq m+1 \end{cases}$$

Tale evoluzione non deve presentare marcati punti di rottura.
Se si analizza la distribuzione dei coefficienti delle m.m. si osserva che molte di queste hanno dei marcati punti di rottura per i valori estremi (figura 3.10 libro).
Per verificare la flessibilità della curva dei coefficienti di utilizza la quantità

$$Q = \sum_{i=-m}^{m+3} ((1-B)^3 \vartheta_i)^2$$

Che misura lo scarto tra una parabola e la curva descritta dai coefficienti. Q è 0 quando tutti i coefficienti descrivono una parabola e cresce per scostamenti via via maggiori.

ESERCIZI:

T	1996	1997	1998	1999	2000	2001	2002	2003	2004
Y_t	8	4	6	7	14	3	13	5	15

1. Data la seguente media mobile $y_t^* = \frac{1}{4}(y_{t-2} + y_{t-1} + y_t + y_{t+1})$, descrivere M e applicare M ai dati.
2. Applicare una media mobile centrata di ordine 3.
3. Applicare un'opportuna media mobile per non perdere i dati più recenti della serie.
4. Applicare la media mobile $M = \left\{[3]; \left[\frac{1}{4}; \frac{\mathbf{2}}{\mathbf{4}}\right]\right\}$ e descriverla.

IL LISCIAMENTO ESPONENZIALE (EXPONENTIAL SMOOTHING)

Uno degli scopi dell'analisi delle serie storiche è fornire indicazioni previsive sul fenomeno d'interesse. Per previsione si intende un descrizione di avvenimenti futuri che si fonda su un insieme coordinato di ipotesi. La previsione si fonda sul presupposto logico che, in talune situazioni, è ragionevole supporre che le informazioni disponibili per il passato, adeguatamente trattate, siano in grado di ridurre l'incertezza circa gli accadimenti futuri.

La previsione si distingue in relazione a tre elementi:

a) orizzonte temporale:
   - breve o *congiunturale* (12-18 mesi)
   - medio (5 anni)
   - lungo termine (non altre 10 anni);

b) obiettivo
   - strumentale (unico scopo avvertimento)
   - tendenziale (evitare il ripetersi di fenomeni)
   - condizionale (ipotesi forti che condizionano le previsione)
   - normativa (come è possibile conseguire certi obiettivi);

c) metodo impiegato
   - informale o previsione *naif* (soggettivo, saggezza accumulata dal decisore)
   - serie storiche (*lisciamento esponenziale*)
   - regressioni e modelli econometrici.

Il lisciamento esponenziale è il metodo più utilizzato specialmente in ambito aziendale per le previsioni nel breve periodo. Qualità: flessibilità e semplicità d'uso.

Data una serie $\{y_t\}_{t=1}^{n}$ supponiamo di volere prevedere y_{n+k} con k>1 è detto *orizzonte temporale*. Si indica con $\hat{y}_{t+k}$ la previsione fatta al tempo t.

Se la serie ha un trend costante alterato da fattori accidentali $y_t = a + \varepsilon_t$,

la previsione è data dalla perequazione di tutte le informazioni disponibili fino a n con uguale peso:

$$\hat{y}_{n+1} = \frac{1}{n}\sum_{t=1}^{n} y_t = \frac{1}{n}\sum_{j=1}^{n} y_{n-j+1}$$

Se invece la serie cambia in modo stocastico è più realistico calcolare la previsione tramite una media ponderata dando peso maggiore alle osservazioni più recenti:

$$\hat{y}_{n+1} = \frac{\sum_{l=1}^{n} w_l y_{n-l+1}}{\sum_{l=1}^{n} w_l} \quad \text{con } w_l \geq 0;\ w_l \leq w_{l-1} \ \text{l=2,3,...,n}$$

scritta anche

$$\hat{y}_{n+1} = \sum_{j=0}^{n-1} c_j y_{n-j} \qquad c_j = \frac{w_{j+1}}{\sum_{l=1}^{n} w_l} \qquad j = 0,1,\ldots,n-1$$

Il lisciamento esponenziale è detto tale perché la serie y_t viene sostituita dalla successione $\hat{y}_{t+1}$ ottenuta da:

$$\hat{y}_{t+1} = c_0 y_t + c_1 y_{t-1} + c_2 y_{t-2} + \ldots$$

Con pesi definiti dalla successione esponenziale

$$c_j = c_0 \delta^j, \quad j = 0,1,\ldots,n-1, \quad 0 < \delta < 1$$

La previsione con costante di lisciamento δ è pari a:

$$\hat{y}_{n+1} = (1-\delta) \sum_{j=0}^{n-1} \delta^j y_{n-j}$$

Dove la previsione è una media ponderata di tutte le osservazioni disponibili; queste influenzano la previsione con intensità decrescente all'aumentare della distanza dal tempo n. Più la costante è vicina ad 1 più la previsione è *rigida* cioè più influenzata dalle osservazioni passate, viceversa più è prossima a 0 più la previsione è *flessibile* cioè influenzata dalle osservazioni più recenti.
La precedente si può anche scrivere:

$$\hat{y}_{n+1} = \delta \hat{y}_n + (1-\delta) y_n \qquad (1.)$$

Cioè come una media ponderata tra la previsione fatta al tempo n-1 e l'ultima osservazione y_n , il cui peso è tanto più forte quanto più piccola è la costante.
Appare chiara la logica di aggiornamento sequenziale del metodo (la previsione viene modificata dall'osservazione più recente) ed il ruolo della costante di lisciamento

Riassumendo:

- per $\delta \to 0$ il lisciamento esponenziale attribuisce sempre più peso ai nuovi dati e l'effetto perequativo è quasi nullo, considerando dunque la serie affidabile e la previsione non fa altro che restituire l'ultima osservazione disponibile.
- per $\delta \to 1$ il lisciamento esponenziale attribuisce peso pressoché nullo ai nuovi dati e la nuova previsione tende a coincidere con la precedente.

La costante esprime la *vischiosità* del sistema (importanza vecchia previsione).

Per stimare la funzione (1.) è necessario avere un valore iniziale di δ e di $\hat{y}_1$.

δ viene posto tra 0,05 e 0,3 oppure stimato mediante algoritmi non lineari. Ad esempio con i minimi quadrati si cerca quel valore che minimizza:

$$\sum_{t=m}^{n-1} (y_{t+1} - \hat{y}_{t+1})^2$$

Per quanto riguarda $\hat{y}_1$ usualmente:

$\hat{y}_1 = y_1$ se δ è molto piccolo

$\hat{y}_1 = (y_1 + y_2)/2$

$\hat{y}_1 = \frac{1}{n}\sum_{t=1}^{n} y_t$ se δ è più vicino ad 1

Se n è grande e/o la costante di lisciamento è piccola, la scelta di tale valore è ininfluente.

La previsione può essere vista come la costante che meglio si adatta alla serie in prossimità di n. Questo fa capire che se la serie ha un trend non costante o fluttuazioni marcate, tale metodo non è appropriato.

Es. fig. 4.1

<u>I metodi di Holt-Winters</u>

Generalizzazione del precedente; cercano la perequazione di y_t tramite la determinazione del livello, del trend e della stagionalità.

- *Metodo non stagionale*

Hp. che in prossimità di n un aggiustamento mediante una retta di equazione $L_n + T_n(t-n)$ sia preferibile alla perequazione fornita da una costante.

Le stime delle formule di aggiornamento sono:

$$\hat{L}_n = \alpha(\hat{L}_{n-1} + \hat{T}_{n-1}) + (1-\alpha)y_n \quad 0<\alpha<1$$

$$\hat{T}_n = \beta\hat{T}_{n-1} + (1-\beta)(\hat{L}_n - \hat{L}_{n-1}) \quad 0<\beta<1$$

Tale metodo è più flessibile del precedente in quanto fa uso di due costanti, anche se sorge qualche problema in più nella scelta dei valori inziali.

- *Metodo stagionale additivo*

Nelle vicinanze di n la serie storica è esprimibile come

$$L_n + T_n(t-n) + S_n$$

Dove S è un fattore stagionale di periodo s.

- *Metodo stagionale moltiplicativo*

Nelle vicinanze di n la serie storica è esprimibile come

$$L_n + T_n(t-n)S_n$$

Valutazione qualità previsioni

a) *date due o più serie di previsioni, quale ha fornito risultati migliori?*
b) *In che misura una serie di previsioni può considerarsi "soddisfacente"?*

Previsione punti di svolta

In genere si misura la distanza tra la previsione e successiva realizzazione. Si pone particolare importanza sui punti di svolta (previsti e realizzatisi).

Indicati con p_t valore previsto al tempo t e r_t valore realizzato, si costruiscono due vettore di tali valori per tutti gli anni considerati. Con tali valori è possibile costruire la seguente tabella.

Si ha:

Valori previsti p_t	Valori realizzati r_t		
	p.s.	No p.s.	Tot
p.s.	n_{11}	n_{12}	$n_{1.}$
No p.s.	n_{21}	n_{22}	$n_{2.}$
Tot	$n_{.1}$	$n_{.2}$	n

Sulla diagonale principale si hanno le previsioni che si sono rivelate esatte. n_{12} rappresenta la frequenza degli errori di prima specie (ps previsto e non realizzato), n_{21} la frequenza degli errori di seconda specie (ps non previsto e realizzato).
Si definisce indice relativo degli errori di prima specie:

$$E_1 = \frac{n_{12}}{n_{1.}}$$

E indice relativo degli errori di seconda specie:

$$E_2 = \frac{n_{21}}{n_{.1}}$$

Errore medio di previsione
Indichiamo con $e_t=p_t-r_t \quad t=1,\ldots,n$ l'errore di previsione.
Media potenziata di ordine s degli errori di previsione:

$$I_s = \left(\frac{1}{n} \sum_{t=1}^{n} |e_t|^s \right)^{\frac{1}{s}}$$

Per s=1 si ha l'errore assoluto medio di previsione:

$$EAM = \frac{1}{n} \sum_{t=1}^{n} |e_t|$$

Spesso affiancato dall'errore medio di previsione:

$$EM = \frac{1}{n} \sum_{t=1}^{n} e_t$$

Se i due indici sono molto vicini, c'è sistematicità negli errori.
Un altro indice è la media quadratica degli errori di previsione, ottenuta ponendo s=2:

$$MQE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} e^2{}_t}$$

Il quadrato di tale quantità è scomponibile in tre addendi che rappresentano quanta parte dell'errore è dovuta a:

- diversa media dei valori previsti e realizzati
- diversa variabilità dei valori previsti e realizzati
- imperfetta correlazione lineare dei valori previsti e realizzati

Nello specifico:

$$ES = \frac{(\bar{p} - \bar{r})^2}{MQE^2} \text{ errore sistematico}$$

$$EV = \frac{(\sigma_p - \sigma_r)^2}{MQE^2} \text{ errore nelle variabilità}$$

$$EC = \frac{2\sigma_p \sigma_r (1 - \rho_{pr})}{MQE^2} \text{ errore nelle covarianza}$$

Con $ES+EV+EC=1$

Per confrontare previsioni relative a più serie storiche si utilizza l'indice di Theil:

$$U = \frac{MQE}{\sqrt{\frac{1}{n}\sum_{t=1}^{n} r_t^2}}$$

Ovvero:

$$U = \sqrt{\frac{\sum_{t=1}^{n}(p_t - r_t)^2}{\sum_{t=1}^{n} r_t^2}}$$

Infine per analizzare l'attendibilità delle previsioni in due archi di tempo successivi si utilizza il coefficiente di Giano.

Supponiamo che l'intero periodo di n tempi sia suddiviso in due sottoperiodi di lunghezza pari a rispettivamente n_1 e n_2 con $n_1 + n_2 = 1$.

$$J = \frac{\frac{1}{n_2}\sum_{t=n_1+1}^{n} e_t^2}{\frac{1}{n_2}\sum_{t=1}^{n_1} e_t^2}$$

Assume valore nullo quando a previsioni non tutte esatte nel primo sottoperiodo corrispondono previsioni perfette nel secondo; viceversa tende ad infinito.

Free econometric software:
http://gretl.sourceforge.net/