Università degli studi di Padova

Facoltà di Ingegneria

Dipartimento di tecnica e gestione dei sistemi industriali

Tesi di Laurea di Primo Livello

Modelli Estrapolativi di Previsione

RELATORE: Prof. Romanin Jacur Giorgio

LAUREANDO:Rizzetto Alessandro

ANNO ACCADEMICO:2010/2011

INDICE

SOMMARIO

In questo elaborato illustrerò i modelli estrapolativi di previsione,approfondendo i concetti relativi a questa metodologia di previsione.

Andremo a trattare le serie storiche,ad esaminare le loro caratteristiche e in particolare le 4 componenti principali.

Considereremo poi tutti i metodi per la valutazione dei modelli,importanti sia per la scelta del modello sia nella fase di controllo e monitoraggio delle previsioni. Definirò quindi l'errore di previsione,le misure di distorsione e di dispersione e il segnale di tracking.

Tratterrò poi il concetto di media mobile e il suo utilizzo nell'effettuazione di predizioni e nella scomposizione delle serie storiche.

Esporrò i vari modelli di smoothing esponenziale,il modello di Brown, di Holt,il modello di Winters e quello a tendenza ridotta

Infine approfondiremo la tecnica di previsione dei modelli autoregressivi,illustrando le formule per ottenere una previsione utilizzando questa tipologia di modelli

INTRODUZIONE

Morse(1967,pag 9)affermò che la ricerca operativa procede analizzando gli aspetti quantitativi delle attività umane(operazioni),elaborando modelli matematici che rappresentano alcune delle interrelazioni fra queste attività,e utilizzando questi modelli per prevedere la reazione delle operazioni alle eventuali variazioni dei fattori esterni o interni. Queste previsioni sono quindi messe a disposizione del responsabile dell'operazione,per aiutarlo nella scelta delle linee d'azione e dei piani da adottare.

Le aziende operano in un contesto sempre più competitivo e variabile ed alla luce della citazione precedente si può dedurre l'importanza della ricerca operativa e in particolare delle previsioni all'interno di un'azienda : è quindi sempre più fondamentale poter fornire delle previsioni accurate per il periodo futuro.

Esistono tre principali tecniche di previsione:i metodo di inferenza statistica,i metodi qualitativi e i metodi quantitativi.

I metodi di inferenza statistica vengono utilizzati per stimare medie e percentuali e di solito vengono utilizzati per indagini di mercato.

I metodi qualitativi si basano invece su valutazioni soggettive di esperti e sono quindi influenzati dalle opinioni personali di chi effettua le previsioni.

I metodi quantitativi sono suddivisi nei modelli esplicativi e nei modelli estrapolati. I primi tentano di fornire un legame tra la grandezza oggetto di studio e un insieme di *n* variabili che possono caratterizzarla e tale legame può essere espresso attraverso una relazione funzionale,in simboli:

$$y = f(x_1, x_2, \dots, x_n)$$

I modelli estrapolativi utilizzano invece le osservazioni di una serie storica per trarne le eventuali regolarità e per dedurne il comportamento futuro. Si potrà quindi riconoscere un'probabile tendenza o una ciclicità delle serie storiche in esame per accertarne lo sviluppo futuro.

La scelta della tecnica più adatta dipende dall'obiettivo e dai parametri delle decisioni,come per esempio le caratteristiche del prodotto in esame o la disponibilità di un

elevato numero di osservazioni. Ad incidere è anche il basso rapporto costo-benefici, poiché un tipo di modello previsionale molto complesso è generalmente poco affidabile rispetto alla variabilità delle serie storiche di natura economica.
Dopo aver scelto il metodo previsionale da adottare risulta allo stesso modo molto importante la determinazione dei parametri del modello,che generalmente si effettua analizzando le osservazioni disponibili e richiede la minimizzazione dello scarto quadratico.
Infine l'ultima fase del processo previsionale prevede il controllo delle previsioni svolte dal modello. Sarà sufficiente compiere un paragone fra le previsioni e le osservazioni reali che verranno raccolte,in modo tale da capire se il modello è efficace o meno. Nel caso in cui la differenza sia considerevole si dovrà modificare i parametri del modello o scegliere una diversa forma funzionale del modello.

CAPITOLO 1

Modelli Estrapolativi e Serie Storiche

1.Introduzione e definizioni

Oltre ai metodi qualitativi e di statistica inferenziale per effettuare previsioni sull'andamento di una grandezza di interesse spesso si utilizzano metodi quantitativi,suddivisi in modelli esplicativi e in modelli estrapolativi.

I modelli estrapolativi consentono di ricavare le eventuali regolarità e identificare l'andamento nel futuro di una grandezza in oggetto di studio,partendo da una serie storica di osservazioni relative alla grandezza stessa.

Andiamo allora ad enunciare le principali definizioni.

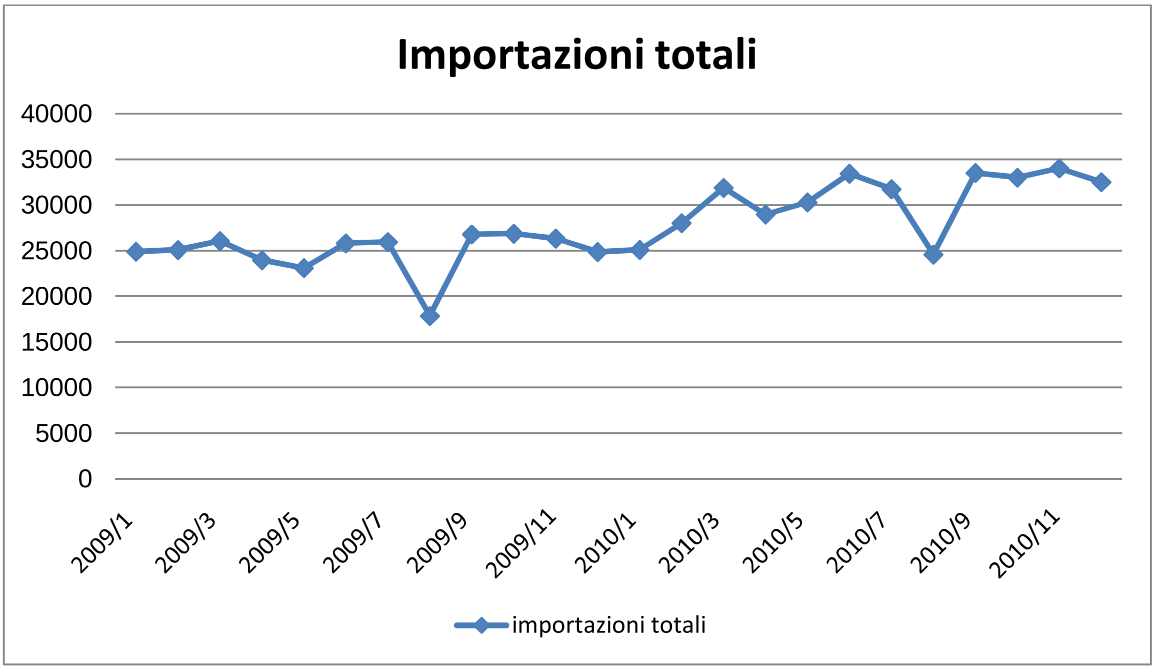
1.1.Definizione di serie storica

Per serie storica intendiamo una sequenza di valori A_t assunti da una grandezza misurabile in coincidenza di determinati istanti temporali *t*,di norma posti uniformemente,come ad esempio giorni,settimane o trimestri. Le serie storiche ci permettono,quindi,di analizzare l'evoluzione della grandezza nel tempo.Le serie storiche possono essere classificate in diverse tipologie,ma sicuramente quelle di maggior interesse sono le serie storiche a carattere economico. Sul sito internet dell'istituto Istat (www.con.istat.it) è possibile ricavare numerose e varie serie storiche,andiamo quindi a proporre un esempio,relativo alle importazioni totali su base mensile per gli anni 2009 e 2010. La seguente tabella rappresenta la serie storica:

t	importazioni totali
2009/1	24890,59907
2009/2	25088,55162
2009/3	26074,27644
2009/4	23968,3531
2009/5	23100,81515
2009/6	25815,17748
2009/7	25951,44068

2009/8	17850,19228
2009/9	26794,37182
2009/10	26879,6974
2009/11	26335,33123
2009/12	24859,85682
2010/1	25093,80233
2010/2	28021,3764
2010/3	31889,58024
2010/4	28976,88092
2010/5	30293,29839
2010/6	33444,09968
2010/7	31745,58005
2010/8	24587,47049
2010/9	33519,21307
2010/10	33015,57142
2010/11	34023,36038
2010/12	32511,32799

Con il temine *t* intendiamo istanti di tempo discreti per cui la rappresentazione grafica più immediata è quella nel piano cartesiano. Per la serie delle esportazioni si ha quindi:



1.2.Variabile serie storica

Una variabile serie storica è una variabile casuale che corrisponde alle osservazioni di una serie storica. Ogni osservazione rappresenta la realizzazione di una variabile casuale. Un modello della serie storica consiste nella definizione della distribuzione di probabilità della sequenza di valori di variabili casuali

1.3.Forma generale di un modello estrapolativo

Consideriamo il valore A_{t+1} di una serie storica corrispondente al periodo *t+1*,ipotizzando di trovarsi nel periodo t e di disporre di un numero pari a k di osservazioni passate della serie storica,è possibile definire la forma generale di un modello estrapolativo:

$$F_{t+1} = f(A_t, A_{t-1}, \dots, A_{t-k+1})$$

Con F_{t+1} viene indicato una predizione del valore della serie storica A_{t+1} del periodo *t+1.*
Nella determinazione di un modello estrapolativo è molto importante la scelta della forma funzionale *f* idonea a rappresentare la serie storica in esame.
Le previsioni si possono riferire anche a periodi successivi a *t+1*,basandosi sull'utilizzo del modello ai valori noti fino al tempo t e a predizione riferite ai periodi successivi sulla base del modello stesso. Indicando con *m* l'orizzonte temporale si ha:

$$F_{t+m} = f(F_{t+m-1}, F_{t+m-2}, .., F_{t+1}, A_1, A_{t-1}, .., A_{t-k+1})$$

Ovviamente le previsioni divengono sempre meno affidabili man mano che spostiamo sempre più in là nel futuro l'orizzonte temporale *m.*

2.Componenti di una serie storica

Data una serie storica A_t si possono considerare 4 componenti della stessa: tendenza,ciclicità,stagionalità e fluttazione casuale.

La *tendenza* T_t fornisce a lungo termine l'andamento nel tempo della serie storica e può essere crescente, stabile o decrescente. La tendenza può assumere un profilo lineare, esponenziale o logaritmico.
La ciclicità C_t si riferisce alle oscillazioni ondulatorie di una serie storica che si dimostrano

con periodicità variabile,per effetto dei cicli economici. Nelle previsioni di breve termine la ciclicità è scambiata con la tendenza,perché generalmente la si evidenzia per periodi di più lunga durata.
La *stagionalità* Q_t deriva dalle oscillazioni ondulatorie di periodicità regolare e di breve periodo,che si evidenziano per esempio durante i mesi dell'anno. Le oscillazioni sono persistenti e si spiegano con la stagionalità del prodotto o con il ciclo naturale di consumo del prodotto a cui si riferisce la serie .
La *fluttuazione casuale* ε_t è invece la componente che identifica tutte le variazioni che non vengono spiegate dalle altre componenti.

Lo studio delle componenti presuppone che la serie storica si possa identificare nella forma:
valori osservati = andamento ricorrente + fluttuazioni casuali

Alcuni modelli estrapolativi vengono interpretati sulla relazione funzionale:

$$A_t = f(T_t, C_t, Q_t, \varepsilon_t)$$

dalla quale si intuisce la dipendenza della serie storica dalle sue componenti .

3.Numeri indice

Per rappresentare una serie storica spesso è conveniente l'uso dei numeri indice.Hanno applicazione soprattutto in campo finanziario e tra i più conosciuti vi è l'indice di borsa FTSE Mib (www.borsaitaliana.it).
Un numero indice semplice è il rapporto,moltiplicato per 100,tra il valore di un' unica osservazione At al tempo t e il corrispondente valore A_1 al tempo t_1. In simboli:

$$I_t = \frac{A_t}{A_1} * (100)$$

Nel caso si voglia,come dice Zani (1991, pag 214) , misurare sinteticamente,per ciascuno dei tempi considerati, le variazioni relative di più fenomeni quantitativi i cui dati sono raccolti in una serie storica multipla,è possibile ricorrere ai numeri indice composti. Ipotizzando di esaminare z serie storiche $A_t^1, A_t^2, \dots, A_t^z$ il numero indice composto è il rapporto tra le somme A_t e A_1 ,moltiplicato per 100.

$$I_t = \frac{A_t}{A_1} * (100)$$

Con:

$A_t = A_t^1 + A_t^2 + \cdots + A_t^z$

$A_1 = A_1^1 + A_1^2 + \cdots + A_1^z$

Si può considerare anche il numero indice composto pesato in cui vengono assegnati dei pesi alle somme A_t e A_1. Indicando con $w^1, w^2, \dots, w^z$ i pesi delle somme si ha:

$$I_t = \frac{A_t}{A_1} * (100)$$

con

$A_t = w^1 A_t^1 + w^2 A_t^2 + \cdots + w^z A_t^z$

$A_1 = w^1 A_1^1 + w^2 A_1^2 + \cdots + w^z A_1^z$.

CAPITOLO 2
Valutazione dei modelli previsionali

1.Introduzione

La qualità nelle previsioni è ovviamente molto importante,se non fondamentale,e lo è principalmente per due motivi. Innanzitutto per poter operare un confronto e quindi una scelta tra modelli posti in alternativa tra di loro. Risulta poi importante un'analisi della qualità delle previsioni,in fase di controllo e di monitoraggio,per poter stabilire se il modello è ancora efficace o se invece diviene indispensabile la sua revisione.

1.1.L'errore di previsione e l'errore di previsione percentuale

Ipotizziamo di avere a disposizione una serie storica di n osservazioni e le rispettive n previsioni,l'errore di previsione E_t al periodo t è dato da:

$$E_t = A_t - F_t$$

Allo stesso modo si definisce l'errore di previsione PE_t percentuale come:

$$PE_t = \frac{A_t - F_t}{A_t} * 100$$

1.2.Misure di distorsione

Le misure di distorsione vengono usate per distinguere i modelli sulla base degli errori medi con segno,e potendo quindi fare un confronto e successivamente una scelta:
Per primo definiamo l'errore medio *ME*:

$$ME = \frac{\sum_{i=1}^{n} E_i}{n} = \frac{\sum_{i=1}^{n}(A_i - F_i)}{n}$$

Possiamo ora definire anche l'errore percentuale medio *MPE*,ricavato dall'errore di previsione percentuale *PE*, e che è dato da:

$$MPE = \frac{\sum_{i=1}^{n} PE_i}{n}$$

Un modello è preferibile ad un altro se garantisce un errore medio e un errore percentuale più vicino a zero.

<u>1.3.Misure di dispersione</u>

Le misure di dispersione si differenziano dalle misure di distorsione perché in questo caso i modelli vengono distinti secondo l'errore medio assoluto.

Il primo tipo di misura di dispersione è lo scarto medio assoluto *MAD* che viene definito come:

$$MAD = \frac{\sum_{i=1}^{n} |E_i|}{n} = \frac{\sum_{i=1}^{n} |A_i - F_i|}{n}$$

Lo scarto percentuale medio assoluto *MAPE* è in simboli:

$$MAPE = \frac{\sum_{i=1}^{n} |PE_i|}{n}$$

Essendo lo scarto medio assoluto una misura non derivabile,viene spesso preferito a quest'ultimo lo scarto quadratico medio *MSE*,che inoltre semplifica la fase di identificazione dei parametri di un modello.

In simboli *MSE* risulta:

$$MSE = \frac{\sum_{i=1}^{n} (E_i)^2}{n} = \frac{\sum_{i=1}^{n} (A_i - F_i)^2}{n}$$

Viene invece definita la deviazione standard degli errori *SDE* come:

$$SDE = \sqrt{\frac{\sum_{i=1}^{n}(E_i)^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(A_i - F_i)^2}{n}}$$

Infine la scelta del modello da utilizzare per le previsioni dovrà ricadere sul modello che fornisce una dispersione minore.

1.4.Segnale di tracking

Il segnale di tracking TS_n è una misura di errore impiegata nella fase di controllo e monitoraggio delle previsioni e si può stimare tramite la seguente formula:

$$TS_n = \frac{\sum_{i=1}^{n} E_i}{\sum_{i=1}^{n}|E_i|}$$

Attraverso le seguenti formule ricorsive si può determinare il segnale di tracking riferito al tempo *t,*spesso usato durante la fase di monitoraggio.Si ha quindi:

$$TS_t = \left|\frac{D_t}{G_t}\right|$$

con:

$$D_t = \beta E_t + (1 - \beta) D_{t-1}$$

$$G_t = \beta E_t + (1 - \beta) G_{t-1}$$

in cui β è parametro compreso tra 0 e 1.

Il monitoraggio del modello avviene con un confronto con un livello di soglia assegnato,compreso anch'esso tra 0 e1.Se il segnale di tracking è maggiore del livello di soglia il modello dovrà essere perfezionato.

CAPITOLO 3

Modelli a Media Mobile

1.Definizioni

Si definisce *media mobile a k punti al tempo* t (in simboli Mt) la media aritmetica di k osservazioni consecutive di una data serie storica,in modo che t sia compreso all'interno dell'intervallo delle k osservazioni considerate.In base alla posizione del periodo t nella serie si possono determinare diversi valori di media mobile.
Viene definita media mobile a k punti *centrata* la media aritmetica delle osservazioni,considerando il tempo t come il punto centrale della serie storica.Vengono quindi considerati 2 diversi casi,ovvero se il numero delle osservazioni k è pari o dispari.
Se per ipotesi k è dispari si ha subito:

$$M_t = \frac{A_{t+{k-1}/_2} + A_{t+{k-1}/_{2-1}} + \cdots + A_{t-{k-1}/_2}}{k}$$

Nel caso opposto,cioè con k pari,si opera con un procedimento di calcolo diviso in 2 fasi,si deve prima centrare l' insieme di medie mobili sui punti intermedi degli intervalli temporali per poi calcolare la media mobile per gli stessi intervalli temporali,imponendo k uguale a 2.In simboli si ha:

$$M_t = \frac{A_{t+k/2} + A_{t+{k}/_{2-1}} + \cdots A_{t-{k}/_{2+1}}}{k} + \frac{A_{t+{k}/_{2-1}} + A_{t+{k}/_{2-2}} + \cdots + A_{t-{k}/_2}}{k}$$

In alcuni casi non tutte le osservazioni hanno la stessa importanza,e conviene allora assegnare dei pesi alle varie osservazioni.Si determina cosi la media mobile pesata

2.Media mobile e predizioni

La media mobile viene utilizzata per effettuare previsioni.La formula con la quale si può ricavare la predizione per il periodo t+1 è in simboli:

$$F_{t-1} = \frac{A_t + A_{t-1} + \cdots + A_{t-k+1}}{k}$$

Considerando di assegnare dei pesi ad ogni osservazione della serie storica si può anche effettuare una predizione pesata. Sia:

$$F_{t+1} = \frac{w_t A_t + w_{t-1} A_{t-1} + \cdots + w_{t-k+1} A_{t-k+1}}{\sum_{i=1}^{k} w_i}$$

Si può dimostrare che la predizione per il tempo *t+1* è la somma tra la predizione per il tempo *t* e 1/k volte la differenza tra l'osservazione più prossima *At e At-k* .La formula è

$$F_{t+1} = F_t + \frac{1}{k}(A_t - A_{t-k})$$

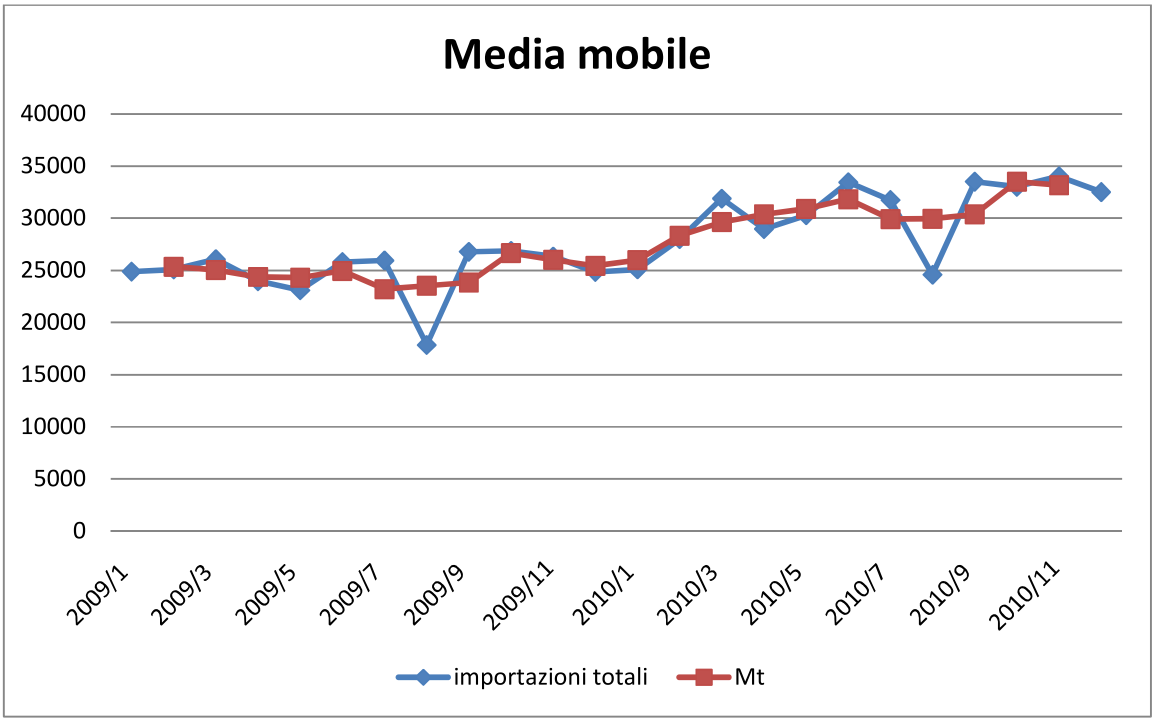
La media mobile è uno strumento che viene utilizzato per depurare la serie storica dalle componenti di stagionalità e fluttazione casuale.Vediamo nel prossimo paragrafo come ciò avviene.

<u>2.1.Esempio:la serie delle importazione totali</u>

Cerchiamo di determinare la media mobile centrata,con k=3,per la serie relativa alle importazioni totali per gli anni 2009 e 2010,evidenziandone l'andamento grafico.

t	importazioni totali	Mt
2009/1	24890,59907	
2009/2	25088,55162	25351,14
2009/3	26074,27644	25043,73
2009/4	23968,3531	24381,15
2009/5	23100,81515	24294,78
2009/6	25815,17748	24955,81
2009/7	25951,44068	23205,6
2009/8	17850,19228	23532
2009/9	26794,37182	23841,42
2009/10	26879,6974	26669,8

2009/11	26335,33123	26024,96
2009/12	24859,85682	25429,66
2010/1	25093,80233	25991,68
2010/2	28021,3764	28334,92
2010/3	31889,58024	29629,28
2010/4	28976,88092	30386,59
2010/5	30293,29839	30904,76
2010/6	33444,09968	31827,66
2010/7	31745,58005	29925,72
2010/8	24587,47049	29950,75
2010/9	33519,21307	30374,08
2010/10	33015,57142	33519,38
2010/11	34023,36038	33183,42
2010/12	32511,32799	



Ponendoci nel periodo di dicembre 2010 possiamo anche determinare la predizione per il periodo di gennaio 2011:

$$F_{2011/1} = \frac{A_{2010/12} + A_{2010/11} + A_{2010/10}}{3} = \frac{33015{,}57 + 34023{,}36 + 32511{,}32}{3} = 33183{,}42$$

3.Scomposizione di una serie storica

La scomposizione di una serie storica è un processo che riguarda soprattutto l'analisi delle serie storiche ,ma che permette anche di poter fare previsioni sull'andamento futuro. Questo procedimento consta nell'identificazione delle quattro componenti principali di una serie,ovvero tendenza ciclicità stagionalità e fluttuazioni casuali. Innanzitutto bisogna ricavare una forma di dipendenza di A_t dalle sue quattro componenti:si può analizzare un modello additivo o in alternativa il modello moltiplicativo,che in simboli sono

$$A_t = T_t + C_t + Q_t + \varepsilon_t$$
$$A_t = T_t * C_t * Q_t * \varepsilon_t$$

Si può anche considerare un modello che sia una combinazione tra i modelli precedenti. Di seguito analizzeremo un modello moltiplicativo ricavandone anche la formula per effettuare delle predizioni.

3.1.Scomposizione di una serie storica nel caso moltiplicativo

Si inizia il procedimento determinando la componente congiunta di tendenza e ciclicità,con il calcolo della media mobile M_t

$$M_t \cong T_t * C_t$$

Successivamente si calcola la componente congiunta di stagionalità e di fluttuazione casuale:

$$Q_t * \varepsilon_t = \frac{A_t}{T_t * C_t} = \frac{A_t}{M_t}$$

Ipotizziamo che la serie storica sia identificabile una stagionalità pari a *L*,cioè che si abbiano *L* periodi per ogni ciclo,ad esempio sette giorni in una settimana o dodici mesi in un anno. A questo punto è possibile introdurre gli indici di stagionalità Q_l (l=1,2,….,*L*),che equivalgono alla media dei prodotti Q_t*ε_t per i periodi analoghi a *l*,così da eliminare l'effetto derivante dalle fluttuazioni casuali:

$$Q_l = \frac{\sum_{t \in n_l} Q_l * \varepsilon_l}{|n_l|}$$

Nella formula n_l rappresenta il numero degli periodi analoghi a *l.*
Nel caso dell'esportazioni su base mensile per esempio i periodi analoghi al mese dicembre sono tutti i mesi di dicembre compresi nella serie.
Si può quindi ora destagionalizzare la serie,dividendo ogni osservazione per il relativo indice di stagionalità:

$$\frac{A_t}{Q_{l(t)}} = T_t * C_t * \varepsilon_t$$

con *l(t)* che rappresenta il tipo di periodo corrispondente a *t*.
Attraverso la formulazione di una retta di regressione delle osservazioni in funzione del tempo è possibile ricavare la componente di tendenza. Se si ipotizza un legame di dipendenza lineare si determina la retta di predizione da:

$$T_t = a + bt$$

Infine per identificare la componente di ciclicità basta togliere dalla serie le altre tre componenti. In simboli diventa:

$$C_t = \frac{T_t * C_t}{T_t} \cong \frac{M_t}{T_t}$$

In conclusione si possono effettuare delle predizioni future basandosi sulla scomposizione della serie storiche. La predizione per gli L periodi successivi si può ricavare dalle componenti di tendenza e ciclicità future e gli indici di stagionalità. In simboli si ha:

$$F_{t+i} = T_{t+i} * C_{t+i} * Q_{l(t+i)}$$

con i=1,2,…,L.

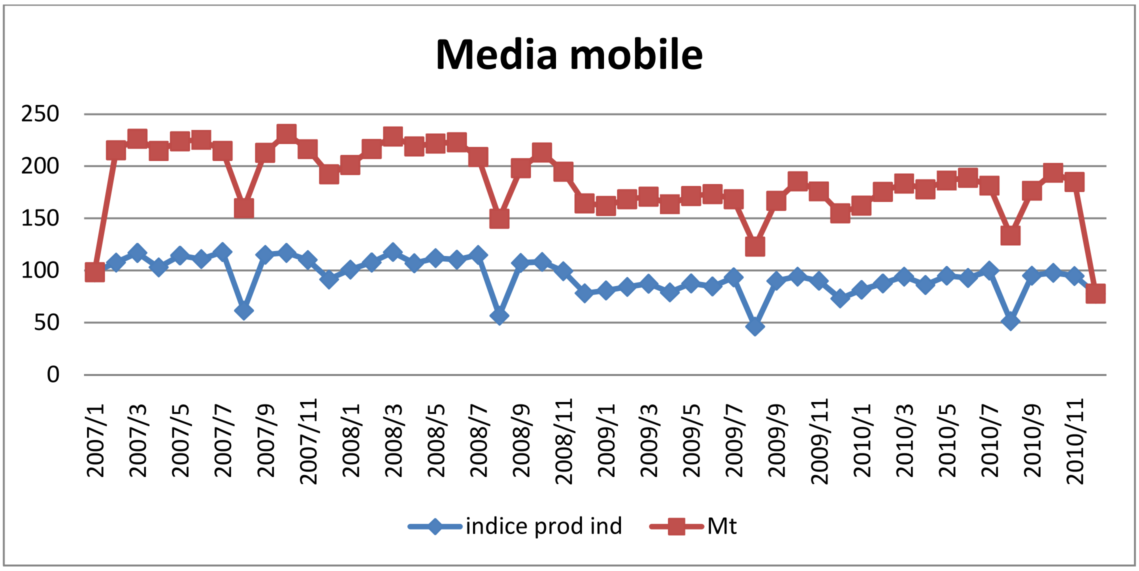
3.1.1.Esempio:Indice della produzione industriale

In questo paragrafo analizziamo una serie storica dell'indice della produzione industriale

su base mensile dall’anno 2007 all’anno 2010 cercando di ottenere una predizione per il periodo di gennaio 2011 con il procedimento descritto nella sezione precedente.
La successiva tabella contiene i valori delle osservazioni della serie storica e il calcolo della media mobile ottenuta ponendo k=3,con di seguito la relativa rappresentazione grafica.

t	indice prod industriale	Mt
2007/1	98,5	
2007/2	107,7	107,7333
2007/3	117	109,2667
2007/4	103,1	111,5
2007/5	114,4	109,4667
2007/6	110,9	114,4
2007/7	117,9	96,8
2007/8	61,6	98,2
2007/9	115,1	97,9
2007/10	117	114,1
2007/11	110,2	106,2333
2007/12	91,5	100,9
2008/1	101	100,1
2008/2	107,8	108,8333
2008/3	117,7	110,8
2008/4	106,9	112,2
2008/5	112	109,8
2008/6	110,5	112,4667
2008/7	114,9	94,03333
2008/8	56,7	92,96667
2008/9	107,3	90,73333
2008/10	108,2	105
2008/11	99,5	95,33333
2008/12	78,3	86,23333
2009/1	80,9	81,13333
2009/2	84,2	84,13333
2009/3	87,3	83,5
2009/4	79	84,63333
2009/5	87,6	83,8
2009/6	84,8	88,63333
2009/7	93,5	74,86667
2009/8	46,3	76,63333

2009/9	90,1	76,86667
2009/10	94,2	91,46667
2009/11	90,1	85,83333
2009/12	73,2	81,6
2010/1	81,5	80,8
2010/2	87,7	87,8
2010/3	94,2	89,36667
2010/4	86,2	91,8
2010/5	95	91,36667
2010/6	92,9	95,96667
2010/7	100	81,43333
2010/8	51,4	82,16667
2010/9	95,1	81,43333
2010/10	97,8	95,9
2010/11	94,8	90,13333
2010/12	77,8	



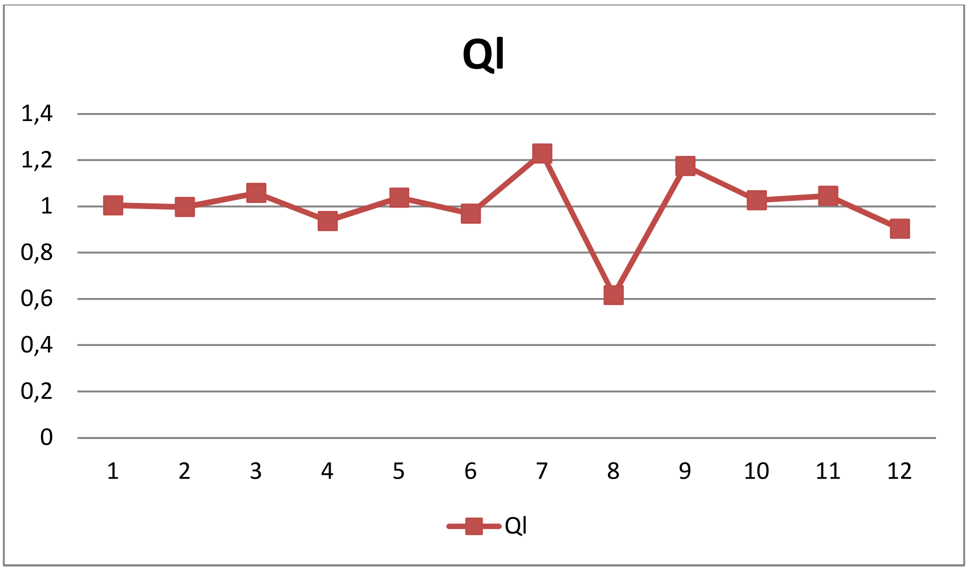
E' ora possibile determinare la componente congiunta di stagionalità e fluttuazione casuale,con il rapporto tra i valori osservati e il corrispondente valore della media mobile. Successivamente si calcolano gli indici di stagionalità $Q_{l.}$.Le tabelle seguenti riportano entrambi i dati:

t	indice prod industriale	Mt	Qt*εt=At/Mt
2007/1	98,5		

2007/2	107,7	107,7333	0,999690594
2007/3	117	109,2667	1,070774863
2007/4	103,1	111,5	0,924663677
2007/5	114,4	109,4667	1,045066991
2007/6	110,9	114,4	0,969405594
2007/7	117,9	96,8	1,217975207
2007/8	61,6	98,2	0,627291242
2007/9	115,1	97,9	1,175689479
2007/10	117	114,1	1,025416301
2007/11	110,2	106,2333	1,03733919
2007/12	91,5	100,9	0,906838454
2008/1	101	100,1	1,008991009
2008/2	107,8	108,8333	0,99050536
2008/3	117,7	110,8	1,062274368
2008/4	106,9	112,2	0,952762923
2008/5	112	109,8	1,02003643
2008/6	110,5	112,4667	0,982513337
2008/7	114,9	94,03333	1,221907125
2008/8	56,7	92,96667	0,60989602
2008/9	107,3	90,73333	1,182586334
2008/10	108,2	105	1,03047619
2008/11	99,5	95,33333	1,043706294
2008/12	78,3	86,23333	0,908001546
2009/1	80,9	81,13333	0,997124076
2009/2	84,2	84,13333	1,000792393
2009/3	87,3	83,5	1,045508982
2009/4	79	84,63333	0,933438362
2009/5	87,6	83,8	1,045346062
2009/6	84,8	88,63333	0,956750658
2009/7	93,5	74,86667	1,24888691
2009/8	46,3	76,63333	0,604175729
2009/9	90,1	76,86667	1,172159584
2009/10	94,2	91,46667	1,029883382
2009/11	90,1	85,83333	1,049708738
2009/12	73,2	81,6	0,897058824
2010/1	81,5	80,8	1,008663366
2010/2	87,7	87,8	0,998861048
2010/3	94,2	89,36667	1,054084297
2010/4	86,2	91,8	0,938997821
2010/5	95	91,36667	1,039766509
2010/6	92,9	95,96667	0,96804446

2010/7	100	81,43333	1,227998363
2010/8	51,4	82,16667	0,625557809
2010/9	95,1	81,43333	1,167826443
2010/10	97,8	95,9	1,019812304
2010/11	94,8	90,13333	1,051775148
2010/12	77,8		

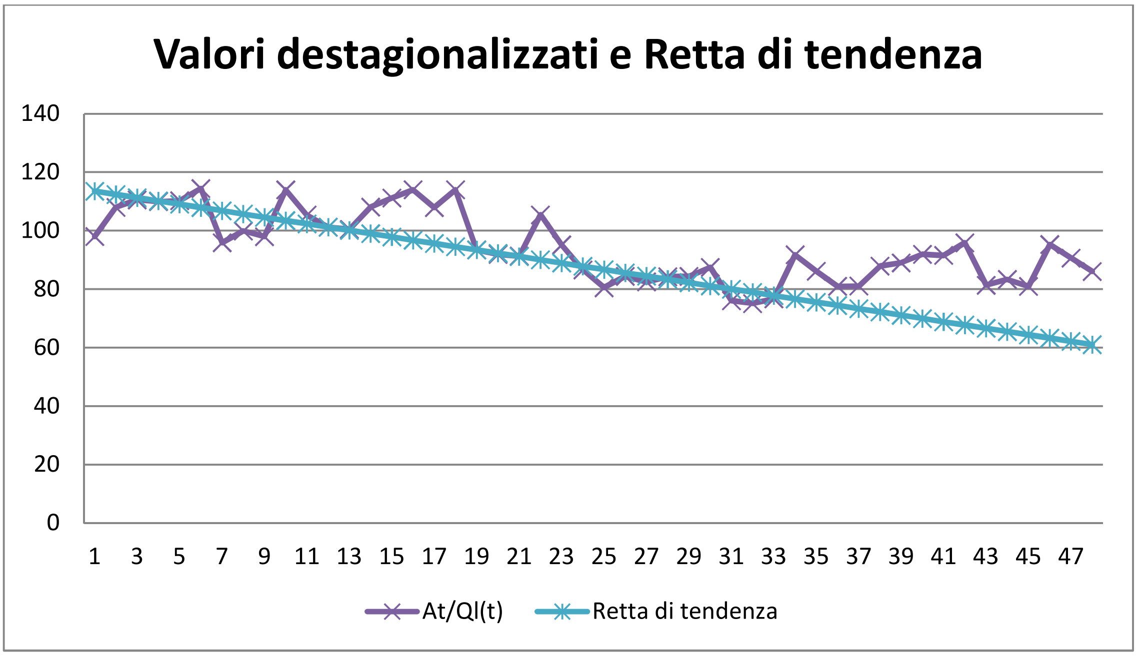
L	QI
1	1,004926
2	0,997462
3	1,058161
4	0,937466
5	1,037554
6	0,969179
7	1,229192
8	0,61673
9	1,174565
10	1,026397
11	1,045632
12	0,903966



Si deve ora destagionalizzare la serie dividendo ogni osservazione per il relativo indice di stagionalità e sarà quindi possibile determinare la componente di tendenza e di ciclicità mediante una curva di regressione,stimata per entrambi i casi con l'equazione geometrica di una retta passante per due punti.Si ha:

t	At/QI(t)	Retta di tendenza	Ct=Mt/Tt	Retta di ciclicità
2007/1	98,01715277	113,5232		
2007/2	107,9740004	112,4064	0,958427	0,999
2007/3	110,5692245	111,2896	0,981823	1,005
2007/4	109,9773575	110,1728	1,012047	1,011
2007/5	110,2593217	109,056	1,003766	1,017
2007/6	114,4268043	107,9392	1,059856	1,023
2007/7	95,91667492	106,8224	0,906177	1,029
2007/8	99,88160137	105,7056	0,928995	1,035
2007/9	97,99368697	104,5888	0,936047	1,041
2007/10	113,9909751	103,472	1,102714	1,047
2007/11	105,3907722	102,3552	1,037889	1,053
2007/12	101,2205904	101,2384	0,996657	1,059
2008/1	100,5048978	100,1216	0,999784	1,065
2008/2	108,0742548	99,0048	1,099273	1,071
2008/3	111,2307498	97,888	1,131906	1,077
2008/4	114,0308392	96,7712	1,159436	1,083
2008/5	107,946189	95,6544	1,147882	1,089
2008/6	114,0140837	94,5376	1,18965	1,095
2008/7	93,47604706	93,4208	1,006557	1,101
2008/8	91,93647399	92,304	1,007179	1,107
2008/9	91,35293321	91,1872	0,995023	1,113
2008/10	105,417295	90,0704	1,165755	1,119
2008/11	95,15772988	88,9536	1,07172	1,125
2008/12	86,61827571	87,8368	0,981745	1,131
2009/1	80,50342801	86,72	0,935578	1,137
2009/2	84,41421384	85,6032	0,982829	1,143
2009/3	82,50165214	84,4864	0,988325	1,149
2009/4	84,26975019	83,3696	1,015158	1,155
2009/5	84,42934071	82,2528	1,01881	1,161
2009/6	87,49678095	81,136	1,092405	1,167
2009/7	76,06623499	80,0192	0,935609	1,173
2009/8	75,07334649	78,9024	0,971242	1,179
2009/9	76,70921978	77,7856	0,988186	1,185
2009/10	91,77734922	76,6688	1,19301	1,191
2009/11	86,16795439	75,552	1,136083	1,197
2009/12	80,97647231	74,4352	1,096256	1,203
2010/1	81,10048681	73,3184	1,102043	1,209
2010/2	87,92311822	72,2016	1,21604	1,215
2010/3	89,02240128	71,0848	1,257184	1,221

2010/4	91,95003122	69,968	1,312028	1,227
2010/5	91,56149963	68,8512	1,327016	1,233
2010/6	95,85437441	67,7344	1,416808	1,239
2010/7	81,35426202	66,6176	1,2224	1,245
2010/8	83,34276478	65,5008	1,254438	1,251
2010/9	80,96611322	64,384	1,264807	1,257
2010/10	95,28476384	63,2672	1,515793	1,263
2010/11	90,66284214	62,1504	1,450245	1,269
2010/12	86,06515773	61,0336		1,275
2011/1		59,9168		1,281



E’ possibile infine ricavare la predizione per il mese di gennaio 2011. Pongo quindi *i=1* e ottengo

$$F_{2011/1} = T_{2011/1} * C_{2011/1} * Q_{l(1)} = 59{,}91 * 1{,}281 * 1{,}004 = 77{,}1315$$

CAPITOLO 4

Modelli di smoothing esponenziale

I modelli di smoothing esponenziale sono un metodo semplice ed efficace per la formulazione di previsione future,anche nelle sue versioni più elaborate.
Questi tipi di modelli possono essere identificati come una generalizzazione dei modelli a media mobile,e abbiamo gia visto che per questo tipo di modelli vale la seguente relazione:

$$F_{t+1} = F_t + \frac{A_t}{k} - \frac{A_{t-k}}{k}$$

Inoltre se il modello non si discosta troppo dalle osservazioni della serie storica,vale

$$\frac{F_t}{k} \cong \frac{A_{t-k}}{k}$$

da cui ricaviamo $F_t \cong A_{t-k}$,e sostituendo nella prima equazione otteniamo

$$F_{t+1} \cong F_t + \frac{A_t}{k} - \frac{F_t}{k} = \frac{1}{k}A_t + \left(1 - \frac{1}{k}\right)F_t$$

dalla quale si determina:

$$F_{t+1} \cong \alpha A_t + (1-\alpha)F_t$$

e sviluppando quest’ultima relazione ricaviamo infine:

$$F_{t+1} = F_t + \alpha(A_t - F_t) = F_t + \alpha E_t$$

che esprime una proprietà di *feedback negativo.* Si ha cioè che la previsione per il tempo *t+1* è data dalla previsione per il tempo *t* corretta per una frazione pari ad α dell'errore di previsione *E* al tempo *t* .Il parametro α è compreso in senso lato tra 0 e 1 e come vedremo successivamente sarà possibile ricavarne una interpretazione.

4.1.Smoothing esponenziale semplice

Un modello che rispetta la relazione di feedback negativo è il modello esponenziale semplice,o modello di Brown.

Dato il parametro α introduciamo la media smorzata al tempo *t*,definita,con $S_1=A_1$, come:

$$S_t = \alpha A_t + (1-\alpha)S_{t-1}$$

la previsione per il tempo *t+1* si determina con

$$F_{t+1} = S_t$$

Per il modello di smoothing esponenziale valgono tutte le successive relazioni:

$S_1 = A_1$

$S_2 = \alpha A_2 + (1-\alpha)S_1$

$S_3 = \alpha A_3 + (1-\alpha)S_2$

……………………

$S_t = \alpha A_t + (1-\alpha)S_{t-1}$

e se per ipotesi si dispone di n osservazioni totali si può scrivere la seguente formula,dalla quale è possibile,come accennato in precedenza,ottenere un interpretazione di α:

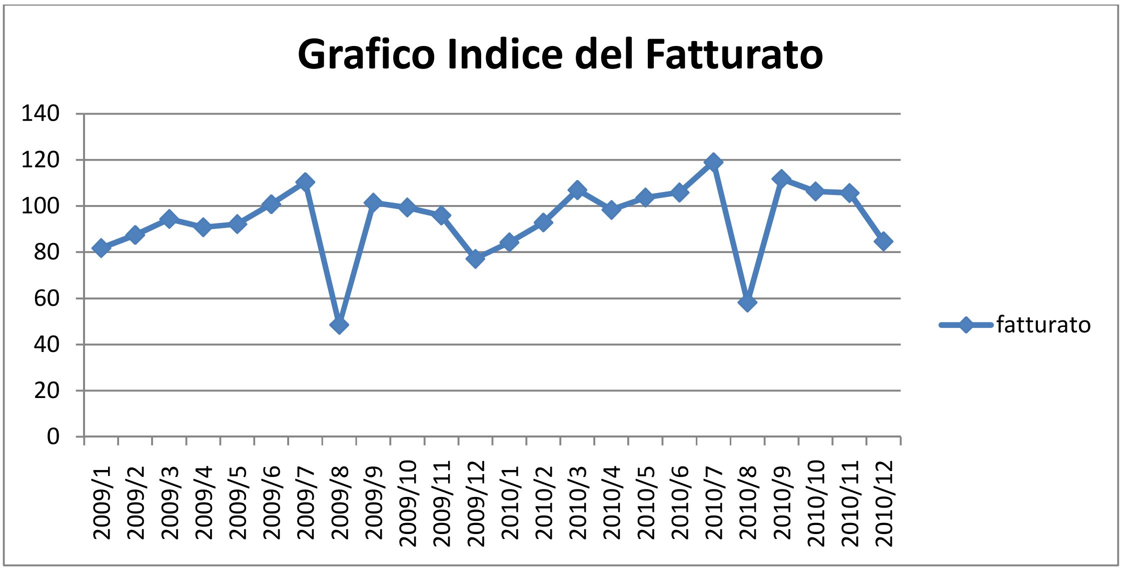
$F_{t+1} = \alpha A_t + \alpha(1-\alpha)A_{t-1} + \alpha(1-\alpha)^2 A_{t-1} + \cdots + \alpha(1-\alpha)^{n-1}A_{t-n+1} = S_t$

Se α tende a 0 il modello è poco reattivo ,cioè assegna lo stesso peso a tutte le osservazioni passate,viceversa se α tende a 1 si ha un modello più reattivo,che assegna cioè un peso maggiore in corrispondenza delle osservazioni più recenti. La scelta di α deve essere effettuata in maniera da minimizzare lo scarto quadratico medio o un qualsiasi altra misura di dispersione già vista.

4.1.1.Esempio:indice fatturato nel settore della produzione di materie plastiche

Con il modello di Brown analizziamo la serie storica su base mensile dell'indice del fatturato nel settore delle materie plastiche,per gli anni 2009 e 2010 .Di seguito rappresento la serie con una tabella e un grafico.

t	fatturato
2009/1	81,8
2009/2	87,5
2009/3	94,4
2009/4	90,9
2009/5	92,2
2009/6	100,8
2009/7	110,4
2009/8	48,6
2009/9	101,5
2009/10	99,4
2009/11	96
2009/12	77,2
2010/1	84,3
2010/2	92,9
2010/3	107
2010/4	98,4
2010/5	103,7
2010/6	105,9
2010/7	119
2010/8	58,3
2010/9	111,8
2010/10	106,4
2010/11	105,7
2010/12	84,7



Attribuendo al parametro α il valore di 0,5(ottenuto con una stima e non minimizzando lo scarto quadratico medio) posso ricavare la media smorzata S_t ,che ricordiamo è data da:

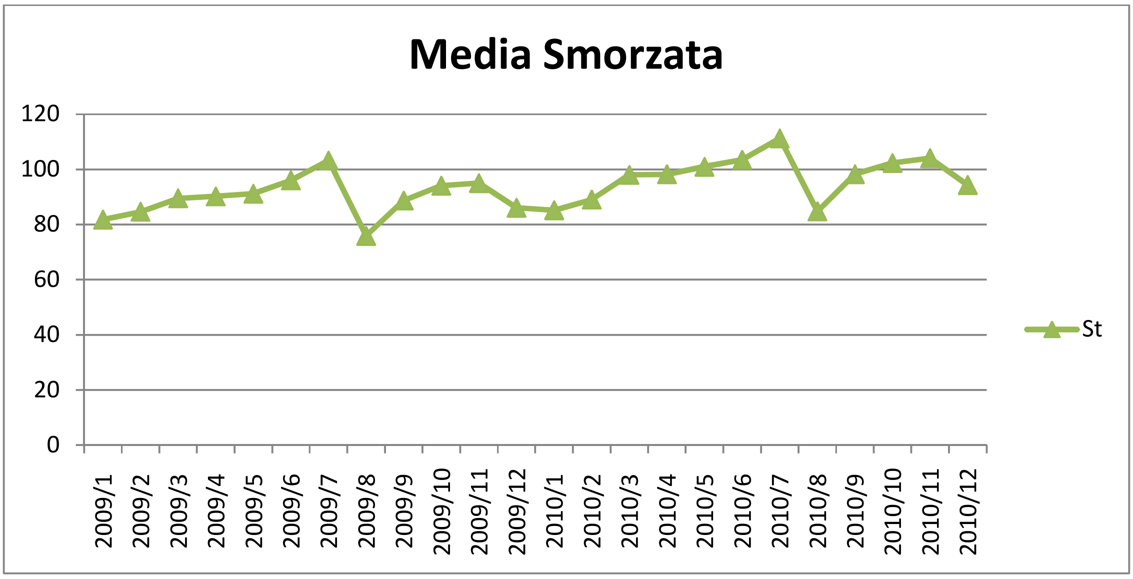
$$S_t = \alpha A_t + (1 - \alpha) S_{t-1}$$

e la previsione per il periodo F_{t+1}:

$$F_{t+1} = S_t$$

t	fatturato	St
2009/1	81,8	81,8
2009/2	87,5	84,65
2009/3	94,4	89,525
2009/4	90,9	90,2125
2009/5	92,2	91,2062
2009/6	100,8	96,0031
2009/7	110,4	103,2015
2009/8	48,6	75,9007
2009/9	101,5	88,7003
2009/10	99,4	94,0501
2009/11	96	95,0250
2009/12	77,2	86,1125
2010/1	84,3	85,2062
2010/2	92,9	89,0531

2010/3	107	98,0265
2010/4	98,4	98,2132
2010/5	103,7	100,9566
2010/6	105,9	103,4283
2010/7	119	111,2141
2010/8	58,3	84,7570
2010/9	111,8	98,2785
2010/10	106,4	102,3392
2010/11	105,7	104,0196
2010/12	84,7	94,3598



La predizione per il periodo di gennaio 2011 è data da:

$$F_{2011/1} = S_{2010/12} = 94{,}3598$$

4.2.Smoothing con tendenza lineare

Se la serie storica in esame presenta una componente di tendenza,il modello di smoothing esponenziale semplice non è capace di individuarne la presenza e risulterà quindi sempre in ritardo rispetto alle osservazioni della serie storica,fornendo quindi delle previsioni distorte.

Esiste comunque la possibilità di ampliare il modello di smoothing esponenziale semplice,introducendo una componente di tendenza in modo da ottenere un modello con

correzione di tendenza,o modello di Holt.
In questo caso oltre alla media smorzata S_t si definisce una tendenza smorzata apparente lineare T_t (ma è anche possibile inserire correzioni di tendenza quadratiche o esponenziali),che permette di approssimare la componente additiva di tendenza. In simboli si ha:

$$S_t = \alpha A_t + (1-\alpha)(S_{t-1} + T_{t-1})$$
$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

per il parametro β valgono le stesse considerazioni fatte per il parametro α,ed è quindi anch'esso compreso in senso lato tra 0 e 1 e si ha che se β tende a 0 il modello assegna un peso uguale alle tendenze passate della serie,mentre se tende a 1 si stabilisce un peso maggiore alle tendenze più prossime. Inoltre anche per il modello di Holt la determinazione di α e β deve essere effettuata in modo da ridurre il più possibile le misure di dispersione.
La previsione per il tempo *t+1* si ricava da:

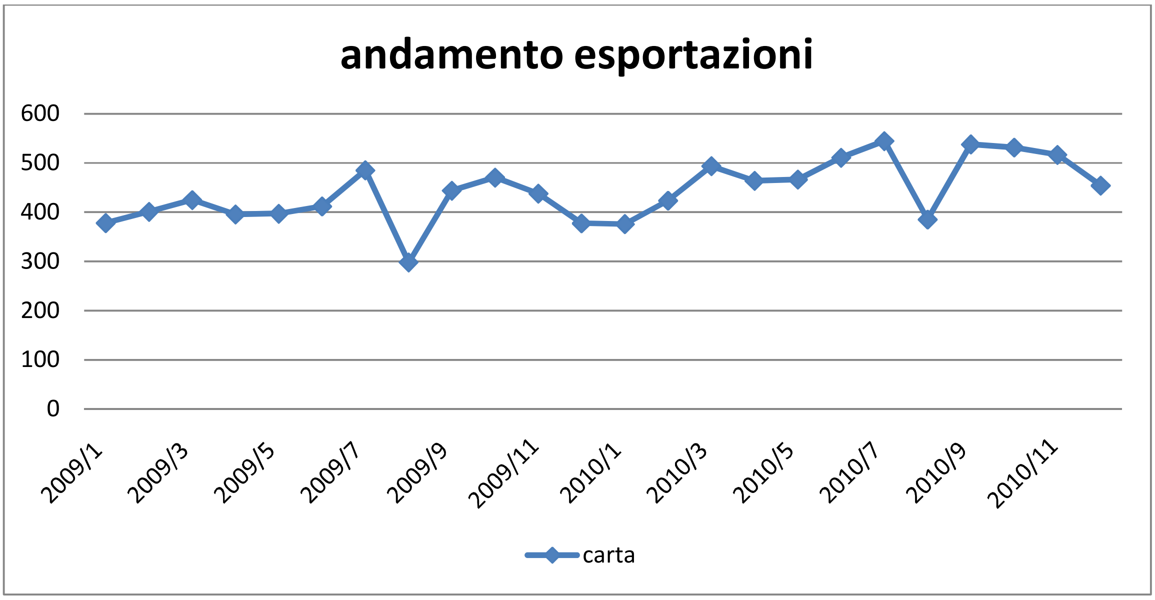
$$F_{t+1} = S_t + T_t$$

4.2.1.Esempio:esportazioni di carta e prodotti di carta
Sempre dal sito internet dell'istituto Istat (www.con.istat.it) dal quale è possibile ricavare delle serie storiche, consideriamo in questa sezione una serie storica,sempre per gli anni 2009 e 2010 e sempre su base mensile,relativa all'esportazioni di carta e di prodotti di carta. Ricaveremo una previsione per il periodo di gennaio 2011 mediante il metodo dello smoothing esponenziale,applicando il modello di Holt.
La serie in esame è rappresentata nella tabella seguente:

t	carta
2009/1	378,12937
2009/2	400,91492
2009/3	424,82471
2009/4	395,52331
2009/5	397,1528
2009/6	411,8641
2009/7	485,09336

2009/8	298,17024
2009/9	443,99553
2009/10	470,44616
2009/11	437,91295
2009/12	377,39256
2010/1	375,85346
2010/2	423,54296
2010/3	493,72131
2010/4	463,78666
2010/5	466,42403
2010/6	510,92407
2010/7	544,59288
2010/8	385,17821
2010/9	537,76156
2010/10	531,46978
2010/11	516,61316
2010/12	454,18085

Visualizziamo anche la rappresentazione grafica della serie:

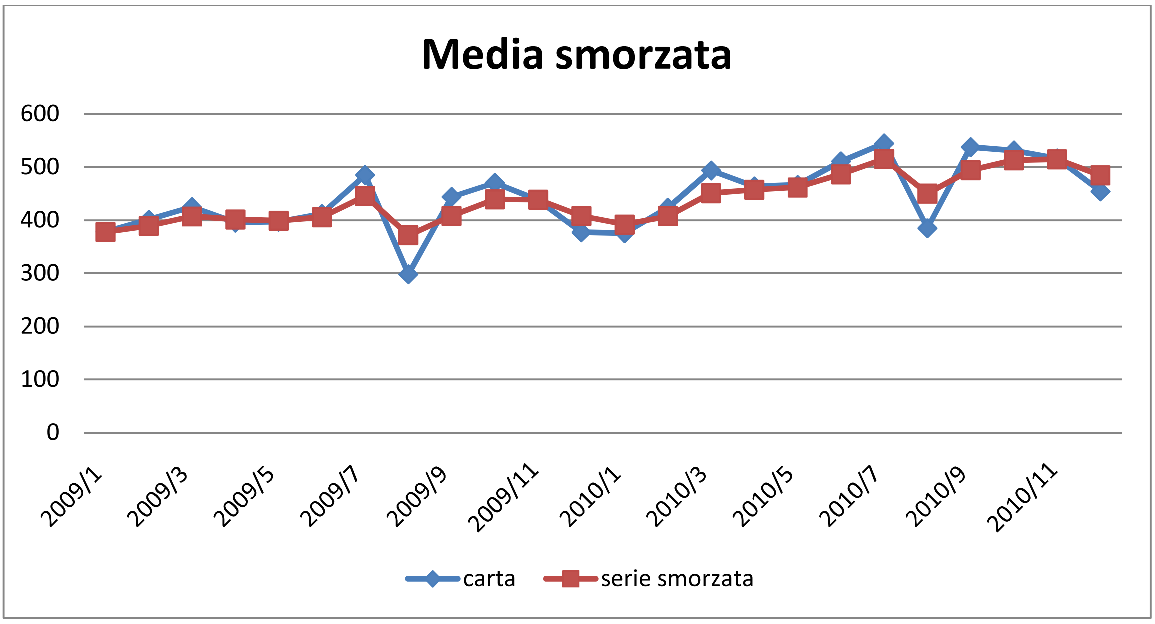


Per prima cosa,data la serie storica,dobbiamo ricavarne la serie smorzata S_t ,che ricordiamo è data da

$$S_t = \alpha A_t + (1 - \alpha) S_{t-1}$$

In tabella raggruppiamo tutti i valori della serie smorzata corrispondenti ai periodi della serie storica,che andremo poi a rappresentare anche in un grafico:

t	carta	media smorzata
2009/1	378,129373	378,1294
2009/2	400,914918	389,522159
2009/3	424,824708	407,1734335
2009/4	395,523308	401,3483708
2009/5	397,152804	399,2505874
2009/6	411,864103	405,5573452
2009/7	485,093362	445,3253536
2009/8	298,170242	371,7477978
2009/9	443,995533	407,8716654
2009/10	470,446163	439,1589142
2009/11	437,912945	438,5359296
2009/12	377,392564	407,9642468
2010/1	375,853456	391,9088514
2010/2	423,542961	407,7259062
2010/3	493,721311	450,7236086
2010/4	463,78666	457,2551343
2010/5	466,424032	461,8395831
2010/6	510,924073	486,3818281
2010/7	544,592877	515,4873525
2010/8	385,178205	450,3327788
2010/9	537,761564	494,0471714
2010/10	531,469782	512,7584767
2010/11	516,613159	514,6858178
2010/12	454,180846	484,4333319



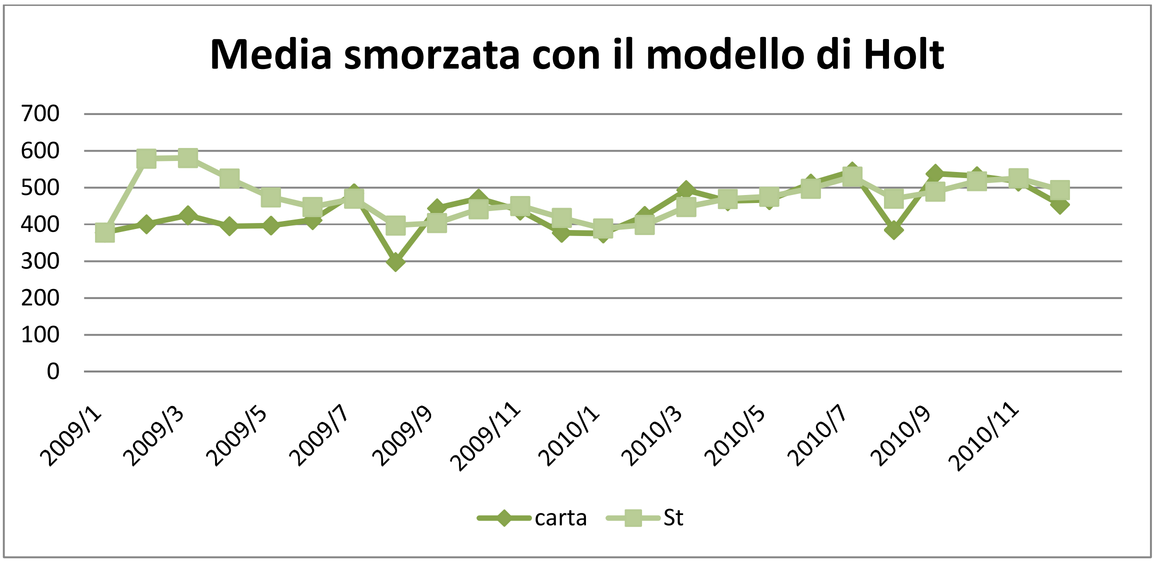
Andiamo ora a determinare la componente di tendenza della serie,che si può ricavare, come descritto in precedenza,da:

$$S_t = \alpha A_t + (1-\alpha)(S_{t-1} + T_{t-1})$$
$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

Anche in questo esempio la scelta dei parametri è dovuta a una semplice stima personale. Ora pongo i dati in una tabella,che contiene oltre alla serie storica originaria la componente di tendenza smorzata T_t e la nuova serie smorzata calcolata considerando T_t

t	carta	T_t	S_t	F_{t+1}
2009/1	378,12937	378,1294	378,12937	
2009/2	400,9149	164,923087	389,522144	756,25877
2009/3	424,82471	80,3151088	407,173426	554,445231
2009/4	395,52331	14,54520352	401,348367	487,4885348
2009/5	397,1528	6,795779008	399,2505855	415,8935705
2009/6	411,8641	11,545091	405,5573443	406,0463645
2009/7	485,09336	48,5555918	445,3253531	417,1024353

2009/8	298,17024	-92,7316353	371,7477976	493,8809449
2009/9	443,99553	50,40252049	407,8716653	279,0161623
2009/10	470,44616	36,0313862	439,1589141	458,2741858
2009/11	437,91295	-5,10737632	438,5359296	475,1903003
2009/12	377,39256	-38,3551791	407,9642468	433,4285532
2010/1	375,85346	-16,2655365	391,9088514	369,6090677
2010/2	423,54296	22,10748842	407,7259062	375,6433149
2010/3	493,72131	50,95000537	450,7236086	429,8333946
2010/4	463,78666	2,419211547	457,2551343	501,673614
2010/5	466,42403	2,550107819	461,8395831	459,6743458
2010/6	510,92407	27,72006773	486,3818281	464,389691
2010/7	544,59288	31,28930949	515,4873525	514,1018958
2010/8	385,17821	-83,1330794	450,3327788	546,776662
2010/9	537,76156	58,29678364	494,0471714	367,1996994
2010/10	531,46978	19,54364426	512,7584767	552,343955
2010/11	516,61316	-1,0965161	514,6858178	532,3021209
2010/12	454,18085	-37,8979942	484,4333319	513,5893017



La predizione per il mese successivo è quindi:

$$F_{2011/1} = S_{2010/12} + T_{2010/12} = 484{,}433 + (-37{,}898) = 444{,}535$$

4.3.Smoothing con tendenza e stagionalità

Nel caso in cui oltre a una componente di tendenza sia presente anche la componente di stagionalità nella serie storica in esame,è possibile modificare ulteriormente il modello di smoothing esponenziale,ottenendo il modello di Winters.

Si definisce un indice di stagionalità smorzato Q_t che consente di approssimare la componente moltiplicativa di stagionalità.

Se per ipotesi abbiamo L periodi per ogni ciclo si può scrivere:

$$S_t = \alpha \frac{A_t}{Q_{t-L}} + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$Q_t = \gamma \frac{A_t}{S_t} + (1-\gamma)Q_{t-L}$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

Anche per γ vale lo stesso discorso fatto in precedenza per gli altri due parametri α e β.Infatti anche per γ vale $0 \le \gamma \le 1$ ed è possibile ricavare l'andamento del modello dal suo valore:si ha che se γ tende a 0 si dà un peso costante alle stagionalità passate,mentre se tende a 1 si assegna un peso superiore alle stagionalità più recenti.

Anche per il modello di Winters i parametri considerati devono essere scelti in maniera da minimizzare lo scarto quadratico medio e le atre misure di dispersione.

La previsione per il periodo *t+1* si determina dalla seguente formula:

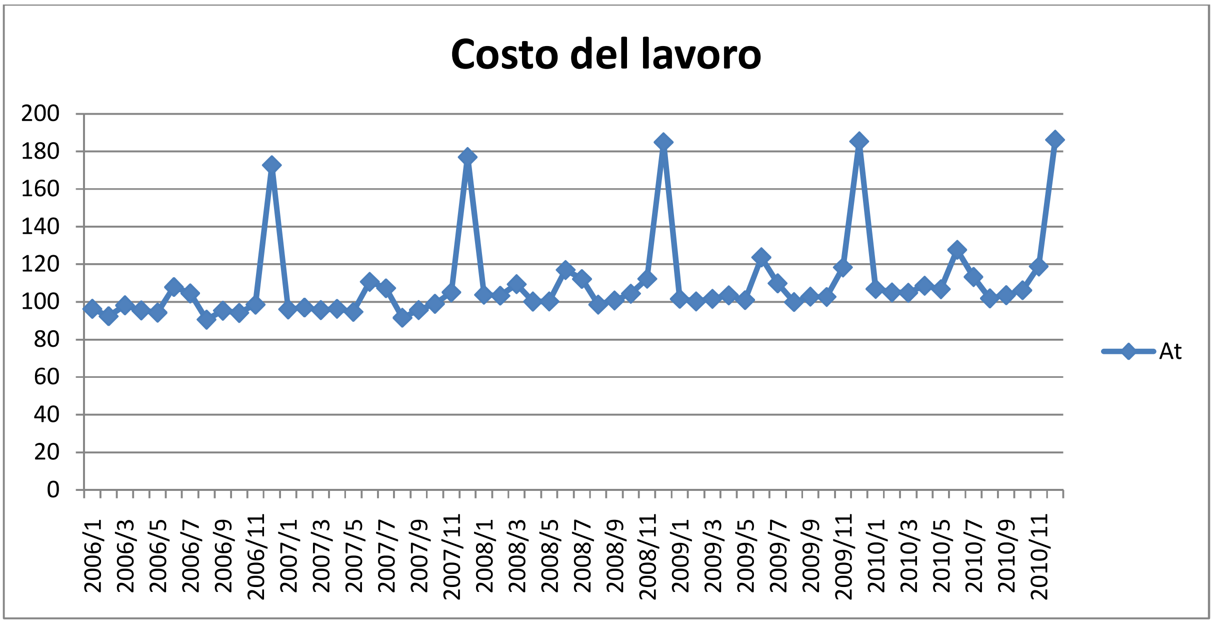
$$F_{t+1} = (S_t + T_t)Q_{t-L+1}$$

4.3.1.Esempio:costo del lavoro nel settore tessile

Come per i modelli precedenti di smoothing esponenziale anche per il modello di Winters analizziamo una serie storica,cercando di ottenere una predizione per il periodo *t+1*. In questo caso si tratta della serie storica dall'anno 2006 al 2010 relativa al costo del lavoro nel settore tessile,su base mensile. Anche in questo caso,attraverso una stima,ho posto i valori dei coefficienti pari a 0,5.

Di seguito presento la serie in esame sia con una sia con un grafico,come negli esempi precedenti:

t	At		
2006/1	96,3	2009/1	101,6
2006/2	92,4	2009/2	100,3
2006/3	98,2	2009/3	101,7
2006/4	95,5	2009/4	103,5
2006/5	94,3	2009/5	101
2006/6	107,9	2009/6	123,7
2006/7	104,6	2009/7	109,9
2006/8	90,6	2009/8	99,9
2006/9	95,4	2009/9	102,8
2006/10	94,2	2009/10	102,7
2006/11	98,6	2009/11	118,4
2006/12	172,7	2009/12	185,4
2007/1	96,1	2010/1	106,9
2007/2	97	2010/2	105,1
2007/3	95,7	2010/3	104,9
2007/4	96,4	2010/4	108,6
2007/5	94,7	2010/5	106,8
2007/6	110,7	2010/6	127,7
2007/7	107,3	2010/7	113,3
2007/8	91,6	2010/8	101,9
2007/9	95,7	2010/9	103,6
2007/10	99	2010/10	106,2
2007/11	105,1	2010/11	118,9
2007/12	177	2010/12	186,2
2008/1	103,8		
2008/2	103,3		
2008/3	109,4		
2008/4	100,2		
2008/5	100,3		
2008/6	117		
2008/7	112,2		
2008/8	98,6		
2008/9	100,8		
2008/10	104,4		
2008/11	112,3		
2008/12	184,9		



Ho calcolato innanzitutto la serie smorzata secondo il modello di Holt la media smorzata S_t e la tendenza smorzata T_t .

t	At	Tt(holt)	St(holt)
2006/1	96,3	96,3	96,3
2006/2	92,4	47,175	142,5
2006/3	98,2	24,55	143,9375
2006/4	95,5	12,08125	131,99375
2006/5	94,3	5,64375	119,1875
2006/6	107,9	6,02343	116,3656
2006/7	104,6	3,7875	113,4945
2006/8	90,6	-1,21835	103,9410
2006/9	95,4	-0,96523	99,0613
2006/10	94,2	-0,96064	96,1480
2006/11	98,6	0,38066	96,8937
2006/12	172,7	19,1458	134,9871
2007/1	96,1	-0,0993	125,1165
2007/2	97	-4,6607	111,0085
2007/3	95,7	-4,9609	101,0238
2007/4	96,4	-3,6207	96,2314
2007/5	94,7	-2,8055	93,6553
2007/6	110,7	2,0996	100,7749
2007/7	107,3	1,9510	105,0872

2007/8	91,6	-2,4988	99,3191
2007/9	95,7	-1,9616	96,2601
2007/10	99	-0,5119	96,64926
2007/11	105,1	1,5034	100,6186
2007/12	177	19,6064	139,5610
2008/1	103,8	0,93059	131,4837
2008/2	103,3	-4,0960	117,8571
2008/3	109,4	-2,8036	111,5805
2008/4	100,2	-4,0796	104,4884
2008/5	100,3	-3,3537	100,3543
2008/6	117	1,8411	107,0003
2008/7	112,2	1,4796	110,5207
2008/8	98,6	-2,3806	105,3001
2008/9	100,8	-2,2005	101,8597
2008/10	104,4	-0,7054	102,0295
2008/11	112,3	1,8197	106,8120
2008/12	184,9	20,1460	146,7659
2009/1	101,6	-1,1338	134,2559
2009/2	100,3	-6,4953	116,7110
2009/3	101,7	-5,8619	105,9578
2009/4	103,5	-3,7880	101,7979
2009/5	101	-2,947588	99,50495
2009/6	123,7	3,674428	110,1286
2009/7	109,9	0,96132	111,8515
2009/8	99,9	-2,4572	106,3564
2009/9	102,8	-1,9726	103,3495
2009/10	102,7	-1,3832	102,0384
2009/11	118,4	3,0348	109,5275
2009/12	185,4	20,1306	148,9812
2010/1	106,9	-0,2530	138,0059
2010/2	105,1	-5,7357	121,4264
2010/3	104,9	-5,7224	110,2953
2010/4	108,6	-3,3635	106,5864
2010/5	106,8	-2,3829	105,0114
2010/6	127,7	3,6829	115,1642
2010/7	113,3	0,6786	116,0736
2010/8	101,9	-3,0920	109,3261
2010/9	103,6	-2,8367	104,9170
2010/10	106,2	-1,4137	104,1401
2010/11	118,9	2,4704	110,8132
2010/12	186,2	19,6488	149,7418

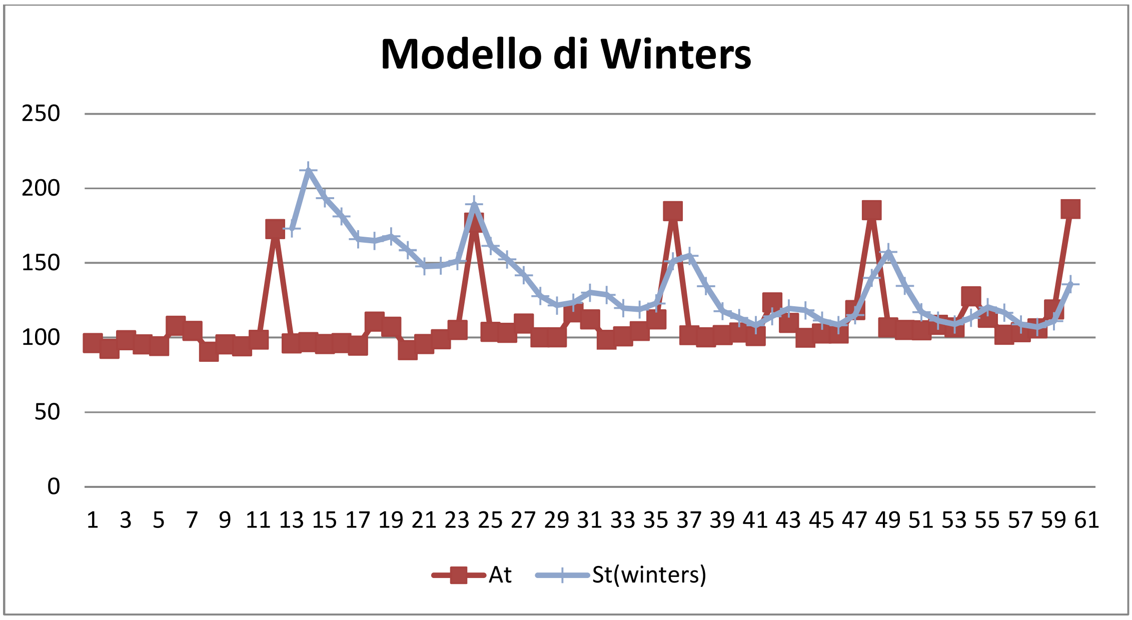
Successivamente ho determinato l'indice di stagionalità smorzato Q_t ipotizzando di avere una stagionalità mensile,cioè 12 periodi per ogni ciclo,la media smorzata e la quindi la previsione per il periodo F_{t+1}.

t	At	Qt	St(winters)	Tt(winters)	Ft+1
2006/1	96,3	0,5			
2006/2	92,4	0,324210526			
2006/3	98,2	0,341120278			
2006/4	95,5	0,361759553			
2006/5	94,3	0,395595176			
2006/6	107,9	0,463624889			
2006/7	104,6	0,460815155			
2006/8	90,6	0,435824104			
2006/9	95,4	0,481519892			
2006/10	94,2	0,489869545			
2006/11	98,6	0,508805004			
2006/12	172,7	0,639690364			
2007/1	96,1	0,634042061	173,1665039	96,1	
2007/2	97	0,599008423	212,1027374	67,518117	134,633252
2007/3	95,7	0,64421047	193,4469956	24,431187	90,65602429
2007/4	96,4	0,681755451	181,2691353	6,1266636	74,32266634
2007/5	94,7	0,703374678	165,9984186	-4,572027	67,79222045
2007/6	110,7	0,781056274	164,8102221	-2,880112	63,85950192
2007/7	107,3	0,740935532	167,8614139	0,0855401	75,0748296
2007/8	91,6	0,679051627	158,6074377	-4,584218	77,39250166
2007/9	95,7	0,737850377	147,7829928	-7,704331	67,12703172
2007/10	99	0,75709596	148,1965708	-3,645377	67,45066187
2007/11	105,1	0,776671363	151,3498905	-0,246029	70,81122769
2007/12	177	0,953976113	189,4092676	18,906674	76,88240108
2008/1	103,8	0,71174653	161,4395465	-4,531523	133,2577007
2008/2	103,3	0,737746465	152,4330213	-6,769024	99,48628633
2008/3	109,4	0,81233389	141,7907229	-8,705661	87,25396105
2008/4	100,2	0,820356521	127,8752191	-11,31058	85,73479005
2008/5	100,3	0,851416315	121,5035232	-8,841139	79,46857635
2008/6	117	0,937255448	123,3988975	-3,472882	79,24386806
2008/7	112,2	0,878064769	130,1358364	1,6320282	93,66896654
2008/8	98,6	0,807711164	128,601431	0,0488114	97,63149289
2008/9	100,8	0,863723207	119,7662765	-4,393172	87,36015641
2008/10	104,4	0,890164325	118,7772453	-2,691101	85,12808899
2008/11	112,3	0,914025279	122,9577813	0,7447173	87,88835069
2008/12	184,9	1,106902745	151,2260893	14,506513	96,07618822

2009/1	101,6	0,734254877	154,8297188	9,0550711	158,1049433
2009/2	100,3	0,79856682	134,5383564	-5,618146	116,6444305
2009/3	101,7	0,886074747	117,7052653	-11,22562	95,11042974
2009/4	103,5	0,91853809	113,1303016	-7,900291	86,49702579
2009/5	101	0,933220578	108,3178882	-6,356352	86,32612543
2009/6	123,7	1,030243448	114,2692302	-0,202505	86,81171527
2009/7	109,9	0,930308521	119,4823604	2,5053125	106,9096596
2009/8	99,9	0,87350274	118,2478538	0,6354029	107,1130779
2009/9	102,8	0,929202767	111,4593847	-3,076533	96,02333363
2009/10	102,7	0,94832367	108,3744614	-3,080728	93,61278417
2009/11	118,4	0,997515697	115,0960445	1,8204275	93,72872497
2009/12	185,4	1,175677416	140,0284228	13,376403	106,864611
2010/1	106,9	0,754429564	157,3508417	15,349411	169,8042227
2010/2	105,1	0,832055633	134,6818509	-3,65979	126,8060028
2010/3	104,9	0,918578598	117,0390317	-10,6513	104,6298706
2010/4	108,6	0,968714579	111,4021436	-8,144096	94,2674784
2010/5	106,8	0,975126147	108,832673	-5,356783	94,84644952
2010/6	127,7	1,069547091	113,2899262	-0,449765	96,56582942
2010/7	113,3	0,953206552	120,3174051	3,2888569	116,2528366
2010/8	101,9	0,90278803	116,7045435	-0,162002	114,9919587
2010/9	103,6	0,958324713	108,8637726	-4,001387	101,8002291
2010/10	106,2	0,984051533	107,0337111	-2,915724	97,43841922
2010/11	118,9	1,03524613	110,961291	0,5059279	98,73755152
2010/12	186,2	1,209575377	135,8302401	12,687439	111,1903005
Ft+1					174,6088807

La predizione per il mese di Gennaio 2011 è quindi:

$$F_{2011/1} = (S_{2010/12} + T_{2010/12}) * Q_{2010/1} = 174{,}608$$



4.4.Smoothing adattivo semplice

Con l'aiuto di formule di aggiornamento adattivo si può far dipendere dal tempo *t* i diversi parametri introdotti nelle sezioni precedenti. Dato un modello di smoothing semplice si può far dipendere il parametro α dal tempo *t* (in simboli α_t) e quindi ottenere un modello adattivo semplice con le successive formule di aggiornamento automatico:

$$S_t = \alpha_t A_t + (1 - \alpha_t) S_{t-1}$$

$$\alpha_t = \left| \frac{D_t}{G_t} \right|$$

$$D_t = \beta E_t + (1 - \beta) D_{t-1}$$

$$G_t = \beta |E_t| + (1 - \beta) G_{t-1}$$

Dall'analisi delle formule di aggiornamento si può dedurre che se il modello non è molto distorto il valore di α_t tende a 0,mentre nel caso opposto α_t tenderà a 1.La scelta dei parametri deve essere eseguita in modo da minimizzare le misure di dispersione.

La previsione per il periodo *t+1* è data da:

$$F_{t+1} = S_t$$

4.6.Smoothing a tendenza ridotta

In genere accade che la componente di tendenza diminuisce col trascorrere del tempo,perciò sono stati trattati modelli di smoothing con tendenza ridotta.

Si può prevedere una riduzione automatica della tendenza con l'uso del parametro φ e in simboli otteniamo:

$$S_t = \alpha A_t + (1-\alpha)(S_{t-1} + T_{t-1}) * \varphi$$
$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} * \varphi$$

Come sempre anche in questo caso i parametri α β e φ vanno selezionati per minimizzare le misure di dispersione.

La previsione si effettua per il periodo *t+m* e risulta:

$$F_{t+m} = S_t + T_t \sum_{i=1}^{m} \varphi^i$$

4.7.Scelta dei valori iniziali nei modelli di smoothing

Quando andiamo ad applicare un modello di smoothing esponenziale per una qualsiasi serie storica si pone il problema della scelta dei valori iniziali. Se supponiamo per esempio di applicare il modello di smoothing esponenziale semplice per il primo periodo nella serie smorzata si avrebbe:

$$S_1 = \alpha A_1 + (1-\alpha)S_0$$

e come si nota è necessario attribuire un valore al termine S_0.

Per semplicità nella trattazione del modello avevamo posto $S_1=A_1$.

Se consideriamo invece il modello con stagionalità e tendenza è ovvio che il ragionamento appena effettuato dovrà essere svolto anche per il valore iniziale di tendenza smorzata T_0 e per gli indici di stagionalità iniziali.

In generale è possibile dire che i valori iniziali della serie smorzata e di tendenza si possono determinare minimizzando le misure di dispersione o utilizzando tecniche di

backforecasting,cioè come indica Vercellis(1990, pag360) si calcolano i valori iniziali come risultato di previsioni ottenute applicando il modello alla serie storica considerata dal periodo t al periodo 1,ovvero dalle ultime osservazioni alle prime.
Per determinare i valori degli indici di stagionalità invece si deve adoperare una complessa pratica che impiega le osservazioni del primo ciclo della serie storica.

CAPITOLO 5
Modelli autoregressivi

1.Introduzione

I metodi autoregressivi nascono dall'identificazione delle analogie fra le diverse osservazioni di una serie storica nei differenti periodi temporali,mediante l'analisi dell'autocorrelazione tra le osservazioni divise da un intervallo costante di tempo.
Dato un certo scarto temporale s,si può definire una nuova serie storica B da:

$$B_t = A_{t-s}$$

con t>s
In pratica con questa operazione si ottiene una nuova serie per traslazione temporale della serie originaria. Nello studio e nell'analisi di un modello autoregressivo si deve considerare tutte le possibili relazioni tra le serie storiche.
Un'ipotesi fondamentale per lo sviluppo di modelli autoregressivi è che la serie storica originaria sia stazionaria ovvero che il suo andamento sia in equilibrio intorno ad una media costante.
Rispetto ai modelli di smoothing esponenziale i modelli autoregressivi sono più adattabili ai diversi casi in esame e hanno un carattere più generale. Ciò nonostante numerose prove empiriche hanno dimostrato che non sempre garantiscono una migliore e accurata previsione dei metodi più semplici.
Andiamo ora a mostrare i vari tipi di metodi autoregressivi.

2.Metodi autoregressivi

2.1.Modelli autoregressivi

Un modello autoregressivo(AR) di ordine *p* si scrive in simboli:

$$A_t = \gamma + \varphi_1 A_{t-1} + \cdots + \varphi_p A_{t-p} + \varepsilon_t$$

I parametri $\gamma, \varphi_1, \ldots, \varphi_p$ sono calcolati con il metodo dei minimi quadrati,così da minimizzare lo scarto quadratico medio. Questo metodo permette di ricavare la funzione che meglio approssima un insieme di dati,indicati con (x_i,y_i),ricavati sperimentalmente,che possono essere anche i valori di una serie storica. La funzione cercata sarà determinata minimizzando la distanza tra le due successioni y_i e $f(x_i)$, ovvero la quantità:

$$\sum_{i=1}^{n}(y_i - f(x_i)^2$$

da cui deriva il nome "*minimi quadrati*".

ε_t viene chiamata *rumore* ed è una variabile casuale che identifica la componente di fluttuazione casuale. In condizioni ideali deve seguire la distribuzione normale.
La previsione si può ricavare con :

$$F_{t+1} = \gamma + \varphi_1 A_t + \varphi_2 A_{t-1} \dots + \varphi_p A_{t-p+1}$$

2.2.Modelli a media mobile

Il modello a media mobile MA di ordine *q* in simboli si scrive:

$$A_t = \gamma - \vartheta_1 E_{t-1} - \vartheta_2 E_{t-2} - \cdots - \vartheta_q E_{t-q} + \varepsilon_t$$

I termini $E_{t-1}, E_{t-2}, \dots, E_{t-q}$ indicano gli errori di predizione nei q periodi precedenti.
Come in precedenza i parametri $\gamma, \vartheta_{1,} \vartheta_2 \dots, \vartheta_q$ sono calcolati il con il metodo dei minimi quadrati maniera tale da minimizzare lo scarto quadratico medio.
Anche in questo caso ε_t indica il rumore e dovrebbe seguire una distribuzione normale.
Si può ricavare quindi la previsione per il periodo t+1 da:

$$F_{t+1} = \gamma - \vartheta_1 E_t - \vartheta_2 E_{t-1} - \cdots - \vartheta_q E_{t-q+1}$$

2.3.Modelli autoregressivi a media mobile

Un modello autoregressivo a media mobile ARMA di ordine p e q si scrive in forma generale come:

$$A_t = \gamma + \varphi_1 A_{t-1} + \cdots + \varphi_p A_{t-p} - \vartheta_1 E_{t-1} - \vartheta_2 E_{t-2} - \cdots - \vartheta_q E_{t-q} + \varepsilon_t$$

Anche in questo caso i parametri della formula devono essere ricavati con il metodo dei minimi quadrati,così da minimizzare lo scarto quadratico.
La previsione per il periodo t+1 si determina da:

$$F_{t+1} = \gamma + \varphi_1 A_t + \cdots + \varphi_p A_{t-p+1} - \vartheta_1 E_t - \vartheta_2 E_t - \cdots - \vartheta_q E_{t-q+1}$$

<u>2.4.Modelli autoregressivi integrati a media mobile</u>

Per una serie storica non stazionaria si può utilizzare il modello ARMA alla nuova serie storica S_t ricavata attraverso *d* differenziazioni. Una serie storica non stazionaria può presentare comunque caratteristiche di stazionarie nelle differenze (A_t-A_{t-1}) e (A_{t-1}-A_{t-2}).
Il parametro *d* indica l'ordine delle differenziazioni che rende il processo pari ad un ARMA. Si ottiene in tal modo un modello autoregressivo integrato a media mobile ARIMA(p,d,q).
In formule si ha:

$$S_t = \gamma + \varepsilon_t + \varphi_1 S_{t-1} + \cdots + \varphi_p S_{t-p} - \vartheta_1 z_{t-1} - \cdots - \vartheta_q z_{t-q}$$

I vari fattori z indicano gli errori di predizione per la serie stazionaria S_t ,mentre anche in questo caso i parametri vengono determinati con il metodo dei minimi quadrati.
La previsione per il tempo t+1 si determina da:

$$S_t = \gamma + \varphi_1 S_t + \varphi_2 S_{t-1} + \cdots + \varphi_p S_{t-p+1} - \vartheta_1 z_t - \cdots - \vartheta_q z_{t-q+1}$$

CONCLUSIONI

In questo elaborato abbiamo visto le diverse tecniche di previsione riguardanti i modelli estrapolativi.
In conclusione credo sia importante osservare che è molto adoperata come tecnica previsionale anche la combinazione dei diversi metodi di previsione. In pratica si considera una somma pesata di più predizioni effettuate con i vari modelli estrapolativi,ottenendo un nuovo modello derivante dalla combinazione di modelli a media mobile con modelli autoregressivi e/o con modelli di smoothing esponenziale. Per i modelli di smoothing è possibile anche variare i valori attribuiti ai parametri del modello,in modo da ottenere numerose predizioni.
Assegnando dei pesi $z_1, z_2, \dots, z_h$ a *h* predizioni ottenute con *h* modelli $F_1, F_2, \dots, F_h$ si può considerare un nuovo modello predittivo F_c ricavato dalla combinazione degli precedenti modelli. In simboli si può scrivere F_c come

$$F_c = \frac{\sum_{i=1}^{h} z_i F_i}{\sum_{i=1}^{h} z_i}$$

Da numerosi test si è notato che il modello F_c fornisce previsioni migliori dei singoli modelli soprattutto per quanto riguarda la riduzione degli errori di dispersione.

Bibliografia e Sitografia

P.M.Morse - Ricerca Operativa e Pianificazione,1970,Marsilio Editori

C.Vercellis - Modelli e Decisioni,1997,Progetto Leonardo Bologna

S.Zani - Statistica,1991,Giuffrè Editore

E.Perucca - Dizionario di Ingegneria,1952,U.T.E.T.

M.Lenti - Enciclopedia dell'ingegneria,1971.Arnoldo Mondadori Editore

Slide delle Lezioni di Laboratorio di Sistemi Informativi Aziendali - Mario Guarracino (A.A. 2006/2007)

Slide di Statistica Economica, seminario n°6 - Rena to Guseo (A.A.2001/2002)

Web of Science

www.wikipedia.it

www.3ndy.biz