Eliminazione del Ciclo-Trend

Come si è visto il trend è individuato in base a due tecniche principali

- **Adattamento con i minimi quadrati (o altro criterio) di una specifica funzione che scaturisce dalla natura del fenomeno analizzato**

$$CT_t = f(t)$$

- **Adattamento con i minimi quadrati (o altro criterio) di un polinomio il cui grado È individuato con tentativi ed errori**

$$CT_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_p t^p$$

Il ciclo-Trend deve essere POI rimosso per poter studiare la stagionalità.

Eliminazione di un trend specifico

La serie mensile qui riportata ha un chiaro andamento logistico

$$y_t = \beta_0 + \frac{1}{\beta_1 + e^{-\beta_2 t}}$$

Tenuto conto che

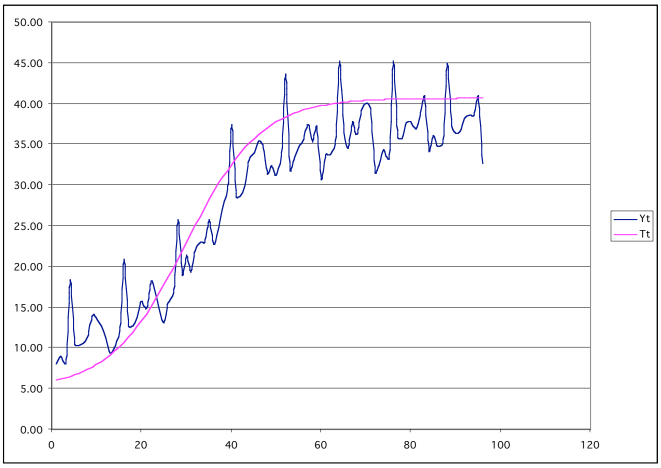
$$\lim_{t\to 0} T_t = \frac{1}{\beta_1 + 1} \cong 6, \quad \lim_{t\to\infty} T_t = \frac{1}{\beta_1} \cong 40$$

Si ha la stima approssimata

$$\frac{1}{\beta_1} - \frac{1}{\beta_1 + 1} = 34 \Rightarrow \beta_1 = 0.028, \beta_0 = 4.286$$

Da completare con

$$Ln\left(\frac{1}{y_t - 4.286} + 0.028\right) = \beta_2(-t) \Rightarrow \beta_2 = 0.031$$



Se il modello è di tipo additivo il trend Andrebbe sottratto alla serie

$$y_t - T_t = S_t + u_t$$

Se il modello è di tipo moltiplicativo il trend divide la serie

$$\frac{y_t}{T_t} = S_t * u_t$$

Differenziazione della serie storica

Un modo semplice di rimuovere il ciclo-trend polinomiale è di trasformare la serie originale in differenze prime, seconde, etc.

$$y_t = x_t - x_{t-1};$$

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = x_t - 2x_{t-1} + x_{t-2}$$

$$\Delta^p y_t = \sum_{j=0}^{p} (-1)^j \binom{p}{j} x_{t-j}$$

Se la serie fosse costituita dal solo polinomio di grado p, le differenze di ordine "p" sarebbero tutte costanti

Questo però è valido solo nel modello additivo. In quello moltiplicativo si procede come in precedenza

Differenziazione della serie storica/2

$$\Delta^p\left(\beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_p t^p\right) = costante$$

Infatti

$$x_t = \beta_0 + \beta_1 t + \beta_2 t^2$$

$$\begin{aligned}\Delta y_t = x_t - x_{t-1} &= \beta_0 + \beta_1 t + \beta_2 t^2 - \beta_0 - \beta_1(t-1) - \beta_2(t-1)^2 \\ &= \beta_0 + \beta_1 t + \beta_2 t^2 - \beta_0 - \beta_1 t + \beta_1 - \beta_2 t^2 - \beta_2 + 2\beta_2 t \\ &= (\beta_1 - \beta_2) + 2\beta_2 t\end{aligned}$$

$$\Delta y_{t-1} = x_{t-1} - x_{t-2} = -\beta_2 + 2\beta_2(t-1) = \beta_1 - 3\beta_2 + 2\beta_2 t$$

$$\begin{aligned}\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} &= -\beta_2 + 2\beta_2 t - \left[-3\beta_2 + 2\beta_2 t\right] \\ &= -\beta_2 + 2\beta_2 t + 3\beta_2 - 2\beta_2 t = 2\beta_2\end{aligned}$$

Differenziazione della serie storica/3

Per decidere quante volte differenziare una serie si parte dal fatto che

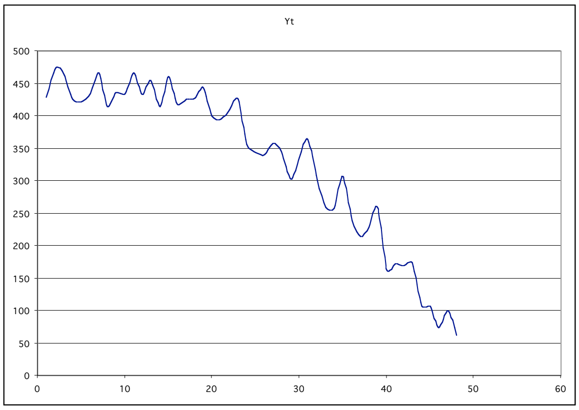
$$\Delta^{p+1}\left(\beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_p t^p\right) = 0$$

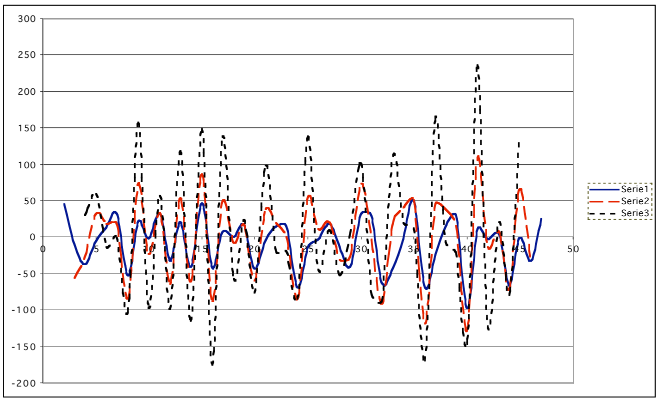
Se quindi le differenze di ordine "p+1" oscillano senza struttura intorno allo zero allora il grado del polinomio è "p"

Esistono anche dei test per una decisione inferenziale su "p", ma l'analisi grafica è di solito sufficiente per decidere.

In caso di dubbio scegliere il grado minore perché una differenziazione eccessiva può introdurre delle oscillazioni" nella serie storica facilmente confuse con componenti strutturali

Esempio





Il trend è quadratico. Le differenze prime sembrano già stabili, ma sono quasi tutte negative.

Le differenze terze oscillano paurosamente superando in ampiezza addirittura il livello medio della serie storica

Le differenze seconde sono più rispondenti ai canoni di stabilità, ma neanche loro convincono troppo

Differenze e logaritmi

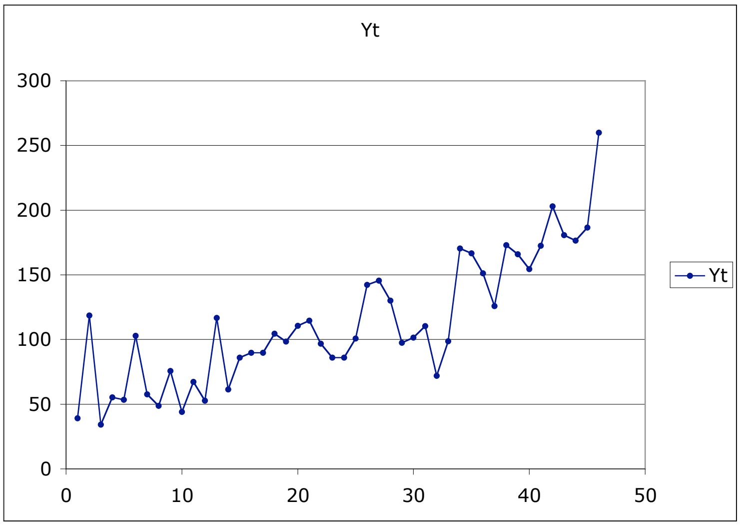
Se I valori sono tutti positivi e se la serie ha un trend esponenziale diventa opportuna la trasformazione logaritmica.

Questa può essere affiancata dalla differenziazione della serie storica

L'uso delle differenze logaritmiche è uno strumento molto efficace di stabilizzazione della serie storica

La serie riportata a destra mostra un evidente trend esponenziale oltreché una spiccata stagionalità.

La sola trasformazione logaritmica forse non basta a stabilizzarla



Stima del trend a mezzo di un filtro

Per stimare il trend abbiamo bisogno di accorpare i valori eventualmente rilevati per sottounità. Questo si può realizzare in modo efficiente a mezzo di filtri.

il filtro di una serie storica consiste nel trasformare una serie in un'altra pari alla prima, depurata di una parte delle componenti

il filtro da noi usato è LINEARE in quanto si concretizza in una combinazione lineare di valori della serie originaria

$$y_t = \sum_{i=-r}^{s} a_i x_{t+i}; \quad \text{dove} \quad \sum_{i=-r}^{s} a_i = 1$$

Ogni valore di "y" è ottenuto come MEDIA PONDERATA dei valori della "x. Poiché, per ogni "t", un termine entra ed un termine esce dalla media, si parla di MEDIA MOBILE.

Media aritmetica mobile

L'idea del filtro è di eliminare la stagionalità prima di misurare il ciclo-trend

La rimozione più semplicistica della stagionalità è di sommare tanti dati quanti ne indica il periodo della stagionalità e dividere per tale numero il totale trovato

A partire dalla prima osservazione utile si sostituiscono i valori osservati con le medie aritmetiche

$$Y_t = \left[\frac{X_t + X_{t-1} + \ldots + X_{t-(s-1)}}{s} \right]$$

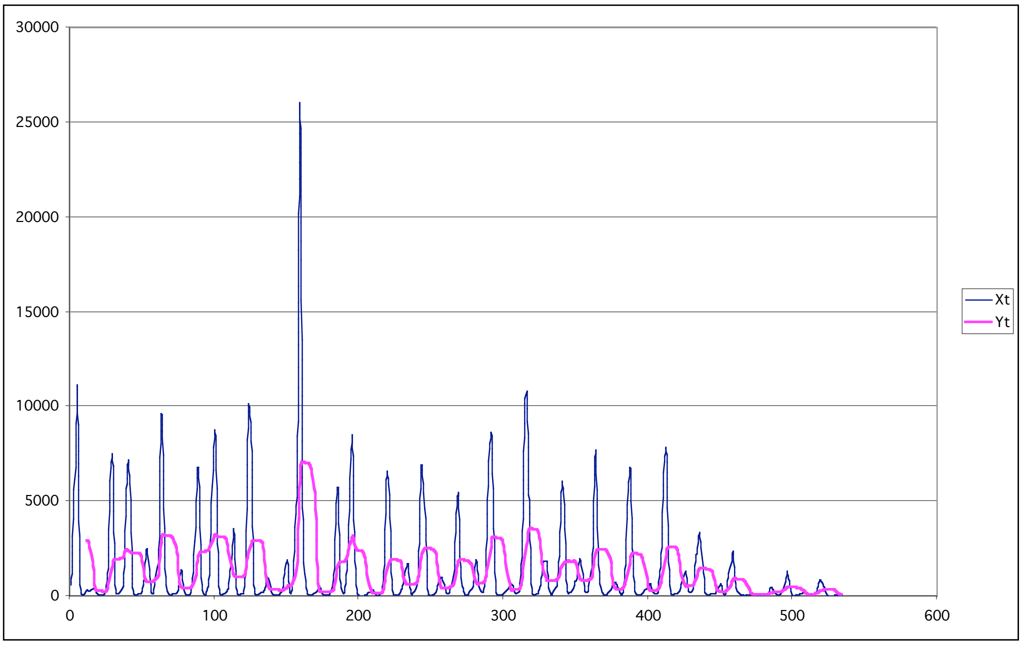
s= periodicità stagionale

Non appena matura un nuovo dato esso verrà passato nel filtro per produrre il dato Depurato.

Per un fissato periodo questa media ignora i valori successivi tenendo conto solo dei precedenti

Esempio

Reported monthly cases of measles, 1928-1972, New York City



L'andamento è smussato, ma la stagionalità non è del tutto sparita. Il filtro non considera i valori successivi

Da notare che il picco della serie filtrata è successivo a quello della serie osservata

Medie Mobili

Consideriamo ora la scelta: r=2, s=1

$$a_{-2} = \frac{2}{5};\ a_{-1} = \frac{1}{5};\ a_0 = \frac{1}{5};\ a_1 = \frac{1}{5}$$

t	x_t	$y_t = a_{-2}x_{t-2} + a_{-1}x_{t-1} + a_0x_t + a_1x_{t+1}$		
1	28			
2	34			
3	36	$y_3 = \frac{2}{5}x_1 + \frac{1}{5}x_2 + \frac{1}{5}x_3 + \frac{1}{5}x_4$	$= \frac{56+34+36+24}{5}$	$=30$
4	24	$y_4 = \frac{2}{5}x_2 + \frac{1}{5}x_3 + \frac{1}{5}x_4 + \frac{1}{5}x_5$	$= \frac{68+36+24+31}{5}$	$=31.8$
5	31	$y_5 = \frac{2}{5}x_3 + \frac{1}{5}x_4 + \frac{1}{5}x_5 + \frac{1}{5}x_6$	$= \frac{72+24+31+38}{5}$	$=33$
6	38			

Le medie si "muovono" lungo la serie storica

Esce X_1

Entra X_5

- **Il sistema dei pesi privilegia più i valori precedenti che quelli seguenti**
- **Il sistema dei pesi è asimmetrico: dati in posizione equidistante hanno pesi diversi**
- **La serie filtrata perde due informazioni all'inizio ed una alla fine. La perdita Dei dati finali è considerata più seria rispetto a quelli persi all'inizio**

Il sistema dei pesi

Di solito il sistema dei pesi (filtro) è SIMMETRICO

$$r = s \quad e \quad a_{-i} = a_i \quad \text{per ogni } i$$

L'esempio più semplice di filtro simmetrico è la media mobile aritmetica

$$\frac{1}{2r+1}\sum_{i=t-r}^{t+r} x_i$$

dove (2r+1) è il numero di termini sommati

<u>Esempio</u>

Per r=3 i pesi sono $a_{-3} = a_{-2} = a_{-1} = a_0 = a_1 = a_2 = a_3 = \frac{1}{7}$

Questa media mobile assegna lo stesso peso a ciascuno degli addendi che compongono il valore della nuova serie

La serie filtrata ha un numero di termini pari a

$$n' = n - 2r \begin{cases} n = numero \quad di \quad dati \\ m \quad numero \quad di \quad termini \quad nella \quad MM \end{cases}$$

Il sistema dei pesi/2

Un altro esempio classico è la scelta dei pesi come i termini dello sviluppo del binomio di Newton

$$\left(\frac{1}{2}+\frac{1}{2}\right)^{2r} \Rightarrow \sum_{i=0}^{2r}\binom{2r}{i}\left(\frac{1}{2}\right)^{2r} = \frac{(2r)!}{i!(2r-i)!}\left(\frac{1}{2}\right)^{2r}$$

Per r=1 si ha

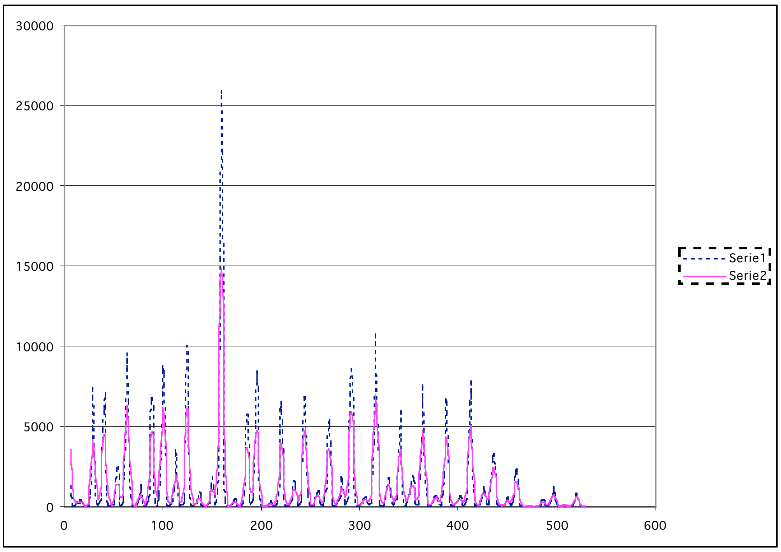
$$a_i = \frac{2!}{i!(2-i)!}\left(\frac{1}{2}\right)^{2} \Rightarrow a_0 = \frac{1}{4}; a_1 = \frac{1}{2}; a_2 = \frac{1}{4}$$

Per r=2 si ha

$$a_i = \frac{4!}{i!(4-i)!}\left(\frac{1}{2}\right)^{4} \Rightarrow a_0 = \frac{1}{16}; a_1 = \frac{4}{16}; a_2 = \frac{6}{16}; a_3 = \frac{4}{16}; a_4 = \frac{1}{16}$$

Questa scelta porta comunque ad un numero di pesi DISPARI simmetricamente disposti intorno ad un centro

Sistema di pesi binario con r=6



-6	0	0.00024414
-5	1	0.00292969
-4	2	0.01611328
-3	3	0.05371094
-2	4	0.12084961
-1	5	0.19335938
0	6	0.22558594
1	7	0.19335938
2	8	0.12084961
3	9	0.05371094
4	10	0.01611328
5	11	0.00292969
6	12	0.00024414
		1

La serie filtrata riproduce fedelmente l'andamento della serie originale.
Non si riscontra un vero e proprio effetto di filtro

Esercizio_5

Utilizzando il sistema di pesi ottenuto per r=1 applicare le medie mobili alla seguente serie per quadrimestri

t	X_t
1	112
2	103
3	108
4	117
5	103
6	119
7	106

Perché filtrare

La stima del Trend-Ciclo è finora avvenuta per dati annuali sommando i dati rilevati per frazioni di anno.

Se per ogni anno non si dispone ad esempio di tutti i mesi, o di tutti i semestri, l'unica possibilità è di usare le medie.

La stima del trend-ciclo avviene tuttavia su di un numero di rilevazioni che è al massimo pari al numero di anni per cui si estende la serie.

Questo si può migliorare

La componente stagionale, per definizione, si compensa nell'arco di 12 mesi o di 4 trimestri, o di 2 semestri etc. e la loro media non ne è più influenzata.

La stima del trend ciclo può essere meglio effettuata se si basa su delle medie mobili delle sottounità stagionali

Medie mobili per dati mensili

Per eliminare la stagionalità dei dati mensili si può usare la M.M. con 12 termini

$$y_t = \frac{1}{12}\sum_{i=-6}^{5} x_{t+i} = \frac{x_{t-6}+x_{t-5}+x_{t-4}+x_{t-3}+x_{t-2}+x_{t-1}+x_t+x_{t+1}+x_{t+2}+x_{t+3}+x_{t+4}+x_{t5}}{12}$$

La serie comincerebbe dal 7° periodo e si chiuderebbe al 5° termine precedente la fine dei dati. Si ottengono perciò n-11 valori utilizzabili

Essendo però pari il numero di termini la MM si colloca fra il 6° ed il 7° mese, diciamo a 6.5 mesi. La MM successiva sta tra il 7° e l'8° mese, cioè a 7.5 mesi.

Per ottenere una media CENTRATA sul 7° mese basterà effettuare la MA semplice delle due medie mobili (6.5+7.5)/2=7

La MM centrata è pari ad una non centrata con un numero di termini dispari e con i due pesi estremi uguali alla metà dei pesi centrali

$$\frac{y_t + y_{t-2}}{2} = \frac{x_{t-6} + 2\left(x_{t-5} + x_{t-4} + x_{t-3} + x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2} + x_{t+3} + x_{t+4} + x_{t+5}\right) + x_{t+6}}{24}$$

Medie mobili per dati mensili/2

L'uso delle M.M. mensili centrate fa però perdere 12 dati: 6 all'inizio e 6 alla fine. L'inconveniente non è troppo grave in fase di stima del trend.

Lo diventa in fase di 'estrapolazione dei dati futuri perché la stima si ferma ai 6 mesi precedenti il primo dato da estrapolare e risulta poco attuale

Esempio di calcolo

t	Xt	yt
1	102	
2	105	
3	118	
4	107	
5	113	
6	119	
7	105	111.38
8	111	113.08
9	113	114.46
10	119	116.00
11	101	117.46
12	114	117.75
13	121	118.54
14	127	
15	129	
16	133	
17	122	
18	117	
19	126	

$$111.38 = \frac{102+2[105+118+107+113+119+105+111+113+119+101+114]+121}{24}$$

$$y_{t+1} = y_t + \frac{x_{t+7} + x_{t+6} - x_{t-5} - x_{t-6}}{24}$$

$$y_8 = y_7 + \frac{x_{14} + x_{13} - x_2 - x_1}{24}$$

$$= 111.38 + \frac{127+121-105-102}{24} = 113.08$$

MM centrate per dati Tri-e quadri-mestrali

Nel caso di dati trimestrali, si può pensare a filtrare i dati con una media mobile a 5 termini

$$y_t = \frac{x_{t-2} + 2\left[x_{t-1} + x_t + x_{t+1}\right] + x_{t+2}}{8}$$

anche qui si tratta di una media mobile centrata, dato che i trimestri formano una suddivisione PARI dell'anno.

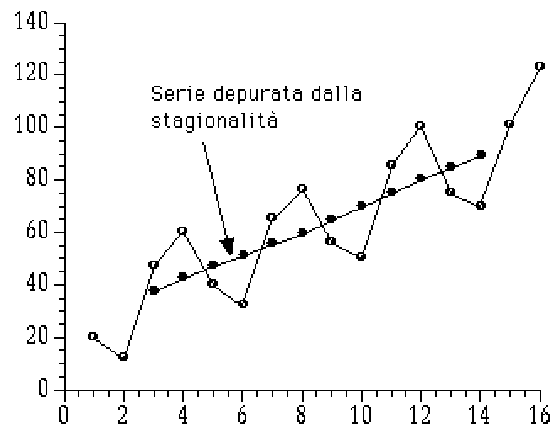
Per i dati quadrimestrali, essendo riferiti ad un frazionamento di anno che si articola con un numero di termini DISPARI, non è necessario CENTRARE le MM perché sono naturalmente riferite al dato intermedio

$$y_t = \frac{x_{t-1} + x_t + x_{t+1}}{3}$$

Esempio di MM

Andamento export di beni durevoli

Anno	Trim.	X_t	Tot.sempl.	MM centr.
1985	1	20		
	2	12	139.00	
	3	47	159.00	37.25
	4	60	179.00	42.25
1986	1	40	197.00	47.00
	2	32	213.00	51.25
	3	65	229.00	55.25
	4	76	247.00	59.50
1987	1	56	267.00	64.25
	2	50	291.00	69.75
	3	85	310.00	75.13
	4	100	330.00	80.00
1988	1	75	346.00	84.50
	2	70	369.00	89.38
	3	101		
	4	123		



Per il calcolo delle MM centrate è forse più semplice procedere in due passi:

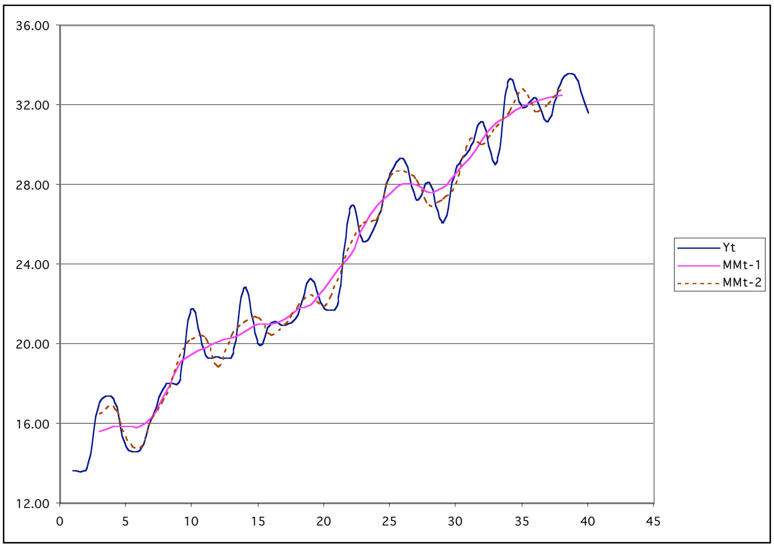
1) calcolare i totali mobili semplici: 139=20+12+47+60; 159=12+47+60+40, etc.

2) La MM centrata è data dalla somma di due totali mobili contigui divisi per 8

$$37.25 = \frac{139 + 159}{8}$$

Operatività delle MM

t	Yt	MMt-1	MMt-2
1	13.65		
2	13.80		
3	17.05	15.59	16.48
4	17.25	15.84	16.88
5	14.85	15.86	15.29
6	14.65	15.88	14.79
7	16.35	16.38	16.33
8	18.05	17.68	17.45
9	18.10	18.96	19.39
10	21.80	19.50	20.24
11	19.40	19.83	20.32
12	19.35	20.12	18.87
13	19.40	20.33	20.51
14	22.85	20.61	21.14
15	20.00	21.02	21.31
16	21.05	21.04	20.47
17	20.95	21.28	21.04
18	21.45	21.78	21.92
19	23.30	21.98	22.46
20	21.75	22.78	21.92
21	21.90	23.69	23.16
22	26.90	24.49	25.08
23	25.15	25.88	26.11
24	26.25	27.01	26.34
25	28.55	27.58	28.42
26	29.30	28.07	28.70
27	27.25	27.99	28.23
28	28.10	27.62	26.96
29	26.10	27.87	27.28
30	28.75	28.57	28.05
31	29.80	29.33	30.28
32	31.15	30.26	30.01
33	29.10	31.09	30.93
34	33.25	31.51	31.62
35	31.90	31.93	32.83
36	32.40	32.19	31.67
37	31.20	32.40	32.07
38	33.30	32.50	32.87
39	33.50		
40	31.60		



=(B45+2*(B46+B47+B48)+B49)/8

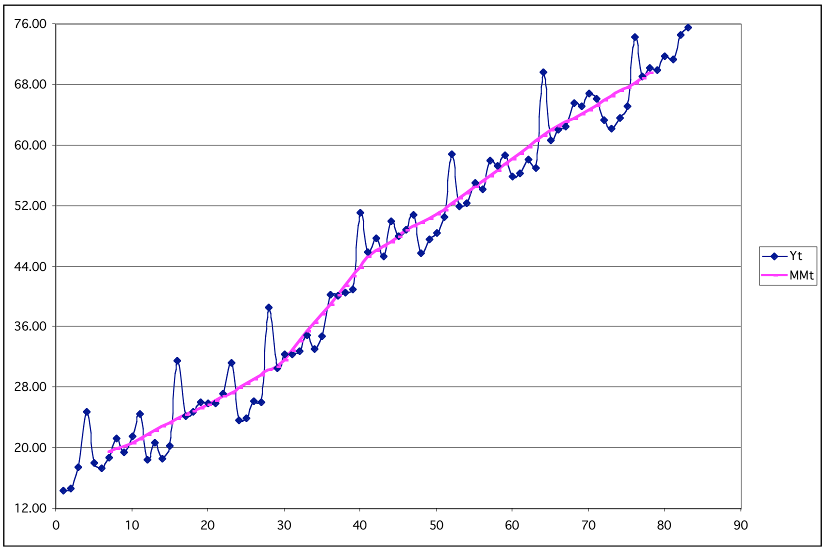
=(-3*B45+12*B46+17*B47+12*B48-3*B49)/35

L'uso combinato di calcoli e grafici consente la facile disamina della serie storica

Il foglio elettronico consente rapidi riscontri delle nostre ipotesi di scenario

Operatività delle MM/2

t	Yt	MMt	t	Yt	MMt
1	14.34		43	45.39	46.57
2	14.62		44	49.94	47.21
3	17.46		45	48.01	47.94
4	24.86		46	48.80	48.66
5	18.02		47	50.87	49.23
6	17.38		48	45.70	49.68
7	18.78	19.50	49	47.57	50.27
8	21.34	19.93	50	48.44	50.84
9	19.50	20.22	51	50.59	51.43
10	21.54	20.61	52	58.78	52.20
11	24.50	21.15	53	51.97	52.88
12	18.50	21.72	54	52.32	53.64
13	20.66	22.34	55	54.99	54.43
14	18.62	22.84	56	54.18	55.20
15	20.30	23.30	57	57.93	55.87
16	31.58	23.81	58	57.32	56.59
17	24.26	24.32	59	58.71	57.41
18	24.82	24.82	60	55.94	58.18
19	26.10	25.17	61	56.29	58.90
20	25.94	25.62	62	58.20	59.69
21	25.98	26.17	63	56.95	60.47
22	27.26	26.70	64	69.74	61.16
23	31.22	27.26	65	60.69	61.87
24	23.62	27.84	66	62.04	62.49
25	23.90	28.42	67	62.51	63.05
26	26.18	28.98	68	65.66	63.52
27	26.06	29.63	69	65.17	64.09
28	38.58	30.25	70	66.80	64.62
29	30.62	30.64	71	66.15	65.16
30	32.42	31.48	72	63.38	65.85
31	32.46	32.85	73	62.25	66.51
32	32.86	34.13	74	63.60	67.07
33	34.86	35.34	75	65.19	67.59
34	33.14	36.48	76	74.34	68.17
35	34.78	37.64	77	69.05	68.89
36	40.26	38.92	78	70.24	69.51
37	40.09	40.10	79	69.95	
38	40.56	41.35	80	71.78	
39	40.91	42.61	81	71.41	
40	51.10	43.81	82	74.56	
41	45.93	45.13	83	75.63	
42	47.76	46.03	84	68.94	



La componente stagionale può essere a volte tanto intensa da mascherare momenti evolutivi importanti

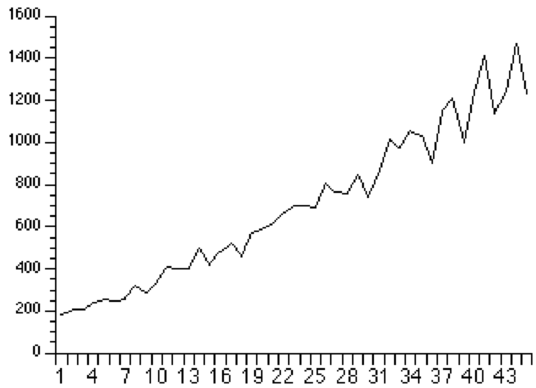
La media mobile centrata a 13 termini evidenzia il netto cambio del trend avvenuto tra la fine del 3° e l'inizio del 4° anno

Questa può essere un'ottima ragione per filtrare

Esercizio_5bis

Data la seguente serie quadrimestrale estesa su 15 anni

Qm	Serie	Qm	Serie	Qm	Serie
1	181.2	16	479.2	31	861.2
2	205.8	17	522.8	32	1016.8
3	205.8	18	456.8	33	973.8
4	236.2	19	569.2	34	1056.2
5	259.0	20	592.0	35	1029.0
6	246.2	21	617.2	36	901.2
7	255.8	22	669.8	37	1141.8
8	320.8	23	697.8	38	1209.8
9	291.2	24	697.2	39	1002.2
10	333.0	25	693.0	40	1233.0
11	409.2	26	808.2	41	1413.2
12	403.8	27	762.8	42	1139.8
13	400.8	28	752.8	43	1242.8
14	496.2	29	845.2	44	1464.2
15	417.0	30	743.0	45	1237.0



Filtrare la stagionalità adottando lo schema: $$Y_t = \frac{3X_{t-1} + 6X_t + 3X_{t+1}}{12}$$

Vantaggi delle MM

Le MM oltre ad essere di facile calcolo e interpretazione sono anche FLESSIBILI cioè seguono abbastanza fedelmente la tendenza evolutiva del fenomeno.

La stima del Trend può perciò avvenire direttamente su di esse e non sui dati grezzi spesso ricchi di asperità che inducono forti errori di fitting.

Il metodo di stima basato sulle medie mobili è molto usato nella pratica perché

- **Ha il pregio di basarsi su premesse logiche semplici, intuibili, accettabili**
- **E' applicato direttamente ai dati grezzi evitando di effettuare delle trasformazioni**
- **Eliminano la stagionalità con una perdita ridotta di informazioni**
- **Riducono l'erraticità grazie alle operazioni perequative**

L'effetto Slutzky-Yule

I dati usciti dal filtro delle MM, PER COSTRUZIONE sono fortemente auto correlati: il dato al tempo (t+1) è linearmente legato al dato del tempo t

Infatti, il calcolo delle MM avviene con uno schema ricorsivo del tipo

$$y_{t+1} = y_t + f(x_t)$$

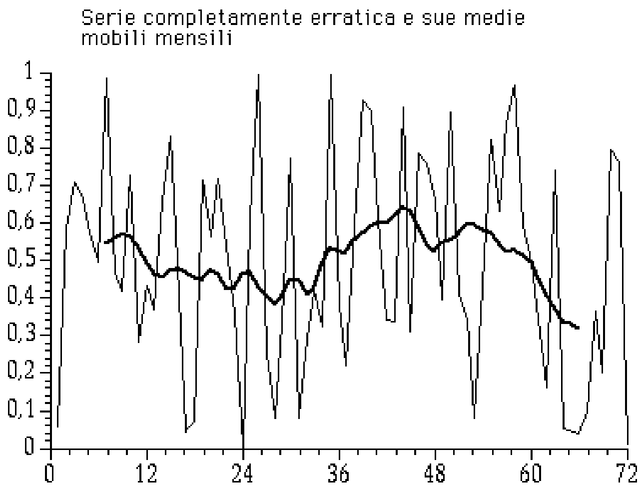
C'è quindi il rischio di introdurre delle strutture artificiali che possono oscurare le tendenze reali presenti nel fenomeno.

Un fenomeno molto noto nelle MM è l'EFFETTO SLUTZKY-YULE:

una serie del tutto erratica, se filtrata con le medie mobili può mostrare oscillazioni cicliche e stagionali molto simile a quelle delle serie reali con Componenti strutturali

Effetto Slutzky-Yule/2

t	e_t	MM_t	t	e_t	MM_t	t	e_t	MM_t
1	0.058		25	0.606	0.467	49	0.397	0.546
2	0.581		26	0.992	0.429	50	0.893	0.555
3	0.711		27	0.254	0.406	51	0.412	0.566
4	0.666		28	0.078	0.384	52	0.340	0.597
5	0.575		29	0.335	0.402	53	0.081	0.597
6	0.495		30	0.772	0.444	54	0.429	0.584
7	0.984	0.546	31	0.081	0.444	55	0.823	0.574
8	0.464	0.562	32	0.290	0.412	56	0.633	0.542
9	0.420	0.570	33	0.419	0.423	57	0.863	0.525
10	0.727	0.564	34	0.322	0.485	58	0.966	0.527
11	0.282	0.531	35	0.994	0.530	59	0.591	0.513
12	0.431	0.491	36	0.383	0.524	60	0.492	0.496
13	0.369	0.462	37	0.221	0.517	61	0.341	0.449
14	0.654	0.455	38	0.593	0.553	62	0.163	0.407
15	0.832	0.471	39	0.928	0.575	63	0.741	0.368
16	0.405	0.476	40	0.892	0.590	64	0.053	0.333
17	0.044	0.471	41	0.607	0.599	65	0.049	0.333
18	0.071	0.456	42	0.344	0.601	66	0.037	0.320
19	0.713	0.448	43	0.339	0.619	67	0.092	
20	0.562	0.472	44	0.908	0.639	68	0.363	
21	0.716	0.462	45	0.311	0.630	69	0.200	
22	0.537	0.425	46	0.788	0.586	70	0.795	
23	0.358	0.423	47	0.754	0.541	71	0.760	
24	0.000	0.464	48	0.660	0.328	72	0.012	



Una serie composta solo da valori casuali assume un aspetto "ragionevole" se filtrata con le MM. Che cosa ci impedisce di scambiare la serie in grassetto per una serie dotata di ciclo e trend?

Esempio sullo Slutzky-Yule

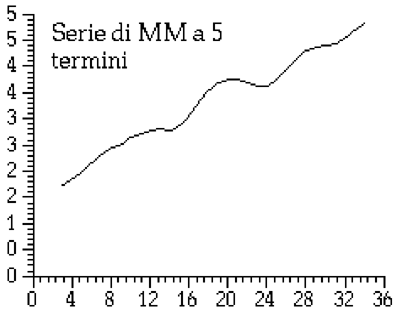
t	1+t/10	u_t	y_t	MMt
1	1.10	0.053	1.15	
2	1.20	0.654	1.85	
3	1.30	0.162	1.46	1.74
4	1.40	0.723	2.12	1.86
5	1.50	0.373	1.87	2.01
6	1.60	0.495	2.09	2.16
7	1.70	0.701	2.40	2.31
8	1.80	0.610	2.41	2.46
9	1.90	0.891	2.79	2.53
10	2.00	0.384	2.38	2.66
11	2.10	0.607	2.71	2.73
12	2.20	0.893	3.09	2.77
13	2.30	0.412	2.71	2.80
14	2.40	0.340	2.74	2.77
15	2.50	0.081	2.58	2.87
16	2.60	0.429	3.03	3.05
17	2.70	0.823	3.52	3.29
18	2.80	0.633	3.43	3.55
19	2.90	0.863	3.76	3.69
20	3.00	0.966	3.97	3.75
21	3.10	0.591	3.69	3.76
22	3.20	0.492	3.69	3.70
23	3.30	0.341	3.64	3.63
24	3.40	0.163	3.56	3.64
25	3.50	0.077	3.58	3.76
26	3.60	0.238	3.84	3.95
27	3.70	0.797	4.50	4.13
28	3.80	0.398	4.20	4.31
29	3.90	0.489	4.39	4.38
30	4.00	0.469	4.47	4.39
31	4.10	0.365	4.46	4.45
32	4.20	0.047	4.25	4.58
33	4.30	0.527	4.83	4.71
34	4.40	0.705	5.11	4.82
35	4.50	0.353	4.85	
36	4.60	0.175	4.78	

Consideriamo il modello per dati annuali

$$y_t = 1 + 0.1t + u_t;$$

con u_t errori pseudocasuali uniformi (0,1)

Non c'è alcun effetto ciclico, ma così non è se si guarda alle medie mobili in questo caso a 5 termini

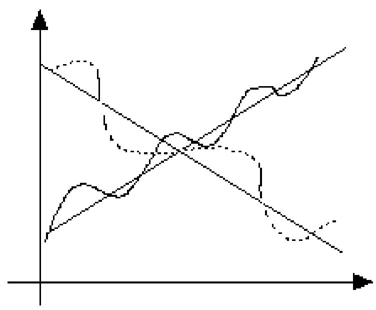


L'effetto si concreta nella somiglianza a serie con componenti cicliche di serie invece prive di ciclicità

L'uso delle M.M è però ancora ragionevole se la componente erratica è minoritaria

La stagionalità

Per STAGIONALITA' di una serie storica si intende una componente oscillatoria che fa ripetere alla serie un comportamento analogo nello stesso periodo (mese, trimestre, semestre, etc.) di anni successivi



Nell'ambito del modello moltiplicativo

$$y_t = T_t * C_t * S_t * u_t$$

La stagionalità si presenta come un indice di variazione rispetto ad un livello base (medio) unitario. Se cioè, per ogni "t" si ha

$$S_t = 1$$

non c'è stagionalità. Gli scarti dall'unità ne indicano invece la presenza.

<u>Esempio</u> $S_t = 0.95$ Riduzione del 5% rispetto al livello base

$S_t = 1.07$ Aumento del 7% rispetto al livello base

La stagionalità/2

Nello studio delle serie storiche è essenziale individuare e poi eliminare gli effetti della stagionalità. Solo così è possibile confrontare dati relativi periodi diversi

Esempio

Il numero destagionalizzato di disoccupati nel primo trimestre di un anno è comparabile con il dato corrispondente all'ultimo trimestre dell'anno precedente.

Questo non è possibile con i dati grezzi perché lo scarto potrebbe derivare da effetti tipici del 1° trimestre ad esempio: Pasqua bassa

La depurazione della stagionalità è una operazione complessa e non univoca: a seconda delle procedure adottate si perviene a dati DESTAGIONALIZZATI diversi e, a volte, molto diversi.

Inevitabili quindi le polemiche quando si tenti di valutare l'evoluzione corrente degli indicatori economici

Stima della stagionalità

La stima della stagionalità deve è da due operazioni

- **Effettuare le correzioni di calendario per eliminare eventuali difformità dovute a soli motivi tecnici**

- **Accertarsi della presenza di stagionalità perché altrimenti le tecniche di depurazione impiegate portano a risultati imprevedibili (effetto S-Y)**

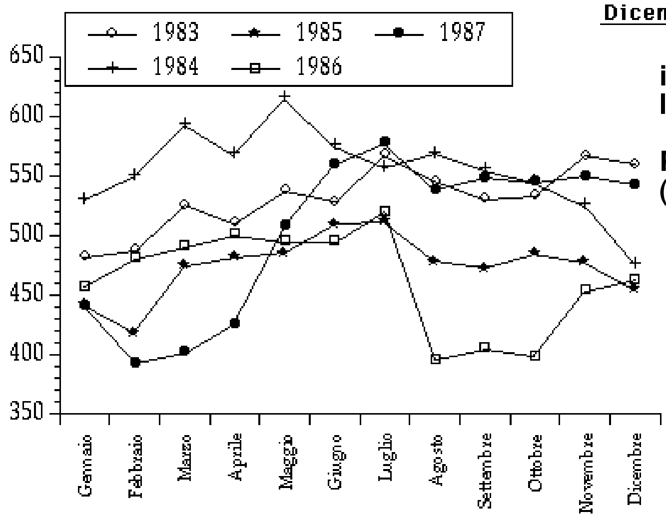
La verifica della presenza significativa della componente stagionale può essere effettuata con tecniche più o meno raffinate. Noi preferiamo procedere per intuizione.

L'idea è che se la serie NON contiene stagionalità allora i grafici redatti periodo per periodo dovrebbero essere tutti diversi ovvero non presentare somiglianze

Verifica della stagionalità

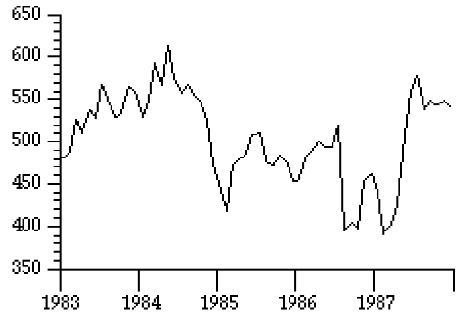
Mese	1983	1984	1985	1986	1987
Gennaio	481	529	441	456	440
Febbraio	487	549	418	480	392
Marzo	525	592	474	490	401
Aprile	510	567	481	500	425
Maggio	537	614	485	494	507
Giugno	528	574	509	494	559
Luglio	567	556	511	519	577
Agosto	545	568	478	395	538
Settembre	530	554	472	404	548
Ottobre	533	544	484	398	544
Novembre	566	524	477	453	549
Dicembre	559	474	454	462	542

Consideriamo i seguenti dati annuali relativi al prezzo medio mensile del cotone sui mercati internazionali



in presenza di stagionalità il grafico ha lo stesso aspetto anno per anno.

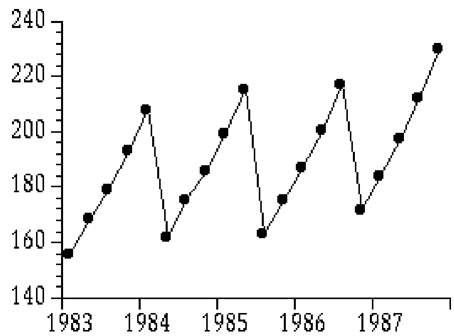
Per i dati in esempio questo è vero (1987 escluso)

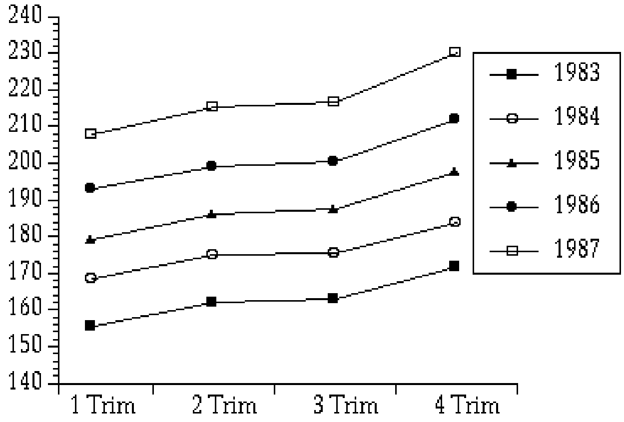


Verifica della stagionalità/2

Dati trimestrali sull'importo complessivo della moneta circolante

Anno	1 Trim	2 Trim	3 Trim	4. Trim
1983	155.7	162.0	162.8	171.9
1984	168.7	175.1	175.3	183.8
1985	179.2	185.9	187.3	197.5
1986	193.2	199.3	200.6	212.0
1987	207.8	215.2	216.8	230.2





In questo caso la presenza della stagionalità è più evidente: si riconosce la stessa struttura all'interno di ogni anno

Esercizio_6

1) Verificate graficamente la presenza di stagionalità nei seguenti dati rilevati quadrimestralmente

Anno	1 Quad	2 Quad	3 Quad
1981	5.5	5.1	6.5
1982	6.0	6.4	5.9
1983	6.2	3.5	4.3
1984	6.4	4.5	3.6
1985	4.0	5.0	6.4
1986	3.2	4.9	3.7
1987	6.0	5.3	3.2

2) Quali possono essere le conseguenze di considerare stagionale una serie storica che non lo sia effettivamente?

Verifica numerica della stagionalità

L'analisi grafica è utile quando i dati producono indicazioni univoche. In caso di interpretazione difficile si ricorre a confronti numerici.

Esistono delle procedure statistiche che sottopongono a verifica l'ipotesi di presenza/assenza di stagionalità, (ad esempio l'analisi della varianza), ma, al solito, si basano su presupposti non sempre validi per le serie storiche.

La nostra verifica numerica partirà invece da considerazioni pratiche.

A questo fine distingueremo fra due scelte possibili

- **Stagionalità costante: per ogni anno agiscono sempre gli stessi stimoli**

- **Stagionalità variabile: ogni anno le forze della stagionalità si esplicano in modo diverso (non troppo diverso, ma diverso)**

Le due scelte si basano entrambe sui rapporti di stagionalità

I rapporti di stagionalità

La serie filtrata con una MM collegata alla stagionalità depura la serie dalle componenti stagionali. Quindi

Componente ciclo-trend

$$\tilde{y}_t = \sum_{i=-r}^{r} a_i y_{t+i} \cong T_t * C_t$$

D'altra parte, in base al modello moltiplicativo si ha

$$y_t = T_t * C_t * S_t * u_t$$

Per cui il rapporto

$$rs_t = \frac{y_t}{\tilde{y}_t} = \frac{T_t * C_t * S_t * u_t}{T_t * C_t} = S_t * u_t$$

Questi sono i rapporti grezzi di stagionalità

può misurare la stagionalità purché l'incidenza degli errori sia piccola (cioè siano tutti vicino alla media unitaria)

Esempio sui rapporti di stagionalità

Consideriamo la serie trimestrale dei valori delle giacenze presso una impresa manifatturiera

anno	Tr.	Riman.	TM	MM	RS
1987	1	15.2			
	2	18.4			
	3	18.7	74.5	18.95	0.9868
	4	22.2	77.1	19.71	1.1262
1988	1	17.8	80.6	20.55	0.8662
	2	21.9	83.8	21.38	1.0246
	3	21.9	87.2	22.03	0.9943
	4	25.6	89.0	22.28	1.1493
1989	1	19.6	89.2	22.24	0.8814
	2	22.1	88.7	22.03	1.0034
	3	21.4	87.5	21.88	0.9783
	4	24.4	87.5	22.04	1.1072
1990	1	19.6	88.8	22.54	0.8697
	2	23.4	91.5	23.35	1.0021
	3	24.1	95.3		
	4	28.2			

Nel modello moltiplicativo le componenti non Trend sono espresse in % quindi RS=0.9868 significa che la stagionalità per il 3° trim '87 riduce il trend dello (1-0.9868)%= 1.32%.

In assenza di stagionalità gli "rs" sarebbero tutti uguali ad uno (a meno dell'influenza di errori).

Ogni scarto da "1", se consistente, ci convince della presenza di stagionalità

Nell'esempio ci sono "rs" con incidenza del ± 14% e quindi l'effetto della stagionalità è rilevante.

Esercizio_7

La seguente serie è relativa ad un fenomeno rilevato quadrimestralmente:

Anno	Q	Serie	Anno	Q	Serie	Anno	Q	Serie
1970	1	145.1	1978	1	174.5	1986	1	330.1
	2	140.4		2	191.6		2	349.0
	3	149.9		3	173.9		3	319.1
1971	1	125.6	1979	1	189.4	1987	1	347.4
	2	146.5		2	208.1		2	367.9
	3	138.6		3	180.0		3	334.6
1972	1	133.9	1980	1	194.1	1988	1	358.5
	2	142.4		2	222.4		2	402.6
	3	146.1		3	190.9		3	370.9
1973	1	134.0	1981	1	199.6	1989	1	392.4
	2	136.1		2	238.5		2	428.1
	3	139.4		3	203.6		3	397.0
1974	1	128.9	1982	1	228.9	1990	1	424.1
	2	139.6		2	256.4		2	445.4
	3	140.5		3	214.1		3	410.9
1975	1	153.6	1983	1	255.0	1991	1	451.6
	2	157.9		2	272.1		2	477.5
	3	158.4		3	257.4		3	455.6
1976	1	151.1	1984	1	270.9			
	2	164.0		2	307.6			
	3	158.1		3	270.5			
1977	1	167.4	1985	1	298.6			
	2	184.9		2	330.9			
	3	166.6		3	296.4			

a) Calcolare i rapporti di stagionalità.

b) Valutare la stagionalità in base all'incidenza dei rapporti.

Stagionalità costante

La STAGIONALITA' COSTANTE implica che gli "rs" siano uguali anno per anno. Per controllare questo aspetto conviene organizzarli nella tabella seguente

	Trimestri			
Anno	1	2	3	4
1987			0.9868	1.1262
1988	0.8662	1.0246	0.9943	1.1493
1989	0.8814	1.0034	0.9783	1.1072
1990	0.8697	1.0021		
Media	0.8724	1.0100	0.9865	1.1276

Da notare che il numero di coefficienti "rs" è sempre pari a: numero di anni - 1

La stagionalità non è mai regolare. La media aritmetica degli rs è una buona scelta che attenua l'erraticità.

Ecco delle alternative

- **Mediana dei pesi sui vari anni**
- **Media potata: si eclude l'rs più piccolo, quello più grande e si fa la media dei rimanenti**
- **Media perequata: si sostituiscono gli rs "anomali" con l'rs medio e si ricalcola la media**

Gli indici di stagionalità (grezzi e netti)

Con i trimestri le scelte si riducono alle prime due e siccome non si notano scompensi eccessivi tra i pesi, effettuiamo la prima scelta ottenendo i

Coefficienti grezzi (o lordi) di stagionalità	1 Trim	2 Trim	3 Trim	4 Trim	Tot.
	0.8724	1.0100	0.9865	1.1276	3.9965

Il totale degli "rs" dovrebbe essere pari a "4" (cioè al numero di frazioni annue)

Occorre effettuare delle correzioni perché altrimenti non si ha il compenso delle variazioni stagionali nel corso dell'anno.

Questo si ottiene moltiplicando i coefficienti grezzi per il fattore

$$\frac{4}{3.9965} = 1.00087577$$

in questo modo si ottengono i coefficienti corretti o netti

INDICI di Stagionalità	1 Trim	2 Trim	3 Trim	4 Trim	Tot.
	0.8732	1.0109	0.9873	1.1286	4.0000

Esercizio_8

A partire dalle seguenti rilevazioni semestrali

Anno	Semestre	Num.Ind.
1980	1	105
	2	110
1981	1	112
	2	118
1982	1	115
	2	122
1983	1	110
	2	128

Calcolare i rapporti di stagionalità

NB. Essendo la frazione d'anno PARI è necessario adoperare una media mobile con un numero DISPARI di termini:

$$MM_t = \frac{x_{t-1} + 2 * x_t + x_{t+1}}{4}$$

Significato degli indici di stagionalità

Il rapporto tra serie originaria ed MM elimina la componente ciclo-trend.

$$\frac{(\cancel{T_t*C_t})*S_t*u_t}{\cancel{T_t*C_t}} \cong S_t u_t$$

Con la media dei coefficienti di si rimuove (in parte) la componente erratica nonché gli eventi sporadici ed eccezionali (non legati ai megatrend)

Quindi, gli indici di stagionalità misurano la componente stagionale S

<u>Esempio</u>

Se, per dei dati quadrimestrali, si ottengono gli indici di stagionalità

Q1	Q2	Q3
0.951	1.062	0.987

Nell'approccio classico tale struttura si ripete da un anno all'alro sempre uguale anche nei numeri

si vede che nel primo quadrimestre c'è una deviazione dal ciclo-trend del 4.9% dovuta ad effetti stagionali; nel secondo quadrimestre c'è un aumento del 6.2% e nel terzo una variazione in meno del 2.3%

Esempio

Serie trimestrale PIL USA

La stagionalità è poco accentuata

Anno	Trim	t	Yt	MMt	
1976	1	1	398.9		
	2	2	426.5		
	3	3	430.8	434.25	0.9921
	4	4	461.8	445.26	1.0371
1977	1	5	436.9	458.78	0.9523
	2	6	476.6	472.75	1.0081
	3	7	488.8	485.45	1.0069
	4	8	515.6	498.85	1.0336
1978	1	9	484.7	513.63	0.9437
	2	10	536.0	530.01	1.0113
	3	11	547.6	548.44	0.9985
	4	12	587.9	565.64	1.0394
1979	1	13	559.8	580.89	0.9637
	2	14	598.5	596.06	1.0041
	3	15	607.1	610.84	0.9939
	4	16	649.8	624.16	1.0411
1980	1	17	616.1	637.19	0.9669
	2	18	648.8	650.91	0.9968
	3	19	661.0	667.11	0.9908
	4	20	705.7	686.94	1.0273
1981	1	21	689.8	709.03	0.9729
	2	22	733.7	729.50	1.0058
	3	23	752.8	743.98	1.0119
	4	24	777.7	754.41	1.0309
1982	1	25	733.6	761.84	0.9629
	2	26	773.4	766.28	1.0093
	3	27	772.5		
	4	28	793.5		

Trimestri				
1°	2°	3°	4°	
		0.9921	1.0371	
0.9523	1.0081	1.0069	1.0336	
0.9437	1.0113	0.9985	1.0394	
0.9637	1.0041	0.9939	1.0411	
0.9669	0.9968	0.9908	1.0273	
0.9729	1.0058	1.0119	1.0309	
0.9629	1.0093			
0.9604	1.0059	0.9990	1.0349	4.0002
0.9604	1.0058	0.9990	1.0348	4.0000

Pesi media mobile a 5 termini

Weig1= 0.125
Weig2= 0.250
Weig3= 0.250
Weig4= 0.250
Weig5= 0.125
1.000

Algoritmo per il calcolo degli indici di stagionalità

L'algoritmo di calcolo degli indici di stagionalità prevede i seguenti passi:

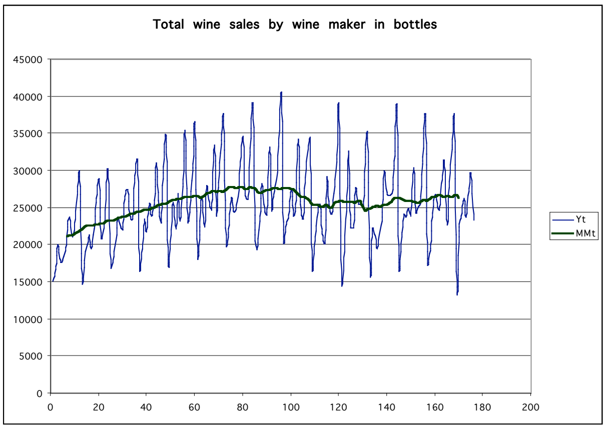
1 **A partire dai dati rilevati su "k" frazioni di anno si calcolano le MM con un numero di termini m=k se k è DISPARI oppure m=k+1 se k è PARI**

2 **Si calcolano i rapporti di stagionalità dividendo i valori originari per i valori filtrati con le MM (ovviamente perdendo k-1 rilevazioni).**

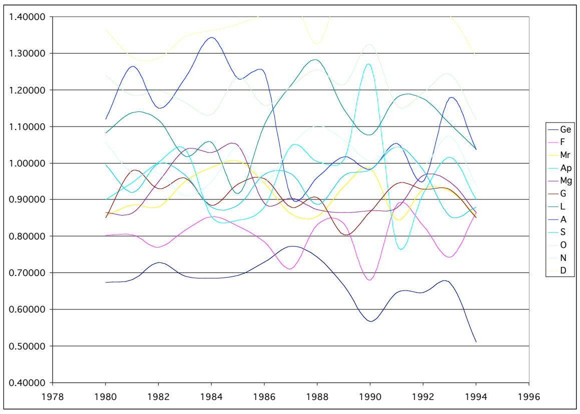
3 **Si sceglie un rapporto di stagionalità rappresentativo per ciascuna delle frazioni di anno (media, mediana, media potata, media perequata, etc.)**

4 **Si normalizzano i rapporti del punto "3" facendo in modo che la loro somma sia uguale a "k"**

Esempio



Coefficienti grezzi di stagionalità



Anno	Ge	F	M r	Ap	Mg	G	L	A	S	O	N	D	
1980	*0.67448*	*0.80316*	*0.85690*	*0.90050*	*0.86616*	*0.85118*	1.08298	1.12049	0.99507	1.05618	1.23993	1.36514	
1981	0.68223	0.80399	0.88678	0.94733	0.86456	0.97958	1.13861	1.26501	0.92058	0.97254	1.18599	1.28983	
1982	0.72841	0.77044	0.88133	1.00389	0.95011	0.93099	1.11778	1.15027	1.00006	0.97202	1.19878	1.28678	
1983	0.69189	0.81740	0.95075	1.03698	1.03698	0.95969	1.01948	1.23376	0.96729	0.91068	1.16533	1.34701	
1984	0.68628	0.85364	0.98883	0.85370	1.02950	0.88493	1.05732	1.34369	0.87950	0.95067	1.13373	1.36185	
1985	0.69405	0.82817	1.00598	0.84489	1.04724	0.94318	0.91800	1.23104	0.88638	1.04543	1.23482	1.38054	
1986	0.73034	0.78531	0.94450	0.88086	0.88983	0.95783	1.10628	1.24648	0.97017	0.94981	1.15771	1.40879	
1987	0.77339	0.71168	0.86160	1.04728	0.90404	0.87959	1.21395	0.90712	0.97139	1.05030	1.21028	1.46162	
1988	0.74307	0.83344	0.85503	1.00501	0.87331	0.90541	1.28177	0.96165	0.88788	1.10274	1.25539	1.32535	
1989	0.66527	0.83373	0.93663	1.01055	0.86571	0.80424	1.14593	1.01717	0.96694	1.06306	1.21548	1.49770	
1990	0.56812	0.68089	0.98598	1.26691	0.87066	0.86857	1.07779	0.98447	0.98276	1.00580	1.32500	1.42568	
1991	0.64757	0.88791	0.84632	0.77650	0.87845	0.94544	1.18059	1.05415	1.04438	1.04531	1.15733	1.47486	
1992	0.64649	0.82888	0.92927	0.91716	0.96659	0.92927	1.17583	0.95156	0.98079	0.99484	1.19305	1.43460	
1993	0.67314	0.74323	0.92359	1.01596	0.94933	0.92773	1.10772	1.17950	0.85473	1.07397	1.24211	1.40321	
1994	0.51248	0.86554	*0.85690*	*0.90050*	*0.86616*	*0.85118*	*1.03865*	*1.03756*	*0.87949*	*0.93837*	*1.11964*	*1.29270*	
C.g.s.	0.67448	0.80316	0.85690	0.90050	0.86616	0.85118	1.03865	1.03756	0.87949	0.93837	1.11964	1.29270	11.25879
C.n.s	0.77879	0.92737	0.98942	1.03977	1.00012	0.98281	1.19928	1.19802	1.01550	1.08349	1.29280	1.49262	13.00000

Destagionalizzazione

Tale operazione consiste nel determinare i valori della serie storica depurati dalle variazioni stagionali cioè i valori che si sarebbero ottenuti, in ogni frazione di anno, qualora la stagionalità fosse poco rilevante.

Per attuare questa ripulitura è necessario dividere i valori originari della serie per gli indici di stagionalità

$$D_t = \frac{y_t}{S_t} = \frac{(T_t * C_t) * \cancel{S_t} * u_t}{\cancel{S_t}} = (T_t * C_t) * u_t$$

La serie destagionalizzata è preferibile ai dati grezzi in quanto sono state eliminate le irregolarità più tormentate e che più complicano la stima del ciclo-trend

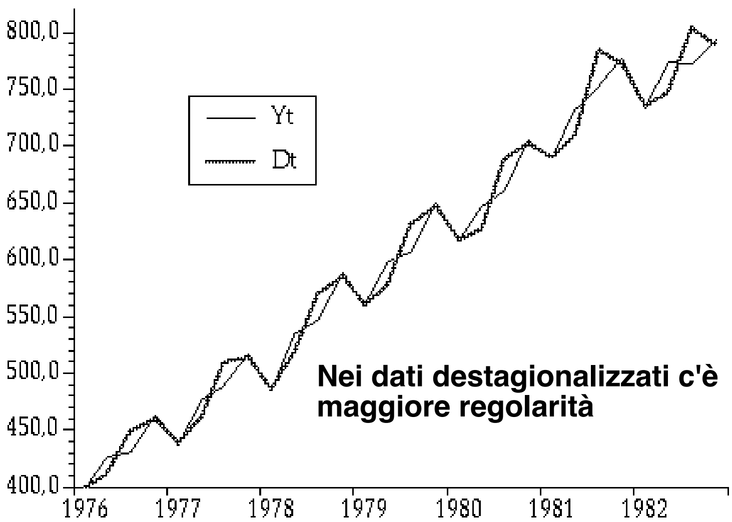
Esistono diverse tecniche di destagionalizzazione in quanto sono diversi i modi di misurare la stagionalità

Esempio di destagionalizzazione

Applichiamo la procedura di destagionalizzazione ai dati sul PIL USA di cui abbiamo già calcolato i coefficienti di stagionalità.

Anno	Trim	y_t	D_t
1976	1	398.9	399.3
	2	426.5	412.2
	3	430.8	448.6
	4	461.8	459.1
1977	1	436.9	437.3
	2	476.6	460.6
	3	488.8	509.0
	4	515.6	512.6
1978	1	484.7	485.2
	2	536.0	518.0
	3	547.6	570.2
	4	587.9	584.5
1979	1	559.8	560.4
	2	598.5	578.4
	3	607.1	632.1
	4	649.8	646.1
1980	1	616.1	616.7
	2	648.8	627.0
	3	661.0	688.3
	4	705.7	701.6
1981	1	689.8	690.5
	2	733.7	709.0
	3	752.8	783.8
	4	777.7	773.2
1982	1	733.6	734.3
	2	773.4	747.4
	3	772.5	804.4
	4	793.5	788.9

Tutti i dati del primo trimestre sono divisi per 0.999, quelli del secondo per 1.0348, etc.

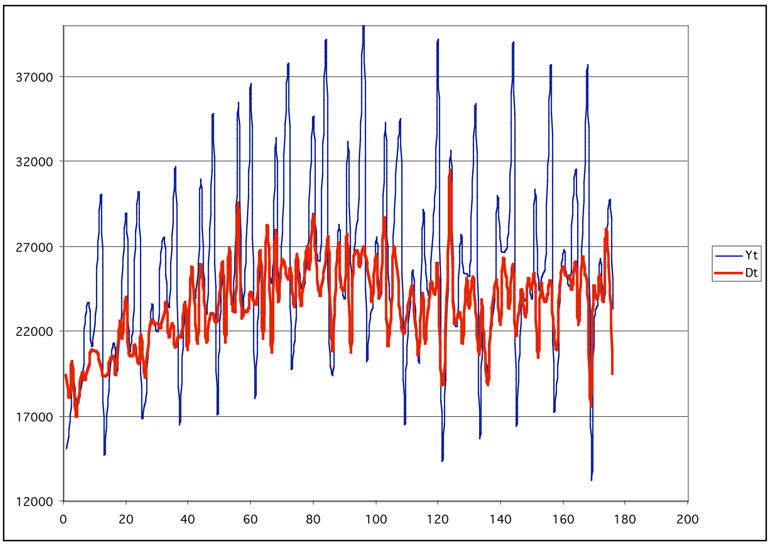


Nei dati destagionalizzati c'è maggiore regolarità

Destagionalizzazione della australian wine

t	Yt	c.n.s.
1	15136	0.77879108
2	16733	0.92737389
3	20016	0.9894212
4	17708	1.03976651
5	18019	1.00011992
6	19227	0.98281211
7	22893	1.19927758
8	23739	1.19802414
9	21133	1.01550396
10	22591	1.0834889
11	26786	1.2927991
12	29740	1.4926216
13	15028	0.77879108
14	17977	0.92737389
15	20008	0.9894212
16	21354	1.03976651
17	19498	1.00011992
18	22125	0.98281211
19	25817	1.19927758
20	28779	1.19802414
21	20960	1.01550396
22	22254	1.0834889
23	27392	1.2927991
24	29945	1.4926216
25	16933	0.77879108
26	17892	0.92737389
27	20533	0.9894212
28	23569	1.03976651
29	22417	1.00011992
30	22084	0.98281211
31	26580	1.19927758
32	27454	1.19802414
33	24081	1.01550396
34	23451	1.0834889
35	28991	1.2927991
36	31386	1.4926216

A fianco della serie è stata inserita una serie artificiale ottenuta duplicando ogni anno gli indici netti di stagionalità



Sulla serie destagionalizzata si può ora realizzare una stima più attendibile del ciclo-trend

Stima del ciclo-Trend da serie destagionalizzate

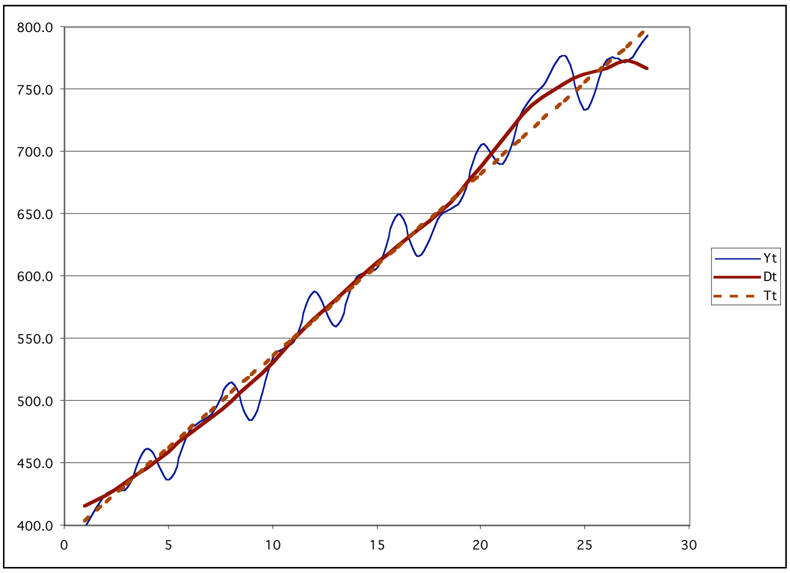
Serie del Pil USA

Stima del Trend sui valori destagionalizzati

14.638	389.598
0.236	3.921
0.993	10.098
3838.682	26.000
391456.159	2651.395

Stima del Trend sui valori destagionalizzati

14.966	385.352
0.415	6.890
0.980	17.742
1299.892	26.000
409194.655	8184.574



I parametri delle due rette stimate sono simili. Questo succede se la stagionalità è poco accentuata

Esercizio_9

1) Depurate dalla stagionalità la seguente serie dei numeri indici mensili relativi all'andamento del commercio al dettaglio nel periodo '84-'87

Mesi	1984	1985	1986	1987
Gennaio	93.09	98.82	105.64	106.39
Febbraio	93.69	95.59	99.66	105.80
Marzo	104.29	110.17	114.24	120.44
Aprile	104.34	113.11	115.71	125.37
Maggio	111.31	120.34	125.42	129.07
Giugno	111.98	114.96	120.35	128.98
Luglio	106.55	115.49	120.74	128.95
Agosto	110.65	121.12	124.06	131.02
Settembre	103.93	114.17	124.65	123.77
Ottobre	109.23	116.14	123.06	127.21
Novembre	113.28	118.56	120.79	125.38
Dicembre	131.65	139.40	151.26	154.75

2) Rappresentate graficamente la serie originaria, la serie ciclo-trend e la serie destagionalizzata

Alternativa di stima

I coefficienti di stagionalità (nel caso questa sia costante) si possono stimare con il modello di regressione.

Partendo dal modello moltiplicativo

Ciclo-trend   stagionalità e casualità

$$y_t = (T_t * C_t) * S_t * u_t$$

facciamo le ipotesi

$$T_t = e^{\sum_{i=0}^{m} b_i (t)^i}$$ Una funzione poli-esponenziale

$$C_t = 1$$ Assenza di effetti ciclici

$$S_t = e^{\sum_{i=1}^{k} a_i S_{ti}}$$ S_{ti} sono le dummies stagionali. $\sum_{i=1}^{k} S_{ti} = 1$

Il modello, nei logaritmi, diventa

k=periodicità dei dati

$$\mathrm{Log}[y_t] = \sum_{i=0}^{m} b_i t^i + \sum_{i=1}^{k} a_i s_{ti} + \mathrm{Log}[u_t]$$

Alternativa di stima/1

La stima dei parametri "b" ed "a" può effettuarsi con la regressione multipla ed anzi, le sue diagnostiche danno suggerimenti sulla stagionalità e sulla sua stabilità

La stima dei parametri "a" deve però rispettare il vincolo

$$\sum_{i=1}^{k} \alpha_i = -b_0$$

che esprime l'esaurirsi degli effetti stagionali ALL'INTERNO DELL'ANNO

Nell'ambito della formulazione ADDITIVA NEI LOGARITMI il significato degli "α" è quello già visto nel modello di regressione

α_i: **misura come si modifica il trend, per gli effetti stagionali, nella frazione i-esima dell'anno**

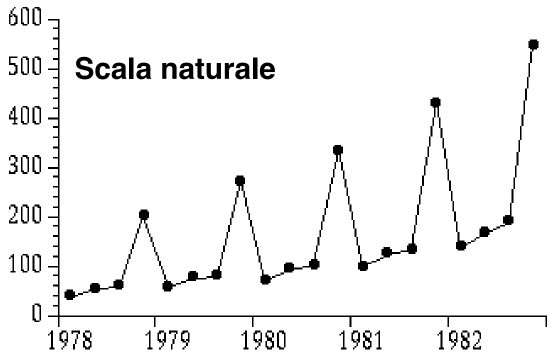
$$\mathrm{Log}\left[\hat{y}_{tj}\right] = \left(\sum_{i=1}^{m} \beta_i t^i\right) + \alpha_j; \quad t = 1, 2, \ldots, \; n; j = 1, 2, \ldots k$$

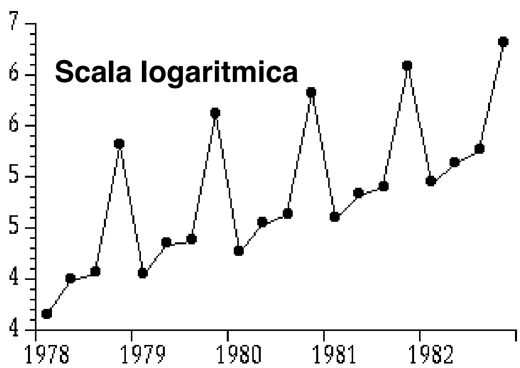
Esempio

La TOYS "R" è forse il maggiore rivenditore di giocattoli nel mondo (solo negli USA controlla 150 supermercati specifici del settore).

Ecco l'andamento delle sue vendite

Anni	Trim.	Vendite
1978	Feb.1-Apr.30	38,0
	Mag.1-Lug.31	53,6
	Ago.1-Ott.31	57,5
	Nov.1-Gen.31	200,0
1979	Feb.1-Apr.30	56,5
	Mag.1-Lug.31	75,8
	Ago.1-Ott.31	78,3
	Nov.1-Gen.31	269,7
1980	Feb.1-Apr.30	70,2
	Mag.1-Lug.31	92,7
	Ago.1-Ott.31	101,8
	Nov.1-Gen.31	332,6
1981	Feb.1-Apr.30	97,3
	Mag.1-Lug.31	123,7
	Ago.1-Ott.31	132,9
	Nov.1-Gen.31	429,4
1982	Feb.1-Apr.30	138,3
	Mag.1-Lug.31	167,6
	Ago.1-Ott.31	189,9
	Nov.1-Gen.31	545,9





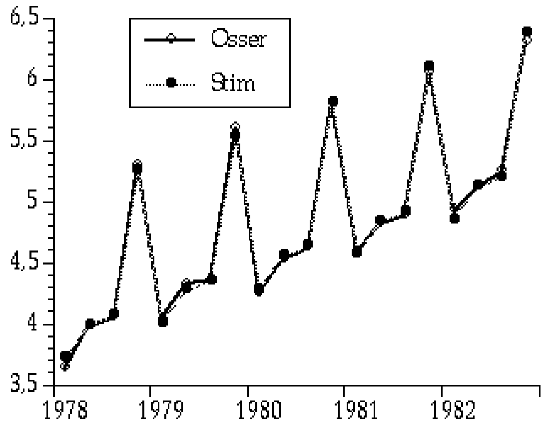
L'andamento in scala logaritmica è più "diretto": un trend esponenziale in "y" diventa lineare in Ln(y)

Esempio_Bis

Ecco la stima del modello

$$\hat{y}_t = \underset{(36.001)}{0.282t} + \underset{(106.431)}{3.439s_{t1}} + \underset{(114.774)}{3.709s_{t2}} + \underset{(117.206)}{3.788s_{t3}} + \underset{(153.695)}{4.967s_{t4}} \quad R^2 = 0.9999$$

Un esame di massima delle diagnostiche ne conferma la validità nonché la forte stagionalità trimestrale



La serie osservata e quella stimata dal modello sono coincidenti.

Il modello stesso può ben sostituirsi alla realtà della serie per l'interpolazione e la estrapolazione

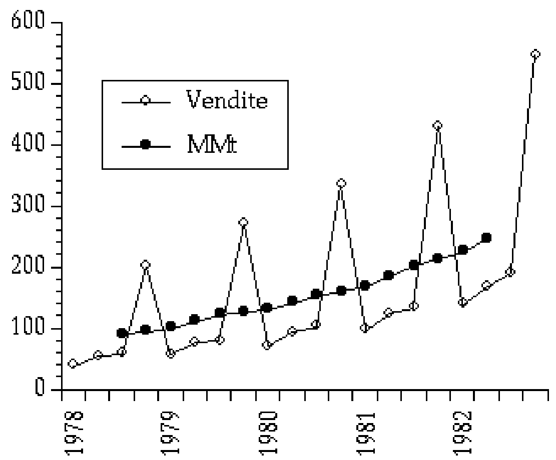
Esempio_ter

Quale sarebbe stato il risultato se avessimo "modellato" la serie storica in base all'approccio classico?

Calcoliamo innanzitutto le medie mobili ed i rapporti di stagionalità

Anni	Vendite	MMt
1978	38.0	
	53.6	
	57.5	89.588
	200.0	94.675
1979	56.5	100.050
	75.8	111.363
	78.3	121.788
	269.7	125.613
1980	70.2	130.663
	92.7	141.463
	101.8	152.713
	332.6	159.975
1981	97.3	167.738
	123.7	183.725
	132.9	200.950
	429.4	211.563
1982	138.3	224.175
	167.6	245.863
	189.9	
	545.9	

rs
0.5758
0.6740
0.6544
2.0959



Le MM eliminano drasticamente la stagionalità che del resto si concentra nell'ultimo trimestre come si vede dal grafico e come mostrano gli "rs"

Esempio_quater

Nelle MM rimane il solo trend che può essere stimato è una funzione lineare nei logaritmi

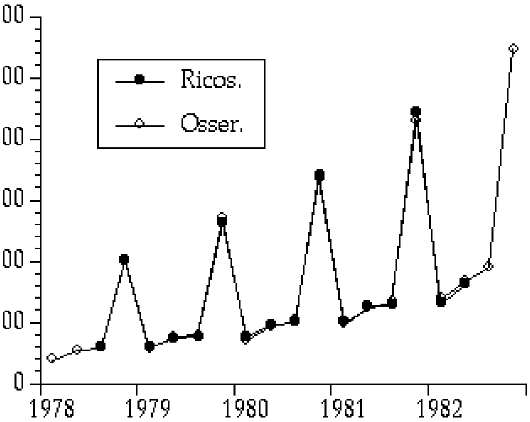
$$\underset{}{Ln(MM_t) = \underset{(434.01)}{4.429} + \underset{(62.581)}{0.066t}}; \quad \bar{R}^2 = 0.996$$

A questo punto possiamo ricostruire la serie storica

$$\hat{y}_t = \hat{T}_t * \hat{S}_t \quad \textbf{(si ignora la casualità!)}$$

$$\hat{T}_t = e^{4.429+0.066t}; \quad t = 1,2,\ldots,n$$

$$\hat{S}_t = rs$$



La ricostruzzione è più che buona. Però mancano i primi e gli ultimi due dati. In questo caso di grande regolarità possono essere estrapolati facilmente

Esercizio_10

I dati seguenti rappresenta l'andamento trimestrale delle spese generali di una grande impresa (dati in miliardi)

Anno	Trim	t	y_t
1986	1	1	10.8
	2	2	9.8
	3	3	9.4
	4	4	9.8
1987	1	5	9.9
	2	6	9.0
	3	7	8.6
	4	8	9.4
1988	1	9	9.7
	2	10	9.1
	3	11	9.0
	4	12	9.8
1989	1	13	9.8
	2	14	9.0
	3	15	8.6
	4	16	9.1

1) Costruire la serie delle medie mobili

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

2) Determinare i rapporti di stagionalità

3) Destagionalizzare i dati

4) Sui dati destagionalizzati individuare e stimare il trend

Confronto

Australian production of electricity
March 1956 - Sep 1994
m. kwH monthly figures aggregate to quarters

Le funzioni periodiche

Una funzione è detta periodica se

$$f(t) = f(t + kp) \qquad \forall \; t, \; k = 0, \pm 1, \pm 2, \dots$$

il valore più piccolo di "p" per cui è vera la relazione è il "periodo" della funzione. Se un tale "p" non esiste, la funzione non è periodica.

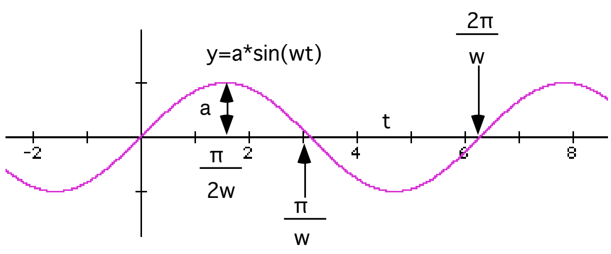
Le funzioni periodiche più comuni sono il seno e coseno trigonometrici

$$a_k * sin(w_k t) \qquad e \qquad b_k * \cos(w_k t)$$

periodiche con periodo:

$$p = \frac{2\pi}{w};$$

$$w = \frac{2\pi}{p} = frequenza \;\; angolare$$



Seno e coseno

Le funzioni seno e coseno sono i prototipi delle funzioni periodiche.

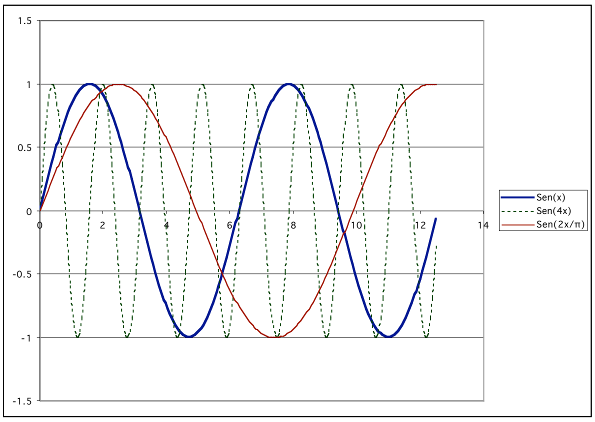
La prima questione da affrontare è come sono fatte le funzioni sen(a x) , dove a è un numero reale positivo fissato.

Considerare ax come argomento del seno corrisponde a cambiare la scala delle ascisse cioè il grafico di sen(ax) è quello di sin(x) ``dilatato'' o “contratto” secondo che a sia minore o maggiore di uno in valore assoluto

sen(x) nell’intervallo (0,4π) ha due picchi e due valli .

sen(4x) ha 8 picchi e 8 valli

sen(2x/π) sale e scende una sola volta



Seno e coseno/2

Se le funzioni seno e coseno hanno periodo “p” allora saranno anche periodiche con periodo “2p”, “3p”, …

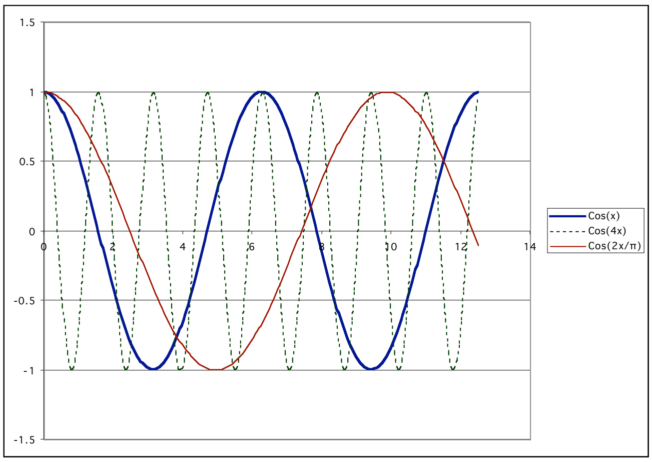
In genere, quando si parla di periodo si sottintende che si considera il periodo minimo, definito come il più piccolo p tale che f(x+p) = f(x) per ogni x

Piú in generale sen(kx) e cos(kx) , con k$ intero, hanno come periodo minimo p=2k/π, ma sono anche periodiche di periodo 2π .

cos(x) nell’intervallo (0,4π) ha tre picchi e due valli.

cos(4x) ha 9 picchi e 8 valli

cos(2x/π) sale una volta e scende due volte



Stagionalità e funzioni trigonometriche

La componente stagionale può essere espressa da una combinazione lineare di funzioni trigono-metriche dette "armoniche":

$$S_t = \sum_{i=1}^{m} a_i \operatorname{sen}\left(\frac{2\pi i}{s} t + \varphi_i\right)$$

Dove:

$S_t =$ *coefficiente di stagionalità*

$s =$ *periodicità della stagionalità*

$f_i = \dfrac{2\pi i}{s}$ *frequenza della armonica*

$\varphi_i =$ *ampiezza o fase della armonica*

$s/i =$ *numero di periodi in cui si completa la stagionalit*

$m =$ *numero massimo di armoniche in* S_t

Ogni armonica è un'onda sinusoidale che completa il suo ciclo in s/i periodi.

Il coefficiente di stagionalità è la somma di tali armoniche

Stagionalità e funzioni trigonometriche/2

La durata più breve in cui si può parlare di "ciclo" è 2 (da un dato a quello successivo). Il massimo numero di armoniche è

$$m = \begin{cases} \dfrac{s}{2} & se\ \ s\ \ è\ \ pari \\ \dfrac{s-1}{2}; & se\ \ s\ \ è\ \ dispari \end{cases}$$

Infatti, in caso di "s" pari, la m-esima armonica completerebbe il suo ciclo in un numero di periodi pari a:

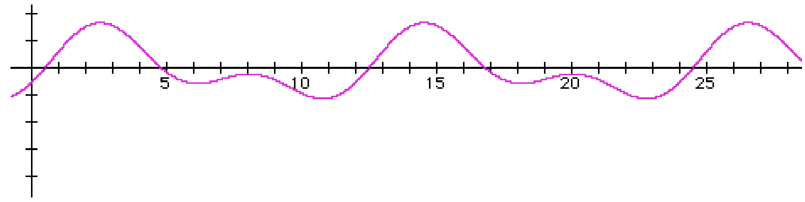
$$\frac{s}{m} = \frac{s}{\frac{s}{2}} = 2$$

D'altra parte non è necessario usare tutte le armoniche dato che le prime due o tre sono già in grado di esprimere complesse strutture stagionali

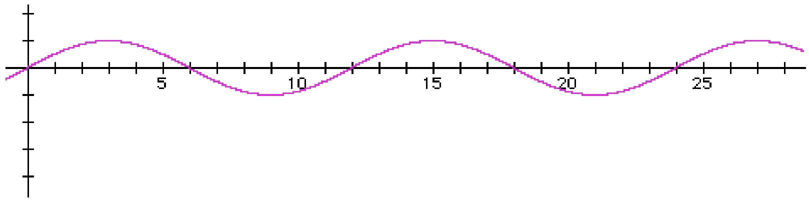
Stagionalità e funzioni trigonometriche/3

Esempio di armonica

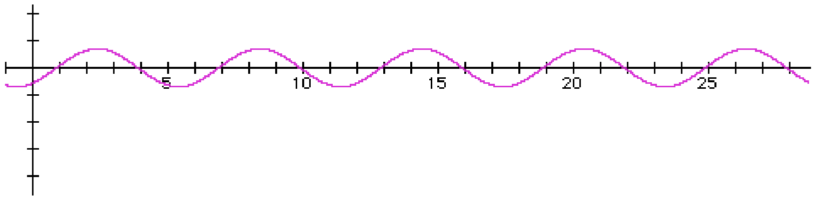
$$a_1 = 0;\ \ \phi_1 = 0;\ \ a_2 = -0.7;\ \ \varphi_2 = 0.6944\pi\ \ (125°)$$



Stagionalità
1ª arm+2ª arm



1ª armonica



2ª armonica

Trend + armoniche

Se il trend è una polinomiale in "t", il modello con stagionalità armonica diventa:

$$y_t = \beta_0 + \sum_{i=1}^{k} \beta_i t^k + \sum_{j=1}^{m} a_i sin\left(f_i t + \varphi_i\right) + u_t$$

il modello non è lineare a causa della presenza di un parametro (ϕ) nello argomento del seno. Tuttavia, tenuto conto che:

$$sin(x+y) = sin(x) * \cos(y) + sin(y) * \cos(x)$$

La relazione può essere linearizzata:

$$\begin{aligned} y_t &= \beta_0 + \sum_{i=1}^{k} \beta_i t^k + \sum_{j=1}^{m} a_i \left[\gamma_{i1} sin\left(f_i t\right) * \cos\left(\varphi_i\right) + \gamma_{i2} \cos\left(f_i t\right) * sin\left(\varphi_i\right)\right] + u_t \\ &= \beta_0 + \sum_{i=1}^{k} \beta_i t^k + \sum_{j=1}^{m} \left[\gamma_{i1} sin\left(f_i t\right) + \gamma_{i2} \cos\left(f_i t\right)\right] + u_t \end{aligned}$$

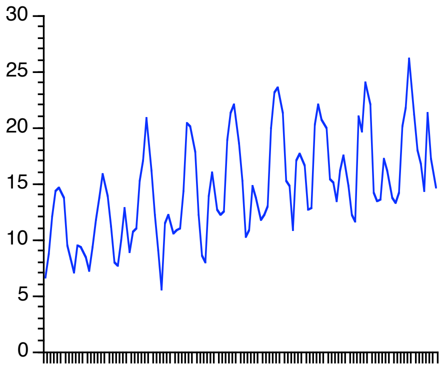
Nella seconda relazione sono stati accorpati i parametri incogniti

Trend + armoniche/2

Dato che la stagionalità deve compensarsi nell'arco dell'anno si pone il vincolo:

$$\sum_{i=1}^{m} \gamma_{1i} sin(f_i t) + \gamma_{2i} \cos(f_i t) = 0 \Rightarrow$$

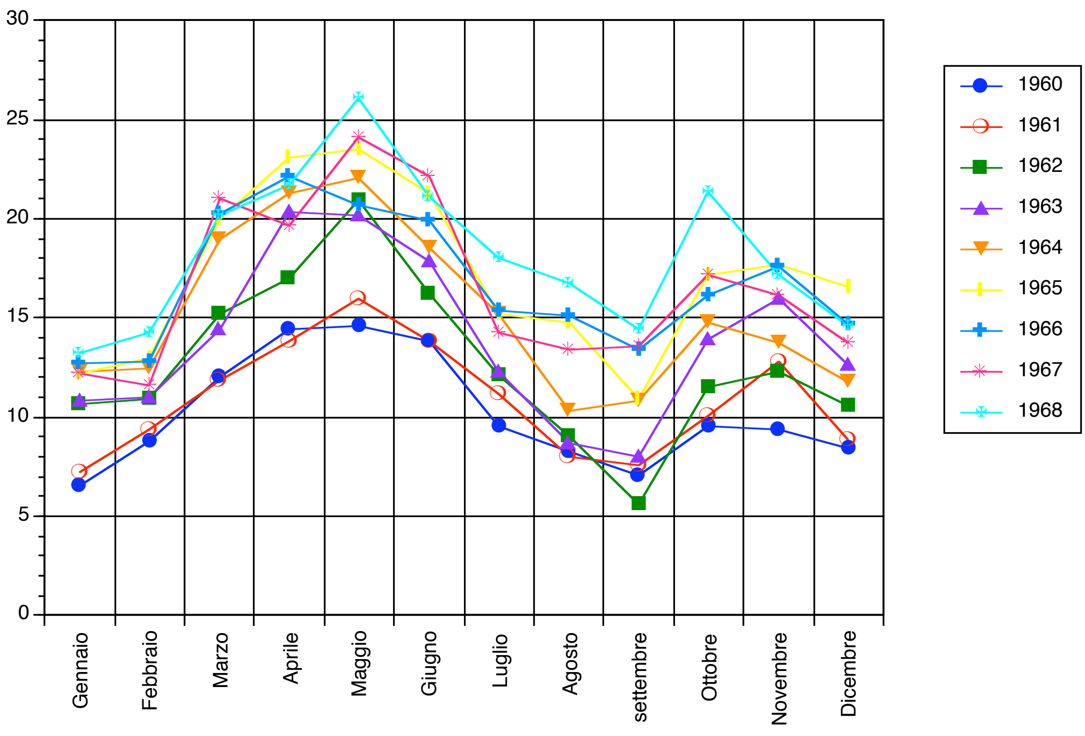
$$\gamma_{2m} = \frac{-\sum_{i=1}^{m-1}\left[\gamma_{1i} sin(f_i t) + \gamma_{2i} \cos(f_i t)\right] - \gamma_{1m} sin(f_m t)}{\cos(f_m t)}$$



Vendite mensile di auto in Quebec-Canada

6.550	8.728	12.026	14.395	14.587	13.791	9.498	8.251	7.049	9.545	9.364	8.456
7.237	9.374	11.837	13.784	15.926	13.821	11.143	7.975	7.610	10.015	12.759	8.816
10.677	10.947	15.200	17.010	20.900	16.205	12.143	8.997	5.568	11.474	12.256	10.583
10.862	10.965	14.405	20.379	20.128	17.816	12.268	8.642	7.962	13.932	15.936	12.628
12.267	12.470	18.944	21.259	22.015	18.581	15.175	10.306	10.792	14.752	13.754	11.738
12.181	12.965	19.990	23.125	23.541	21.247	15.189	14.767	10.895	17.130	17.697	16.611
12.674	12.760	20.249	22.135	20.677	19.933	15.388	15.113	13.401	16.135	17.562	14.720
12.225	11.608	20.985	19.692	24.081	22.114	14.220	13.434	13.598	17.187	16.119	13.713
13.210	14.251	20.139	21.725	26.099	21.084	18.024	16.722	14.385	21.342	17.180	14.577

Applicazione alle vendite auto



La stagionalità è presente stabilmente. C'è anche il trend

Applicazione alle vendite auto/2

variable	BETA	St. Err. of BETA	St. Err. B	of B	t(94)	p-level
T	.584010	.0317114	.08438	.004582	18.4164	.0000000
ARM1S	.387389	.0316323	2.46699	.201442	12.2466	.0000000
ARM1C	-.424435	.0315324	-2.70433	.200912	-13.4603	.0000000
ARM2S	-.474496	.0315446	-3.02193	.200898	-15.0421	.0000000
ARM2C	.102745	.0315327	.65460	.200899	3.2584	.0015602
ARM3S	-.077766	.0315332	-.49534	.200855	-2.4662	.0154688
ARM3C	-.047818	.0315332	-.30461	.200873	-1.5164	.1327701
ARM4S	.002316	.0315430	.01476	.200980	.0734	.9416296
ARM4C	-.029896	.0315344	-.19038	.200819	-.9480	.3455432
ARM5S	.125483	.0316254	.80024	.201683	3.9678	.0001417
ARM5C	-.011514	.0315384	-.07326	.200667	-.3651	.7158611
ARM6S	.012266	.0633775	.49275	2.545934	.1935	.8469526

Il modello ha R^2 corretto: 0.894

Le prime tre armoniche sono significative. Anche una componente della quinta. Conviene però semplificare il modello.

Il periodogramma

I coefficienti stimati delle componenti di ogni armonica danno indicazione sul comportamento della stagionalità.

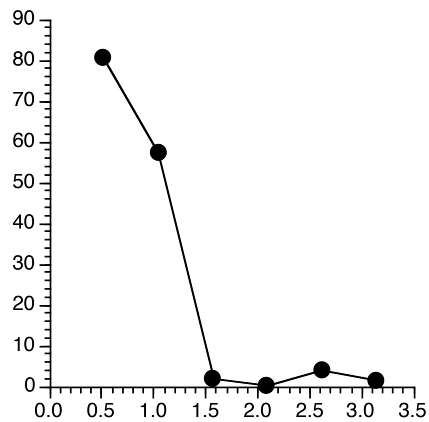
In particolare:

$$I_i = \frac{s}{2}\left[a_i^2 + b_i^2\right]; \quad i = 1,2,\ldots,m \quad dove \;\; m = \begin{cases} \frac{s}{2} \;\; se \;\; s \;\; è \;\; pari \\ \frac{s-1}{2}; \;\; se \;\; s \;\; è \;\; dispari \end{cases}$$

misura il contributo della k-esima armonica

Lo scatterplot di I_i contro f_i dovrebbe chiarire quali armoniche abbiano un ruolo più importante.

La 1ª e la 2ª dovrebbero bastare

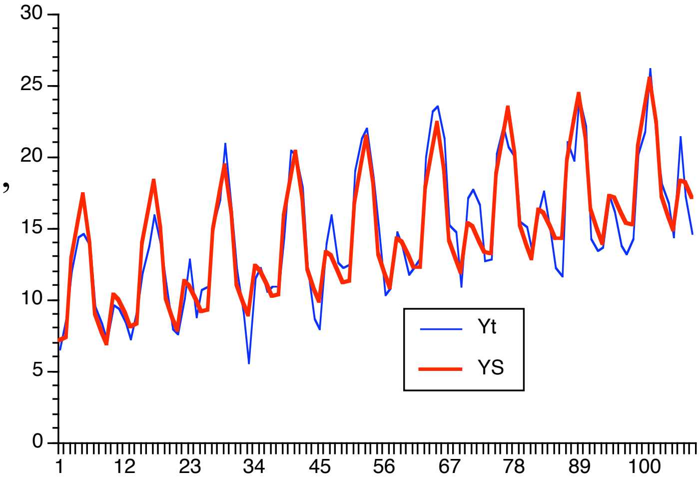


Applicazione alle vendite auto/3

variable	BETA	St. Err. of BETA	B	St. Err. of B	t(101)	p-level
Constant			9.9968	0.2905	34.42	.0000000
T	.583843	.0320633	.08435	.0046325	18.2091	.0000000
ARM1S	.387378	.0320220	2.46692	.2039238	12.0972	.0000000
ARM1C	-.424437	.0319130	-2.70434	.2033369	-13.2998	.0000000
ARM2S	-.474501	.0319314	-3.02196	.2033617	-14.8600	.0000000
ARM2C	.102732	.0319130	.65452	.2033226	3.2191	.0017295
ARM5S	.125209	.0319067	.79849	.2034769	3.9242	.0001590

R^2 corretto: 0.891

Il risultato è soddisfacente, ma si può fare di meglio.



Verifica del modello di stagionalità

Il modello di stagionalità costante è molto conveniente in quanto semplifica il problema della destagionalizzazione.

Basta infatti dividere il dato stagionale per il coefficiente di stagionalità della frazione d'anno interessata ed il fattore "S" sparisce (mai del tutto)

Che ragioni abbiamo per adottare tale modello? Ancora una volta sono i grafici a confermare o sconfermare la nostra ipotesi.

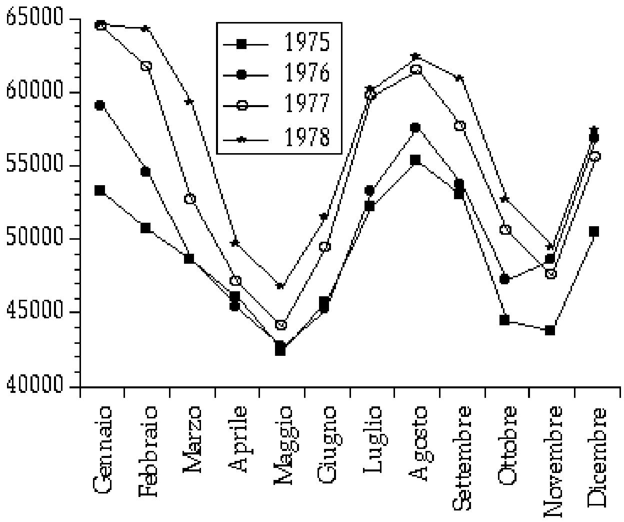
I primi suggerimenti ci arrivano dal grafico in cui abbiamo già verificato la presenza o meno di stagionalità.

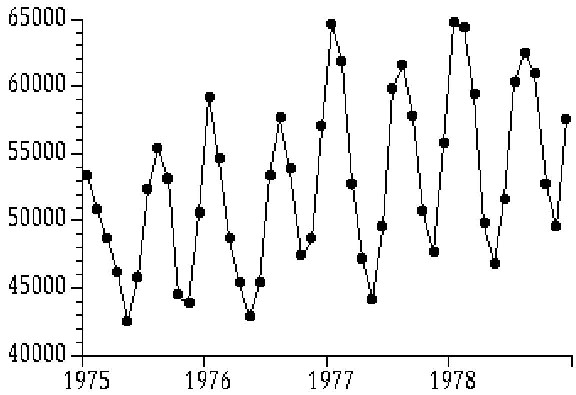
Nel momento in cui la accertiamo siamo anche in grado di valutarne la stabilità

Esempio

Andamento dei consumi domestici di elettricità

Mese	1975	1976	1977	1978
Gennaio	53299	59088	64516	64624
Febbraio	50716	54530	61705	64283
Marzo	48595	48656	52686	59283
Aprile	46036	45365	47118	49722
Maggio	42424	42786	44086	46764
Giugno	45741	45262	49481	51533
Luglio	52275	53312	59748	60266
Agosto	55310	57556	61541	62366
Settembre	53057	53746	57687	60883
Ottobre	44430	47296	50599	52656
Novembre	43824	48582	47568	49440
Dicembre	50442	56893	55611	57458





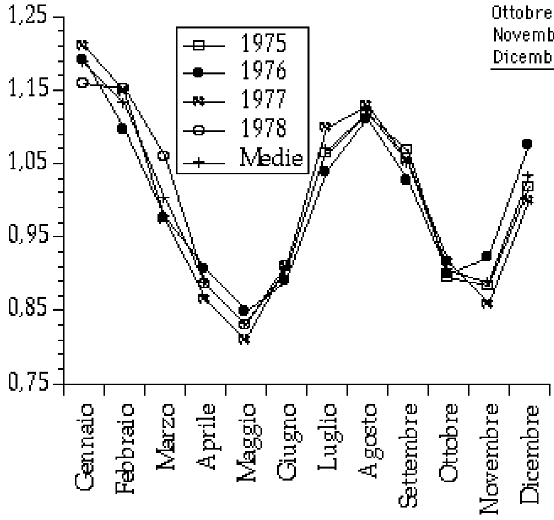
La presenza della stagionalità è lampante ed è pure evidente la sua stabilità

Un altro ausilio grafico

Calcoliamo i coefficienti di stagionalità

Invece dei dati stagionali si possono rappresentare i rapporti grezzi di stagionalità

Mese	1975	1976	1977	1978	Medie	RS
Gennaio		1.1906	1.2093	1.1588	1.1862	1.1865
Febbraio		1.0957	1.1473	1.1515	1.1315	1.1318
Marzo		0.9753	0.9736	1.0588	1.0026	1.0028
Aprile		0.9067	0.8659	0.8846	0.8857	0.8859
Maggio		0.8497	0.8088	0.8295	0.8293	0.8295
Giugno		0.8906	0.9093	0.9116	0.9039	0.9041
Luglio	1.0649	1.0389	1.0990		1.0676	1.0679
Agosto	1.1177	1.1102	1.1296		1.1192	1.1194
Settembre	1.0687	1.0275	1.0515		1.0492	1.0495
Ottobre	0.8953	0.9000	0.9159		0.9038	0.9040
Novembre	0.8834	0.9223	0.8576		0.8878	0.8880
Dicembre	1.0169	1.0754	0.9991		1.0304	1.0307
					11.9972	



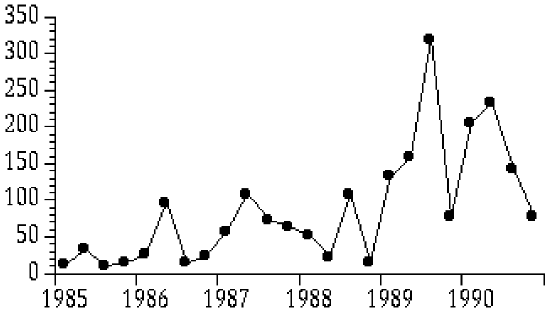
Anche il grafico dei coefficienti grezzi conferma l'ipotesi di stagionalità costante.

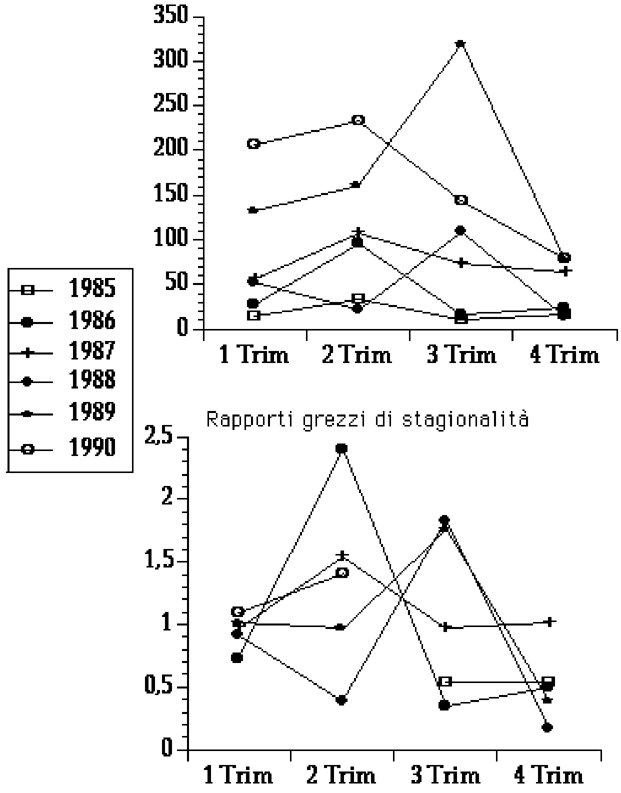
Infatti, la struttura degli "rs" rimane identica nel corso degli anni

Esempio

Andamento trimestrale delle esportazioni di una impresa dolciaria

Anno	Trim	yt	Anno	Trim	yt
1985	1	13.59	1988	1	52.87
	2	32.75		2	21.44
	3	10.69		3	108.88
	4	16.03		4	15.42
1986	1	27.99	1989	1	133.11
	2	96.48		2	160.18
	3	15.60		3	319.40
	4	24.76		4	77.59
1987	1	56.47	1990	1	205.70
	2	108.44		2	232.68
	3	72.54		3	143.59
	4	64.38		4	78.71





Esempio_bis

Che succede se ignoriamo la irregolarità delle fluttuazioni stagionali e usiamo il modello costante?

Ecco la stima dei rapporti di stagionalità.

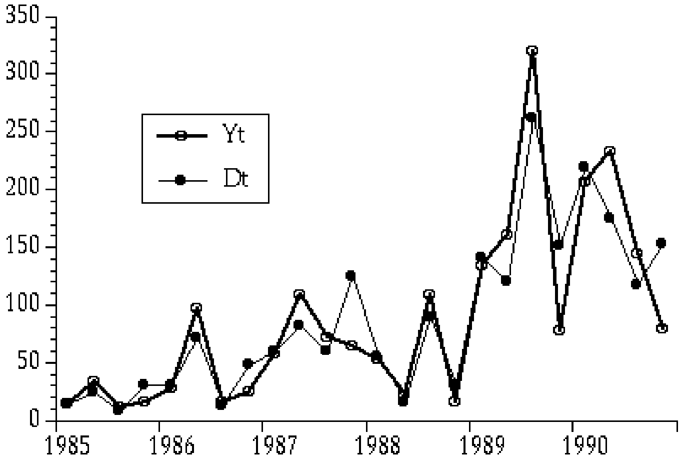
Trim.	1985	1986	1987	1988	1989	1990	Media	rs
1		0.7287	0.9665	0.9219	1.0184	1.1008	0.9473	0.9397
2		2.4050	1.5380	0.3845	0.9720	1.4099	1.3419	1.3312
3	0.5328	0.3486	0.9671	1.8242	1.7584		1.2246	1.2148
4	0.5373	0.4969	1.0110	0.1771	0.3884		0.5183	0.5142
							4.0320	4.0000

I valori segnalano variazioni molto alte rispetto al trend

La serie depurata non differisce da quella originaria.

Anzi, applicando il modello sbagliato si può rischiare di inserire strutture inesistenti come l'effetto Slutzky-Yule.

Se la stagionalità non è costante è preferibile non procedere alla depurazione ed utilizzare altri metodi.



Esercizio_11

Per i seguenti dati quadrimestrali

Anno	t	y_t
1981	1	12.2
	2	14.7
	3	18.6
1982	4	20.2
	5	21.6
	6	23.1
1983	7	27.7
	8	27.0
	9	29.6
1984	10	30.9
	11	34.7
	12	35.6
1985	13	39.2
	14	41.1
	15	40.9
1986	16	43.2
	17	45.8
	18	50.6

1) Accertare la presenza di stagionalità

2) Nel caso sia presente valutare l'ipotesi di stagionalità costante