ANALISI DELLE COMPONENTI DI UNA SERIE TEMPORALE

NOZIONI TEORICHE

L'analisi delle serie temporali si può utilizzare con diversi obiettivi, che possono andare da una mera analisi descrittiva dell'evoluzione della serie, fino all'obiettivo di predire a breve o lungo periodo valori futuri della serie.
L'impostazione classica delle serie temporali prende come punto di partenza la decomposizione della serie temporale in quattro fattori o componenti. Secondo questo impostazione, ogni valore osservato di una serie temporale è il risultato della combinazione di quattro fattori o componenti che possono essere presenti o determinare l'evoluzione lungo il tempo della serie. Queste componenti ipotetiche, che non si osservano nella realtà, sono le seguenti:

Tendenza (T_t):
Riflette il movimento a lungo termine della serie. Cioè, la crescita; decrescita o ristagno che si produce di formazione lenta in un lungo periodo di tempo.
Per esempio il numero di stranieri che visitano il nostro Paese mostra una chiara tendenza crescente.
Per potere osservare questa componente bisogna disporre di un numero sufficientemente grande di osservazioni.

Fluttuazioni stagionali (S_t)
Sono i movimenti della serie a breve periodo che si ripetono periodico ogni anno. Per esempio, nella costa del sole, ogni mese di agosto si producono tassi di occupazione alberghiera più elevati, mentre nei mesi invernali, anno dietro anno, si sperimentano sempre tassi di occupazione più ridotti dell'anno. Queste fluttuazioni sono causati per motivi tanto diversi come le causate climatiche, i periodi vacanzieri, etc. Per potere osservare questa componente è necessario che la serie abbia una periodicità inferiore all'anno (serie mensili, trimestrali...).

Fluttuazioni cicliche (C_t)
Questa componente riflette i movimenti oscillatori della serie a medio termine. Esistono molte serie che mostrano picchi in epoche di crescita economica e discese importanti in momenti di recessione. Il periodo di queste fluttuazioni sono poco esatte e suole variare tra quattro ed otto anni. Nella pratica, è difficile separare le fluttuazioni cicliche dalla tendenza, motivo per il come, alle volte, si considera una sola componente detta tendenza-ciclo.

Variazioni accidentali o irregolari (I_t)
La componente irregolare di una serie assorbe la variazione a breve termine che generalmente, restano fuori dell'analisi nelle nei tre componenti descritte anteriormente. Include tanto movimenti della serie causati da fattori accidentali con un effetto significativo (effetti di uno sciopero, di una catastrofe naturale,...), come movimenti di minore ammontare determinati da fattori non individuabili.
En l'evoluzione di una serie temporale si posso presentare tutti o solo alcune delle componenti. Per esempio, in una serie con dati annuali non si possono evidenziare fluttuazioni stagionali.
Si possono impostare teorico diversi formazioni di combinando delle componenti (non osservate) per generare i valori osservati della serie.
Le più utilizzate sono l'ipotesi additiva e l'ipotesi moltiplicativa, che consistono nella somma o nel prodotto delle quattro componenti per ottenere li valori della serie.
Ipotesi additiva: $Y_t = T_t + S_t + C_t + I_t$
Ipotesi moltiplicativa: $Y_t = T_t \times S_t \times C_t \times I_t$
Sotto l'ipotesi di additiva, le quattro componenti sono espresse nella stessa unità di misura che la serie Y_t.
Sotto l'ipotesi moltiplicativa, invece, solo una componente (generalmente la tendenza) è espressa

nell'unità della serie e le restanti sono adimensionali e si esprimono in unità o percentuali rispetto alla tendenza.
In molte occasioni, nell'analisi di una serie temporale si è interessati allo studio di solo una dei componenti, procedendo all'eliminazione dalla serie delle resti componenti (se presenti).
Le componenti che maggiore interesse suscitano nello studio del turismo è la tendenza e la componente stagionale.

SIMA DELLA TENDENZA MEDIANTE L'ANALISI DELLA REGRESSIONE LINEARE

L'analisi della regressione lineare, si può impiegare per la stima della tendenza. Per utilizzare l'analisi della regressione per la della tendenza di una serie, bisogna adottare l'ipotesi additiva. Secondo questa ipotesi, i valori osservati della serie sono la somma delle quattro componenti. La componente di tendenza si sostituisce per un modello matematico, nel nostro caso un modello lineare che dipende del tempo ($T_t = a + bt$), ed il resto delle componenti si includono nei residui della regressione ($e_t = S_t + C_t + I_t$)

$$Y_t = a + bt + e_t$$

La stima dei parametri della tendenza lineare, a e b, si effettua mediante il metodo dei minimi quadrati ordinari.
Questo metodo consiste nel determinare una retta interpolante alla nuvola di punti risultante dalla rappresentazione grafica dei valori della serie Y_t sull'asse verticale, rispetto ad una variabile t, che rappresenta il tempo, rappresentato sull'asse orizzontale.
La retta aggiustata, per tanto, è una rappresentazione della tendenza lineare della serie. Il procedimento di aggiustamento consiste nel trovare la retta che renda minima la somma dei quadrati degli errori (differenziate tra i valori della serie ed i corrispondenti valori sulla retta). Nel quadro seguente si riassumono le quantità necessarie per la stima della tendenza lineare e le corrispondenti analogie impiegate nell'analisi di una relazione causale.

	Modello Casuale	Tendenza Lineare
Numero di dati		
Variabile Dipendente		Y_t
Variabile indipendente	X_i	t
Modello	$Y_i = a + b X_i + e_i$	$Y_t = a + b t + e_t$
Parametro a	$a = \bar{y}_i - b\bar{x}$	$a = \bar{y}_t - b\bar{t}$
Parametro b	$b = \dfrac{S_{x,y}}{S_x^2} = \dfrac{\dfrac{\sum x_i y_i}{N} - \overline{xy}}{\dfrac{\sum x_i^2}{N} - \bar{x}^2}$	$b = \dfrac{S_{t,yt}}{S_t^2} = \dfrac{\dfrac{\sum t y_t}{T} - \overline{t y_t}}{\dfrac{\sum t^2}{T} - \bar{t}^2}$
Coefficiente di determinazione	$R^2 = \dfrac{S_{xy}^2}{S_x^2 S_y^2} = \dfrac{\left(\dfrac{\sum x_i y_i}{N} - \overline{xy}\right)^2}{\left(\dfrac{\sum x_i^2}{N} - \bar{x}^2\right)\left(\dfrac{\sum y_i^2}{N} - \bar{y}^2\right)}$	$R^2 = \dfrac{S_{t y_t}^2}{S_t^2 S_{y_t}^2} = \dfrac{\left(\dfrac{\sum t y_t}{T} - \overline{t y_t}\right)^2}{\left(\dfrac{\sum t^2}{T} - \bar{t}^2\right)\left(\dfrac{\sum y_t^2}{T} - \bar{y}_t^2\right)}$

Questo metodo offre inoltre la possibilità di ottenere una misura della bontà dell'aggiustamento, il coefficiente di determinazione R^2, che ci informa in che grado la tendenza stimata descrive il comportamento della serie lungo il tempo, così come dell'affidabilità che meritano le predizioni realizzate con il modello aggiustato.
La tendenza che si stima con il suddetto metodo, si chiama tendenza deterministica che si mantiene costante durante tutto il periodo analizzato. Questo metodo è applicabile solo se la nube dei punti assume una forma che possa essere interpolata da una funzione di primo grado.

Cioè, non si deve applicare se la serie mostra un comportamento crescente in un subperiodo di tempo e decrescente in altro, o viceversa, o se la tendenza osservata è di formazione non lineare.
D'altronde, se il componente stagionale della serie è molto accentuato, dato che questa componente è contenuta nei residui, questa avrà un'importante molto grande nel modello e ne rende difficile la stima, in questi casi si deve eliminare la componente stagionale prima di applicare il metodo della regressione per la stima della tendenza.

STIMA DELLA TENDENZA MEDIANTE MEDIE MOBILI

Non sempre si trovano nella pratica di tendenze deterministiche. Di fatto, è molto frequente trovare tendenze che vanno cambiando lungo il tempo che non seguono una funzione matematica semplice, che permetta la sua stima mediante il metodo della regressione lineare. In questo caso si può impiegare il metodo delle medie mobili.
Il metodo delle medie mobili consiste nel sostituire ad ogni valore della serie la media aritmetica degli n valori più vicini (compreso lo stesso valore). Per esempio, una serie di medie mobili di ampiezza tre, relativa a sei osservazioni è la seguente:

Periodo t	0	1	2	3	4	5
Serie	Y_0	Y_1	Y_2	Y_3	Y_4	Y_5
Media Mobile (3)	…..	$\frac{Y_0+Y_1+Y_2}{3}$	$\frac{Y_1+Y_2+Y_3}{3}$	$\frac{Y_2+Y_3+Y_4}{3}$	$\frac{Y_3+Y_4+Y_5}{3}$	…….

L'esempio serve per constatare che non si può calcolare la media mobile del primo e dell'ultimo periodo di tempo per ampiezza tre. Se l'ampiezza della media è di cinque periodi, non si possono calcolare i corrispondenti 4 periodi (primi due e ultimi due).
Quando si prendono medie mobili con un numero pari di periodi è necessario affrontare il problema del valore centrale. Per esempio, se prendiamo un'ampiezza di quattro periodi, la media dei periodi 0,1,2 e 3 corrisponde al periodo 1,5, che non è osservato. Per ciò, nella pratica quello che si fa è prendere un periodo in più e ponderare il primo e l'ultimo per 0,5. Un esempio è riportato nella tabella seguente.

Periodo t	0	1	2	3	4	5
Serie	Y_0	Y_1	Y_2	Y_3	Y_4	Y_5
Media Mobile (4)	…..	…..	$\frac{\frac{Y_0}{2}+Y_1+Y_2+Y_3+\frac{Y_4}{2}}{4}$	$\frac{\frac{Y_1}{2}+Y_2+Y_3+Y_4+\frac{Y_5}{2}}{3}$	…..	…….

La serie risultante applicando il metodo delle medie mobili "liscia" la serie, eliminando i salti o i picchi che mostra la serie originale, (i valori più elevati saranno sostituiti con valori compensati più bassi). D'altronde, se la serie presenta un comportamento stagionale, che si ripete periodicamente ogni anno, una serie di valori ottenuti con medie mobili di ampiezza uguale al numero di osservazioni che contiene ogni anno elimina detto effetto di stagionalità.

STIMA DELLA COMPONENTE STAGIONALE MEDIANTE COEFFICIENTE DI MEDIE MOBILI

È molto frequente, negli studi turistici, l'interesse di calcolare la componente stagionale di una serie con l'obiettivo di realizzare un'analisi specifica su questa componente. Il metodo più semplice per isolare la componente stagionale è il coefficiente di medie mobili. Se si adotta l'ipotesi moltiplicativa, il quoziente tra la serie originale Y_t e la serie di medie mobili MM_t, di periodo uguale al numero di osservazioni contenute in un anno, include solo le componenti stagionale ed irregolare:

$$\frac{Y_t}{MM_t} = \frac{T_t \times S_t \times C_t \times I_t}{T_t \times C_t} = S_t \times I_t$$

Questo quoziente si chiama *indice specifico di variazione stagionale* (ISVS) e si può esprimere in unità o percentuali (moltiplicando lo per 100).
Se la stagionalità è stabile e non esistono componenti irregolari di grandezza considerevole, gli indici specifici di variazione stagionale, si possono utilizzare per ottenere gli indici generali di variazione stagionale (IGVS) che contengono solo la componente stagionale. Anche gli indici generali si possono esprimere in unità o in percentuali. In una serie mensile, l'Indice generale di gennaio si ottiene calcolando la media aritmetica dell'indice specifico di tutti i mesi di gennaio osservati.
L'interpreta dell'Indice generale di variazione stagionale, non presenta grandi difficoltà di interpretazione.
In una serie mensile, un valore dell'Indice generale di 1,4 corrispondente al mese di agosto indica che per causa dell'effetto stagionale della serie, nel mese di agosto si osserva abitualmente un valore superiore del 40% a quello che mostra la tendenza-ciclo della serie.

ELIMINAZIONE DELLA COMPONENTE STAGIONALE

Nelle serie temporali turistiche, l'effetto di stagionalità può essere presente in una forza tale da rendere difficile l'analisi dell'altri componenti della serie.
Per questo è necessario utilizzare qualche metodo che permetto di eliminare in una serie la componente stagionale. Il procedimento di eliminazione dell'effetto stagionale di una serie si denomina destagionalizzazione e la serie che si ottiene si denomina serie destagionalizzata. Tra : i diversi metodi esistenti, il più semplice consiste dividendo la serie originale, Ytper l'indice generale di variazione stagionale (e moltiplicare per 100 se l'indice è espresso in percentuale). La serie destagionalizzata risultante, sotto l'ipotesi moltiplicativa, non contiene la componente stagionale.

$$\frac{Y_t}{IGVS_t} = \frac{T_t \times S_t \times C_t \times I_t}{S_t} = T_t \times C_t \times I_t$$