



Statistica Aziendale

Lezione 1

Orario Lezioni

Lezioni

Lunedì 9-11 Aula A8

Venerdì 14-16 Aula A8

Laboratorio informatico

Dott. Antonio Lucadamo

Le date saranno comunicate successivamente

Gli studenti **DEVONO** essere dotati di password per l'accesso al laboratorio informatico (richiesta al dott. Delli Veneri)

Programma del Corso

Parte Teorica

Richiami di Statistica descrittiva e inferenziale;

Introduzione all'analisi classica delle serie storiche e dei metodi di previsione;

Elementi di algebra delle matrici;

Introduzione all'Analisi Mutidimensionale dei dati: il Modello di Regressione Multipla;

l'Analisi in Componenti Principali e l'Analisi Fattoriale Discriminante;

Parte Applicativa

L'analisi statistica dei dati di bilancio: analisi univariata e multivariata.

Metodi statistici multivariati per l'analisi della concorrenza;

Modelli statistici per la previsione delle insolvenze.

Si consiglia la frequenza al Corso

Il Corso prevede lezioni frontali in Aula e in laboratorio informatico

E' previsto l'utilizzo di Software statistico per l'elaborazione dei dati

L'esame potrà essere svolto secondo due modalità:

- a) Prova di laboratorio informatico e colloquio;
- b) Esame scritto e colloquio

Libri Consigliati:

Appunti forniti dal docente disponibili sul sito internet della Facoltà

Statistica Aziendale

Definizione

Non c'è univocità sulla definizione di statistica aziendale né sulle tematiche di sua pertinenza.

Tutti i fenomeni aziendali che si prestano ad analisi statistica dovrebbero rientrare, anche se in modo non esclusivo, nell'ambito dello studio della statistica aziendale (Di Fonzo, 2002)

(es. Produzione, commercializzazione, personale, magazzino, direzione generale....)

Schema di struttura aziendale: aree e attività

AREA	ATTIVITA'
Produzione	Previsioni della produzione Controllo degli stock e delle scorte Pianificazione (manodopera, produzione, acquisti) Controllo di qualità
Contabilità e finanza	Formulazione del bilancio Planning Controllo di gestione
Servizi alla produzione	Modelli finanziari Analisi dei dati Simulazione Disegni sperimentali
Marketing e vendite	Previsioni delle vendite Analisi dei dati di vendita Ricerche di mercato
Ricerca e sviluppo	Analisi di dati sperimentali Simulazione

INFORMAZIONE ED ELABORAZIONE STATISTICA

L'informazione è una risorsa essenziale per il decisore d'azienda

La bontà di una decisione dipende, in larga misura, dalla quantità e dalla qualità dell'informazione disponibile

Cos'è "L'INFORMAZIONE"?

Dato, trasformato in forma utile per il ricevente, che possiede un valore reale o potenziale sulle decisioni attuali o future.

Il dato non costituisce di per sé un'informazione, ma richiede un'elaborazione.

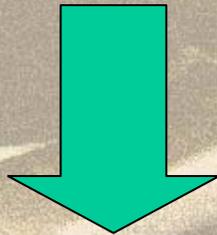
Relazione tra Statistica e Informazione

Il metodo statistico permette di:

- generare informazione, elaborando dati grezzi inizialmente privi di valore informativo
- valorizzare l'informazione esistente mediante indici sintetici
- leggere e impiegare in modo razionale le informazioni probabilistiche ottenute, ad es., tramite campionamento
- generare informazioni prospettiche (previsioni, modelli di simulazione)

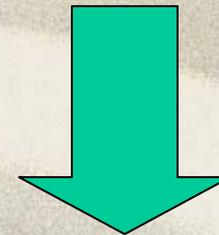
Le fonti dei dati

Dati Primari



Informazioni che devono essere raccolte per la prima volta mediante:
Osservazioni, Esperimenti e Questionari
(Metodi appositamente predisposti per risolvere un dato problema)

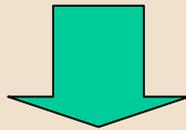
Dati Secondari



Informazioni già esistenti, utili per scopi scientifici ed Economici
Disponibili internamente o esternamente
Pregi:
· economicità
· velocità d'acquisizione
· tutela della privacy

Dati Secondari

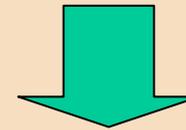
Interni



Informazioni aziendali su:

- produzione
 - costi
 - vendite
 - distribuzione
- ecc.

Esterni



Statistiche e rapporti
pubblicati da:

- enti pubblici
- associazioni commerciali di categoria
- altre organizzazioni

Le principali fonti di dati in Italia

Fonti istituzionali

- Ministero dell'Economia e delle Finanze (già del Bilancio e della Programmazione Economica): Relazione previsionale e programmatica (annuale)
- Banca d'Italia (Relazione del Governatore, annuale)

Fonti statistiche ufficiali (Istat, www.istat.it)

- Annuario Statistico Italiano (annuale)
- Annuario di Statistiche Industriali (annuale)
- Bollettino mensile di statistica (annuale)

Enti per la ricerca economica e sociale

- ISAE (già ISCO, analisi della congiuntura)
- CENSIS (Centro Studi Investimenti sociali): *Rapporto sulla situazione sociale del Paese*, col patrocinio del CNEL
- INDIS (Istituto nazionale distribuzione e servizi)
- Associazioni di categoria (Confindustria, Confcommercio,...)
- Camere di Commercio
- Banche e Industrie
- Istituti di ricerca privati
- Annuari (Kompass, Seat)
- Quotidiani e riviste (Il Sole 24 ore, Financial Times, Mondo Economico, ...)

L'impiego della statistica in azienda (Tassinari-Freo)

Area commerciale

- segmentazione di mercato
- prodotto
- prezzo
- distribuzione
- promozione

Area risorse umane

- test attitudinali e di abilità
- *performance* dei lavoratori

Area operativa

- disegno del prodotto
- pianificazione del processo
- produzione
- logistica
- aggregazione dei piani di produzione
- inventario
- linee di attesa
- controllo del processo produttivo

Area contabilità e gestione

- analisi degli indici di bilancio
- stima di funzioni di costo

Area finanziaria

- progetti di investimento
- struttura del capitale
- gestione del capitale corrente
- valutazione delle attività finanziarie
- analisi dei mercati dei capitali

Serie Storica

Una serie storica è una sequenza di osservazioni ordinate rispetto al tempo (ad esempio: il fatturato mensile, i prezzi giornalieri delle azioni, il tasso di interesse settimanale, il profitto annuo, ecc.). Lo **scopo dell'analisi** delle serie storiche consiste nello studio dell'evoluzione passata del fenomeno rispetto al tempo; la previsione viene ottenuta ipotizzando che tali regolarità di comportamento di ripetano nel futuro.

Esempio: Popolazione residente in Italia dal 1979 al 2000.

Anno	Popolazione	Anno	Popolazione
1979	56.318	1990	56.719
1980	56.434	1991	56.751
1981	56.510	1992	56.859
1982	56.544	1993	57.049
1983	56.564	1994	57.204
1984	56.577	1995	57.301
1985	56.593	1996	57.397
1986	56.596	1997	57.512
1987	56.602	1998	57.588
1988	56.629	1999	57.646
1989	56.672	2000	57.728

Dati Sezionali o Cross Section

- Dati caratterizzati da unità individuali (persone, imprese, ecc.)
- L'ordinamento dei dati non è usualmente rilevante

Esempio: *PIL regionale anno 1999 (miliardi di Lire)*

Regione	PIL	Regione	PIL
Piemonte	186328,9	Marche	55344,1
V. d'Aosta	5806,1	Lazio	216629,9
Lombardia	436875,2	Abruzzo	39631,8
Trentino A.A.	45347,1	Molise	9544,2
Veneto	195651,8	Campania	139188,6
Friuli V.G.	49461,6	Puglia	100616,6
Liguria	64664,1	Basilicata	16154,3
Emilia R.	188508,8	Calabria	47131,2
Toscana	144547,6	Sicilia	124999,4
Umbria	29973,9	Sardegna	47438,3

I rapporti statistici

I Rapporti Statistici sono degli indicatori che risultano dal rapporto di **due dati statistici**;

Permettono di **confrontare** l'intensità di un fenomeno, registrata **in luoghi o tempi diversi**.

Differenze (Variazioni) assolute

Date A e B, due grandezze omogenee espresse nella stessa unità di misura, viene definita differenza assoluta l'espressione:

$$d=A-B$$

Esempio:

In due hotel nel mese di febbraio si sono avuti rispettivamente 280 e 240 turisti.

$$d=280-240=40$$

$d=40$ esprime che la prima prima quantità è maggiore di 40 rispetto alla seconda.

Differenze (Variazioni) relative

Benché calcolabili, molto spesso le variazioni assolute possono dare origine a confronti non significativi. Si supponga il caso di voler confrontare il reddito di due Paesi di dimensioni notevolmente diverse.

$$d=A/B$$

Nel caso in cui una grandezza sia di rilievo maggiore nel confronto si è soliti utilizzare le differenze relative.

$d_1=(A-B)/A$ se la grandezza A si ritiene essere la più importante;

$d_2=(A-B)/B$ se la grandezza B si ritiene essere la più importante;

$d_3=(A-B)/(A+B)/2$ se si ritiene le due grandezze ugualmente importanti.

Tassi (Saggi)

Nel caso che si prenda in esame un fenomeno collettivo, osservato in due tempi diversi, le differenze relative riferite all'unità di tempo. Si chiamano: tassi d'incremento o di decremento e vengono espresse in percentuali. Servono quindi a misurare l'intensità della variazione nell'unità di tempo.

$$I=d/n$$

Dove n rappresenta le unità di tempo considerate

Esempio

ad esempio:

Si consideri il P.I.L. : 1994 $\rightarrow X = 1.641.105$

1991 $\rightarrow Y = 1.429.453$

Si ha che:

Inc. assoluto: $X - Y = 211.652$

Saggi relativi ai 3 anni intercorsi:

Inc. relativi: $d_1 = \frac{X - Y}{X} = \frac{211.652}{1.641.105} = 14,8\%$

$$I_1 = \frac{d_1}{n} = \frac{14,8}{3} = 4,3\%$$

$$d_2 = \frac{X - Y}{Y} = \frac{211.652}{1.429.652} = 12,9\%$$

$$I_2 = \frac{d_2}{n} = \frac{12,9}{3} = 4,9\%$$

$$d_3 = \frac{X - Y}{\frac{1}{2}(X + Y)} = \frac{211.652}{1.535.279} = 13,78\%$$

$$I_3 = \frac{d_3}{n} = \frac{13,78}{3} = 4,6\%$$

I Rapporti Statistici

Definiamo rapporto statistico un
quoziente tra due fenomeni, legati da
un nesso logico



I Rapporti di Composizione

- I rapporti di composizione sono rapporti fra le intensità di un certo fenomeno e l'intensità complessiva, cioè fra dati omogenei.
- Nel caso di distribuzione di frequenza, i rapporti di composizione coincidono con le frequenze relative.

Esempio: Rapporti di Composizione

Si consideri l'ammontare degli incassi di una determinata provincia turistica relative alle principali forme di accomodation.

Incassi	Anno 2003	Anno 2004
Hotel	35000000	41540000
Agriturismi	4250000	3598000
Guesthouses	7650000	2795000
Totale	46900000	47933000

Si calcolino i rapporti di composizione.

Esempio: rapporti di composizione

	FUMATORI	NON FUMATORI	TOTALI
MASCHI	23	20	43
FEMMINE	18	10	28
TOTALI	41	30	71

Si determinino i rapporti di composizione.

$$\frac{\text{fumatori maschi}}{\text{totale}} = \frac{23}{71} \cdot 100 = 32,39$$

	FUMATORI	NON FUMATORI	TOTALE
MASCHI	32,39%	28,17%	60,56%
FEMMINE	25,35%	14,08%	39,44%
TOTALI	57,75%	42,25%	100,00%

I Rapporti di derivazione

I rapporti di derivazione confrontano due dati statistici il primo dei quali deriva, o è causato, dal secondo.

Sono rapporti di derivazione:

- il quoziente di natalità, rapporto fra i nati in un anno e la popolazione.
- il quoziente di mortalità, rapporto fra i morti in un anno e la popolazione
- il quoziente di nuzialità, rapporto fra il numero dei matrimoni in un anno e la popolazione.
- il rapporti fra gli studenti promossi e il numero totale degli studenti.
- il rapporto fra il numero dei reati commessi e la popolazione.

I rapporti di derivazione sono molto usati per confrontare uno stesso fenomeno in tempi o in luoghi diversi.

Generalmente al numeratore viene riportato un fenomeno di flusso (fenomeni rilevati in un arco temporale) ed al denominatore fenomeni di stato (rilevati in un particolare istante).

Tipico es. è il *quoziente generico di natalità*

(dato di flusso – intervallo di tempo)

Nati nel 2002
Popolazione al 31/12/2002

(dato di stato)

Si usa moltiplicato per 1.000: indica il numero medio annuo di nati ogni 1.000 abitanti

I Rapporti di Coesistenza

I rapporti di coesistenza si calcolano fra fenomeni che, pur essendo indipendenti fra loro, coesistono insieme; possono riferirsi a uno stesso fenomeno in due luoghi diversi oppure a due fenomeni diversi nello stesso luogo.

Sono rapporti di coesistenza:

- Totale delle importazioni (riferito a un certo paese in un determinato anno);
Totale delle esportazioni
- Numero delle nascite (riferito a uno stesso luogo nello stesso periodo di tempo);
Numero delle morti
- Totale delle importazioni dell'Italia (riferito a uno stesso anno);
Totale delle importazioni della Francia

Esempio:

$$\frac{\text{Nati maschi}}{\text{Nate femmine}} = 1,06$$

Si usa leggerlo
in forma %

sono nati 106 maschi ogni 100 femmine

I numeri indici

Per studiare i cambiamenti di una grandezza in funzione di uno dei suoi valori, che si prende come riferimento, si usano *i numeri di indici*. Le osservazioni della variabile possono essere trasversali o temporali, sebbene il più frequente è che i numeri di indici si applichino all'evoluzione di serie temporale. Il valore rispetto a quello che si fa la comparazione, dare essere " normale ", nel senso di cui non dare trattare influenzato per alcuno avvenimento straordinario accaduto in codesto periodo, giacche il resto dei valori della serie si esprimono come percentuali di variazione rispetto a questo, e se è un dato atipico, le informazioni che apportano i numeri di indici può condurre ad equivoci.

La comparazione tra le grandezze si fa mediante percentuali, cioè, quello che ci interessa sapere è quanto è cresciuto (o decresciuto) il valore della variabile in due istanti di tempo, in termini sempre percentuali.

Queste percentuali facilitano la descrizione dell'evoluzione temporale della serie.

Inoltre, eliminano le unità di misura (sono adimensionali), il che permette di confrontare diverse serie in cui le variabili vengono espresse in differenti unità.

Come si Costruiscono

I numeri di indici si costruiscono attraverso un quoziente tra la grandezza ed il valore di riferimento (moltiplicato per 100, per ottenere percentuali). Così, se si ottiene un quoziente uguale a 100, la variabile non ha sperimentato nessuno cambio quantitativo tra due momenti di tempo; se il risultato della divisione è superiore a 100, la variabile è cresciuta in una percentuale uguale a quello che eccede di 100 codesto quoziente, e se il risultato è inferiore a 100, la variabile ha sperimentato una discesa nel suo valore in una percentuale uguale al differenziale a 100.

Numeri indici

semplici

complessi

Prestando attenzione al numero di variabili di cui si studia l'evoluzione, distinguiamo tra numeri di indici semplici e complessi.

Numeri Indici Semplici

I numeri di indici semplici descrivono l'evoluzione di un'unica variabile. Misurano la variazione, in percentuale, di una grandezza in funzione di uno dei suoi valori osservati che si prende come riferimento, ed è quello che denominiamo valore del periodo base.

Gli indici semplici si classificano in: *indici a base fissa* ed *indici a base mobile*.

Numeri Indici a base Fissa

Si caratterizzano perché la base o periodo di riferimento rimane fissa. Cioè, si studia l'evoluzione di tutti i valori della variabile rispetto al valore preso in un istante di tempo concreto.

Il valore dell'indice della variabile x , per il periodo t , prendendo come riferimento il valore in $t=0$, si calcola secondo la seguente formula:

$$I_t^0 = \frac{x_t}{x_0} \times 100$$

Il valore dell'indice nel periodo base è sempre uguale a 100. Invece, se il valore dell'indice è 125, diciamo che la variabile ha sperimentato un incremento del 25% rispetto al suo valore in t rispetto al periodo base, mentre se il risultato è 83 diciamo che la variabile ha subito in t un decremento del 17%.

Indici a Base Mobile

Gli indici a base mobile riflettono l'evoluzione della variabile, periodo per periodo. Perciò, in ogni periodo si prende come base il periodo anteriore. Il periodo base non è fisso, ma cambia lungo il tempo. Si calcolano come segue:

$$IM_t = \frac{x_t}{x_{t-1}} \times 100$$

Esempio

Si riportano nella tavola seguente, i dati relativi ai turisti alloggiati presso una determinata struttura alberghiera di Marina di Campo, Isola D'Elba:

<i>Anno</i>	<i>Viaggiatori (in migliaia)</i>
<i>1996</i>	<i>7252</i>
<i>1997</i>	<i>8020</i>
<i>1998</i>	<i>8981</i>
<i>1999</i>	<i>10672</i>
<i>2000</i>	<i>11168</i>
<i>2001</i>	<i>11309</i>

Si calcolino la serie di indici a base fissa, prendendo come periodo di riferimento il 1996, e gli indici a base mobile.

Cambiamento di base

Da Base Fissa a Mobile

Gli indici a base mobile si possono ottenere utilizzando i valori degli indici a base fissa, invece di calcolarli direttamente dalla serie. Ricordando che gli indici a base fissa risultano dividendo la serie originale di valori *per* una costante (il valore del variabile nel periodo di riferimento), il rapporto tra due indici consecutivi è costante, la relazione è la seguente:

$$IM_t = \frac{x_t}{x_{t-1}} \times 100 = \frac{\frac{x_t}{x_0} \times 100}{\frac{x_{t-1}}{x_0} \times 100} \times 100 = \frac{I_t^0}{I_{t-1}^0} \times 100$$

Cambiamento di base

Da Base Mobile a Fissa

Di converso è possibile ottenere indici a base fissa a partire degli indici a base mobile:

$$I_t^0 = \frac{x_t}{x_0} \times 100 = \frac{x_t}{x_{t-1}} \frac{x_{t-1}}{x_{t-2}} \dots \frac{x_2}{x_1} \frac{x_1}{x_0} \times 100$$

$$I_t^0 = \frac{IM_t}{100} \frac{IM_{t-1}}{100} \frac{IM_{t-2}}{100} \dots \frac{IM_3}{100} \frac{IM_2}{100} \times IM_1$$

I Tassi di Variazione

A partire degli indici semplici, tanto a base fissa che a base mobile, si possono calcolare i tassi di variazione della variabile. Questi tassi riflettono la percentuale di variazione dei valori della variabile in un periodo rispetto ad altro. Sebbene si possono calcolare a partire degli indici a base fissa o mobile, il suo uso è più comune rispetto a questi ultimi. È giusto segnalare che l'indice a base mobile e il tasso di crescita tra due periodi non coincidono numericamente. Il tasso di crescita si ottiene sottraendo 100 all'indice a base mobile. Per calcolare il tasso di crescita a partire dell'indice a base mobile, procediamo nel seguente modo:

$$tasso_t = \frac{x_t - x_{t-1}}{x_{t-1}} \times 100 = \frac{x_t}{x_{t-1}} \times 100 - 100 = IM_t - 100$$

Si può provare senza difficoltà che quando il valore della serie decresce tra due periodi consecutivi, l'indice è minore di 100 ed il tasso è negativo. In caso contrario, l'indice è superiore a 100 ed il tasso di crescita è positivo.

Si riportano i dati relativi al personale impiegato in un determinato complesso alberghiero di Quartu Sant'Elena, Cagliari.

<i>Anno</i>	<i>Personale totale</i>
1996	18435
1997	18844
1998	20007
1999	23060
2000	25168
2001	26632

Risultati

<i>Anno</i>	<i>Personale</i>	Indice % a base mobile	Tasso di Variazione %
1996	18435	-----	-----
1997	18844	102,2186	2,218606
1998	20007	106,1717	6,171726
1999	23060	115,2597	15,25966
2000	25168	109,1414	9,14137
2001	26632	105,8169	5,81691

Per il calcolo dei tassi a partire dagli indici a base fissa, procede di formazione analoga, solo che il tasso indica la crescita o diminuzione della variabile in ogni periodo rispetto ad uno stesso periodo base che rimane fisso.

Tasso di crescita medio

Per concludere lo studio dei tassi di variazione, si deve segnalare che per calcolare i tassi di crescita non si deve utilizzare la media aritmetica. Se abbiamo una serie temporale $y_1, y_2, y_3, \dots, y_n$, il tasso medio di crescita complessivo degli N periodi è quel tasso K che permette ottenere il valore y_t applicando successivamente una crescita del $K\%$ ad ognuno dei periodi osservati

$$y_2 = y_1 \times \frac{(100 + K)}{100}$$

$$y_3 = y_2 \times \frac{(100 + K)}{100} = y_1 \times \frac{(100 + K)}{100} \times \frac{(100 + K)}{100}$$

.

.

.

$$y_t = y_{t-1} \times \frac{(100 + K)}{100} = y_1 \times \frac{(100 + K)}{100} \times \dots \times \frac{(100 + K)}{100}$$

Isolando K nella ultima equazione otteniamo che

$$K = \left[\left(\sqrt[t-1]{\frac{y_t}{y_1}} \right) - 1 \right] \times 100$$

La tabella seguente riporta in numero di disoccupati registrato nel settore alberghiero della provincia di Malaga.

Si calcoli il tasso complessivo medio annuale.

<i>Anno</i>	<i>Disoccupati</i>
1991	12957
1992	12526
1993	12474
1994	12290
1995	11221
1996	9283
1997	8520
1998	7802
1999	6786
2000	6384
2001	6343

Il tasso di crescita medio K

È pari a

$$K = \left[\left(\sqrt[10]{\frac{6343}{12957}} \right) - 1 \right] \times 100 = -6,89371$$

Il valore -6,89371 indica che nel periodo 1991-2001, il numero di disoccupati registrato nel settore alberghiero nella provincia di Malaga si è ridotto mediamente ciascun anno del 6,89371%

Numeri Indici Complessi

I numeri di indici complessi riflettono l'evoluzione congiunta di un gruppo di variabili. Per tanto, la sua utilità consiste nel descrivere come aumentano o diminuiscono in percentuali valori di differenti grandezze o variabile considerate congiuntamente. I due metodi più utilizzati per ottenere indici complessi sono la media aggregativa semplice e la media aritmetica.

- a) *Metodo della media aggregativa semplice.* Si sommano i valori di ciascuna variabile in ogni periodo considerato e sulla colonna dei valori aggregati si calcolano gli indici semplici. Questo metodo è applicabile esclusivamente nel caso di cui analizziamo variabili espresse nella stessa unità di misura.
- b) *Metodo della media aritmetica semplice.* Consta di due fasi. Nella prima si calcolano gli indici semplici (con lo stesso periodo base) per ognuna delle variabili individualmente e, dopo, si calcola la media aritmetica degli indici per ogni periodo.

Indici dei Prezzi

Come per qualsiasi altra variabile, è possibile calcolare serie degli indici dei prezzi. I più usati sono gli indici di prezzi al consumo, che riflettono l'evoluzione dei prezzi dei prodotti consumati dalla popolazione, nell'ambito del territorio di riferimento. La loro importanza risente nel fatto che offre informazioni circa la capacità d'acquisto dei soggetti. Ogni paese determina il suo indice di prezzi di consumo proprio, in Italia l'ISTAT ottiene l'indice di prezzi al consumo (IPC), l'EUROSTAT, in Europa, dal 1996, offre l'indice di prezzi al consumo armonizzato (IPCA) per tutti i paesi dell'unione europea.

Indici dei Prezzi al Consumo

IPC l'indice di prezzi al consumo è una misura statistica dell'evoluzione dell'aggregato di prezzi dei beni e servizi (paniere di beni) che consuma la popolazione presente (non residente) in Italia. Si tratta di un indice dei prezzi complesso ponderato. IL paniere dei beni, si calcola l'IPC, si determina mediante procedimenti che tengono conto delle quantità complessive acquistate dalle famiglie e si aggiorna ogni certo tempo. I prezzi di ognuno dei prodotti che compongono il paniere si ponderano per ottenere l'indice complesso tenendo in conto la percentuale che rappresenta il suo consumo nel consumo totale delle famiglie.

Indici dei Prezzi al Consumo Armonizzato

- *IPCA l'indice di prezzi di consumo armonizzato* è un indicatore statistico: il cui obiettivo è fornire una misura comune dell'inflazione che permette, comparazioni internazionali.

Anche esso è un indice di prezzi ponderati, ed in ogni paese copre i valori che superano l'uno per mille del totale di spesa del paniere della spesa nazionale. In ogni stato di membro è stato necessario realizzare particolari aggiustamenti per consentire la confrontabilità.

Se si considera un gruppo di due o più beni, la misurazione delle variazioni del complesso di tali beni avviene mediante numeri indici complessi

Laspeyres

Paasche

Tecniche di Previsione

Introduzione

In molte situazioni che interessano l'impresa e il *management*, emerge la necessità di pianificare le azioni future.

La previsione è uno strumento importante per una pianificazione efficiente.

- Fra gli scopi per i quali lo strumento previsivo è di cruciale importanza ricordiamo.
- 1. *Utilizzo efficiente delle risorse (breve periodo)*: programmazione della produzione, organizzazione dei trasporti, del personale, ecc.. I fenomeni da prevedere sono costituiti da: livello della domanda, del materiale, del lavoro, ecc.
- 2. *Approvvigionamento delle risorse (medio periodo)*: è importante prevedere le necessità future di materiali, prodotti, ecc. perché esiste un certo intervallo di tempo per ottenerli.
- 3. *Determinazione dell'ammontare delle risorse necessarie (lungo periodo)*: tutte le organizzazioni devono determinare le risorse che saranno necessarie nel lungo termine.

Tecniche Previsive

Tecniche Quantitative

I metodi di tipo quantitativo possono essere impiegati quando:

- (i) sia disponibile una sufficiente informazione sull'evoluzione passata del fenomeno;
- (ii) tale informazione possa essere quantificata,
- (iii) si possa assumere che le caratteristiche dell'evoluzione passata continuino a sussistere nel futuro, al fine di effettuare la previsione.

Tecniche Qualitative

Si basano principalmente sui giudizi. Vengono utilizzati quando vi è una *limitata o nessuna informazione quantitativa ma esiste sufficiente informazione di tipo qualitativo*

Tecniche Quantitative

Metodi quantitativi: *è disponibile sufficiente informazione quantitativa.*

Analisi delle serie storiche (*time series*):
prevede la ripetizione, nel futuro, del sentiero storico (es. andamento delle vendite, del PIL, ecc.).

Metodi esplicativi: impiegano modelli di regressione per misurare quanto una variabile esplicativa influenza la variabile da prevedere (es. effetti sulle vendite della promozione pubblicitaria e/o del prezzo).

Tecniche Quantitative

Metodi Esplicativi

assumono che la variabile da prevedere possa essere messa in relazione con una o più variabili indipendenti o esplicative.

Ad esempio, la domanda di beni di consumo di una famiglia dipende dal reddito percepito, dall'età dei componenti, ecc. Tali tecniche di previsione impiegano i metodi di regressione e quindi la fase principale dell'analisi consiste nella specificazione e stima di un modello che mette in relazione la variabile da prevedere (variabile risposta) e le variabili esplicative.

Analisi delle Serie Storiche

Il fenomeno da prevedere viene trattato come una *scatola nera* in quanto non si cerca di individuare i fenomeni che lo possono influenzare. L'obiettivo di questo approccio consiste nell'identificazione dell'evoluzione passata del fenomeno e nella estrapolazione del *sentiero* passato per ottenere la previsione. In altre parole, il fenomeno da prevedere viene modellato rispetto al tempo e non rispetto ad una variabile esplicativa.

Quando utilizzare l'Analisi delle Serie Storiche?

1. il fenomeno è poco conosciuto oppure è difficile individuare (misurare) le relazioni che questo ha con altri fenomeni;
2. siamo interessati a conoscere *ciò che accadrà* e non *il modo in cui accadrà*.

Fasi di un'analisi di previsione

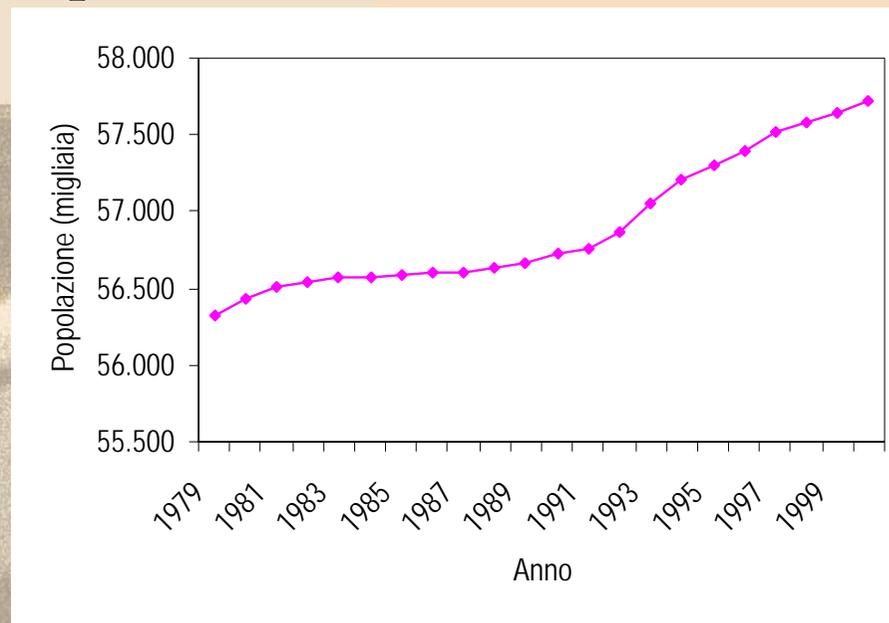
1. Definizione del problema
2. Raccolta di informazioni
3. Analisi preliminare dei dati
4. Scelta e adattamento del modello
5. Valutazione del modello e suo impiego a fini previsivi.

Analisi Grafiche Preliminari

La prima cosa importante da fare quando ci accingiamo ad analizzare una serie storica è quello di visualizzare i dati mediante una rappresentazione grafica. Lo scopo è quello di individuare eventuali regolarità di comportamento che sono utili nel suggerire l'approccio modellistico.

Il grafico più semplice è il cosiddetto time plot (oppure line plot) che consiste nella rappresentazione dei dati rispetto al tempo.

Time Plot Popolazione residente in Italia dal 1979 al 2000.



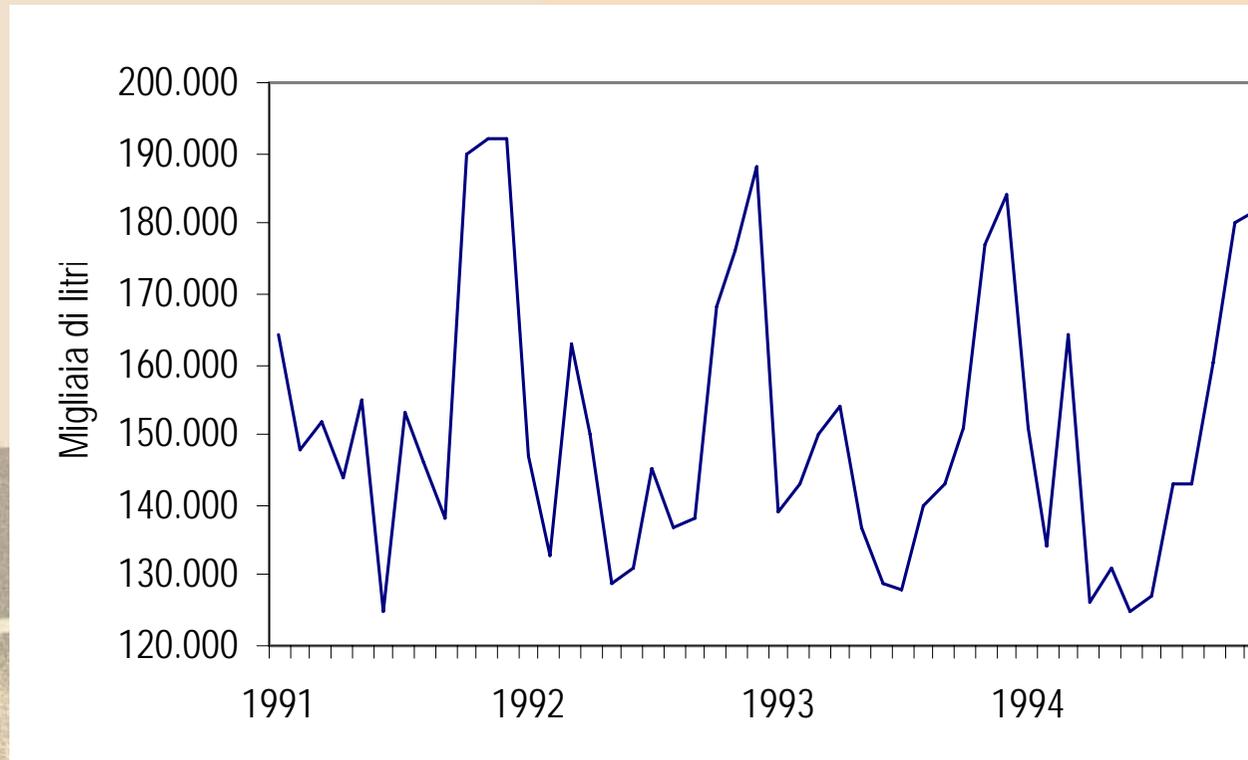
Si evidenzia un andamento sistematicamente crescente nel lungo periodo (*trend* crescente)

Esempio

Vendite mensili di birra in Australia, per il periodo 1991-1994

Anno	Mese	Quantità	Anno	Mese	Quantità
1991	Gennaio	164.000	1993	Gennaio	139.000
	Febbraio	148.000		Febbraio	143.000
	Marzo	152.000		Marzo	150.000
	Aprile	144.000		Aprile	154.000
	Maggio	155.000		Maggio	137.000
	Giugno	125.000		Giugno	129.000
	Luglio	153.000		Luglio	128.000
	Agosto	146.000		Agosto	140.000
	Settembre	138.000		Settembre	143.000
	Ottobre	190.000		Ottobre	151.000
	Novembre	192.000		Novembre	177.000
	Dicembre	192.000		Dicembre	184.000
1992	Gennaio	147.000	1994	Gennaio	151.000
	Febbraio	133.000		Febbraio	134.000
	Marzo	163.000		Marzo	164.000
	Aprile	150.000		Aprile	126.000
	Maggio	129.000		Maggio	131.000
	Giugno	131.000		Giugno	125.000
	Luglio	145.000		Luglio	127.000
	Agosto	137.000		Agosto	143.000
	Settembre	138.000		Settembre	143.000
	Ottobre	168.000		Ottobre	160.000
	Novembre	176.000		Novembre	180.000
	Dicembre	188.000		Dicembre	182.000

Time Plot Vendite mensili di birra in Australia



Si evidenzia un andamento meno *liscio*; essendo i dati a cadenza mensile, in essi è presente il fenomeno denominato *stagionalità*. Si può notare che i picchi elevati si registrano sempre nei mesi caldi (novembre e dicembre) e vicini alle feste natalizie.

Andamenti Tipici

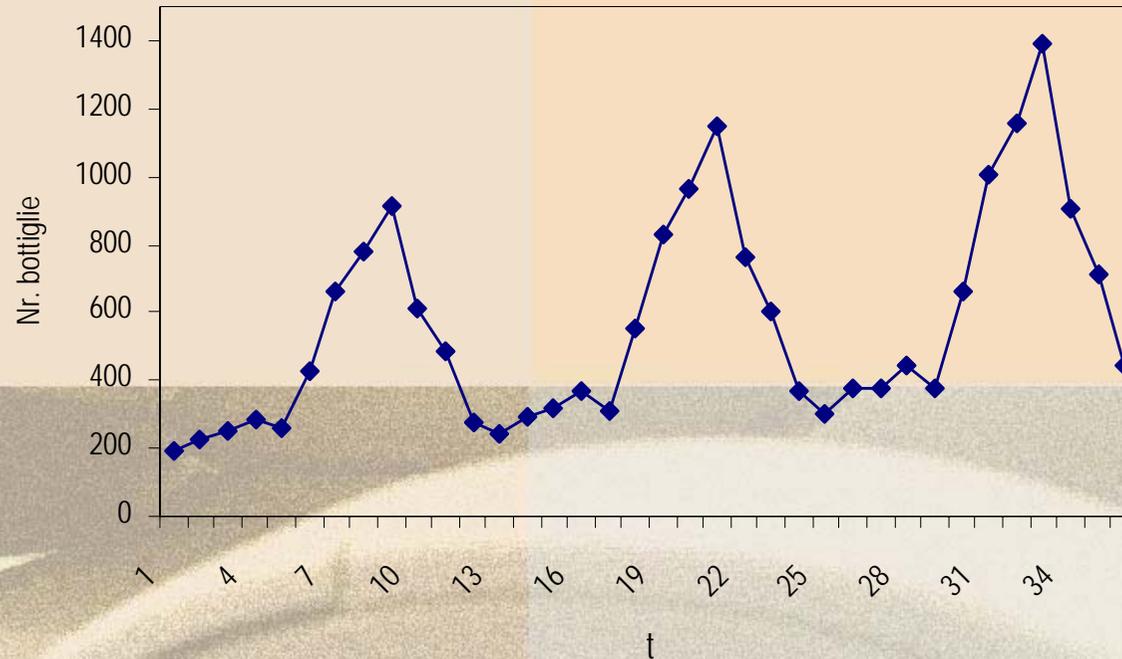
- 1. *Pattern orizzontale*. In questo caso la serie oscilla intorno ad un valore costante (media della serie). Tale serie è detta *stazionaria in media*.
- 2. *Pattern stagionale*. Questo esiste quando la serie è influenzata da fattori stagionali (es. mensile, semestrale, trimestrale, ecc.). Prodotti come gelati, bibite analcoliche, consumo di elettricità sono soggette al fenomeno stagionale. Le serie influenzate dalla stagionalità sono dette anche *serie periodiche* poiché il ciclo stagionale *si ripete in un periodo fisso*. Nei dati di tipo annuale la stagionalità non è presente (v. Fig. 2.1).
- 3. *Pattern ciclico*. Questo tipo di andamento è presente quando la serie presenta aumenti e diminuzioni che *non sono di periodo fisso*. Questa è la principale differenza fra le fluttuazioni cicliche e quelle stagionali. Inoltre, l'ampiezza delle oscillazioni cicliche è generalmente più grande di quella dovuta alla stagionalità. Nelle serie economiche il *pattern* ciclico è determinato dalle espansioni e contrazioni dell'economia dovuti a fenomeni congiunturali.
- 4. *Trend o tendenza di fondo*. E' caratterizzato da un andamento crescente o decrescente di lungo periodo. La serie della popolazione residente in Italia è un esempio di andamento tendenziale o *trend* di tipo crescente; la serie delle vendite mensili di birra, invece, non presenta alcun trend. Ha un *pattern* di fondo di tipo orizzontale.

Molte serie evidenziano una combinazione di questi *pattern*.

Esempio: Vendite mensili di bottiglie di bibita XXX (da ½ litro)

Anno	Mese	t	Nr.	Anno	Mese	t	Nr.	Anno	Mese	t	Nr.
1999	1	1	189	2000	1	13	244	2001	1	25	298
1999	2	2	229	2000	2	14	296	2001	2	26	378
1999	3	3	249	2000	3	15	319	2001	3	27	373
1999	4	4	289	2000	4	16	370	2001	4	28	443
1999	5	5	260	2000	5	17	313	2001	5	29	374
1999	6	6	431	2000	6	18	556	2001	6	30	660
1999	7	7	660	2000	7	19	831	2001	7	31	1004
1999	8	8	777	2000	8	20	960	2001	8	32	1153
1999	9	9	915	2000	9	21	1152	2001	9	33	1388
1999	10	10	613	2000	10	22	759	2001	10	34	904
1999	11	11	485	2000	11	23	607	2001	11	35	715
1999	12	12	277	2000	12	24	371	2001	12	36	441

Time plot Vendite mensili di bottiglie di bibita

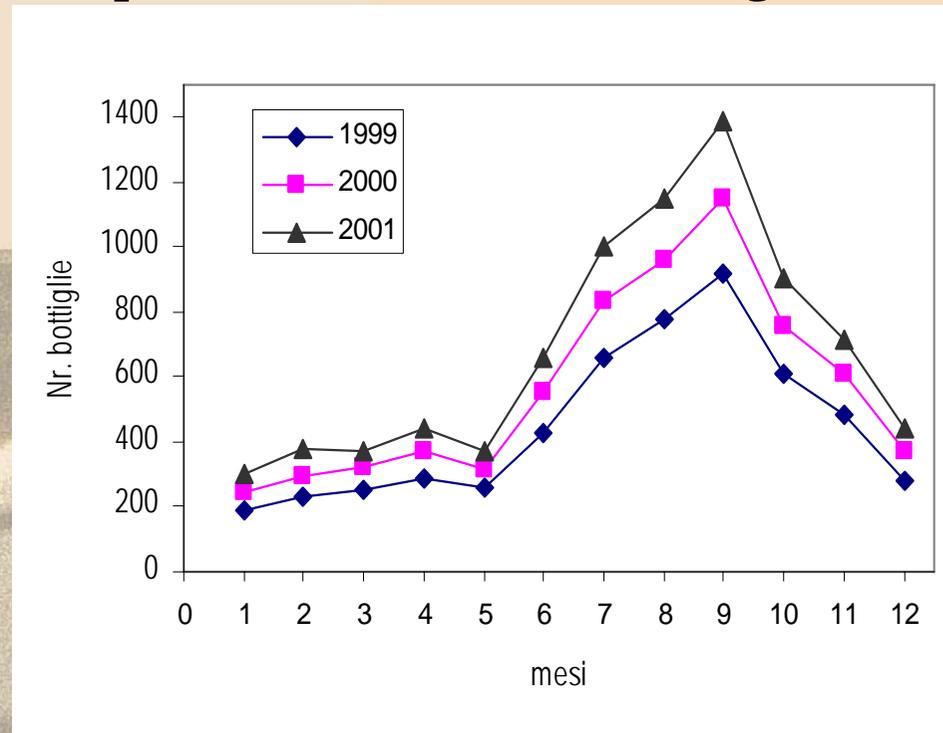


E' presente sia la componente Trend che la Stagionalità

Seasonal Plot

Se una serie storica esibisce un'oscillazione stagionale è utile eseguire un *seasonal plot* o grafico stagionale, che consiste nella rappresentazione dei valori della serie (in ordinata) *versus* i periodi stagionali come: mesi, trimestri, semestri, ecc. (in ascissa).

Esempio: Vendite mensili di bottiglie di bibita

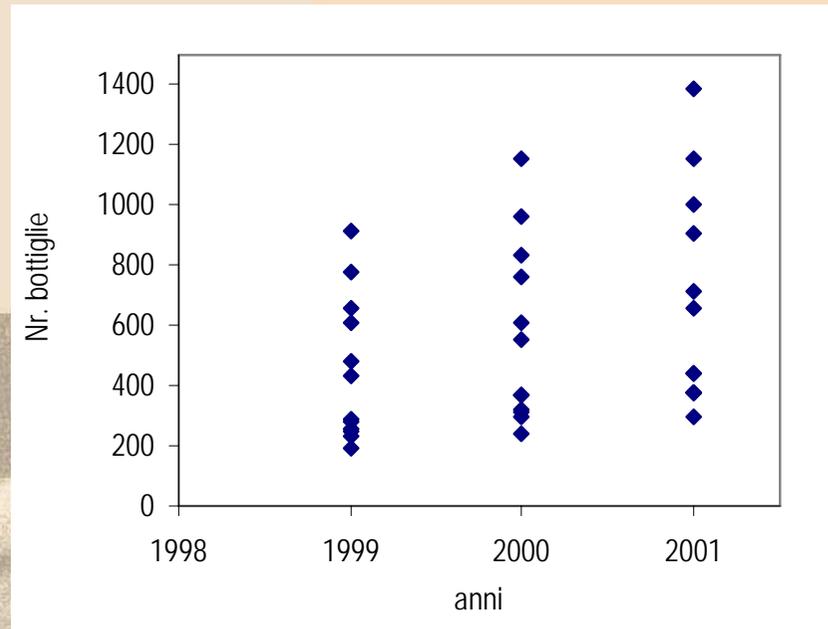


Si evidenzia la presenza di oscillazioni stagionali in espansione durante i mesi estivi con inizio da giugno, in contrazione a partire da ottobre. Il *seasonal plot* è in grado di mostrare anche l'eventuale presenza del *trend*

Grafico a punti Valori/Anni

I valori della serie (in ordinata) sono rappresentati rispetto ai periodi annuali (in ascissa).

Esempio: Vendite mensili di bottiglie di bibita



Si Evidenzia:

- la presenza del *trend*; si nota come le tre 'colonne' di punti tendono, con gli anni, ad essere posizionate leggermente più in alto;
- il tipo di oscillazione stagionale interna all'anno; si nota come, all'aumentare del *trend*, il *range* di variazione dei valori stagionali è più ampio.

Sintesi Numeriche

In aggiunta ai grafici, è utile approfondire l'analisi statistica, calcolando indici sintetici quali: media, mediana, campo di variazione, varianza, deviazione standard (*standard deviation*), ecc.

Indici di posizione e di variabilità frequentemente usati

Serie storica: $y_1, y_2, \dots, y_t, \dots, y_n$ $t=1, \dots, n$

Media aritmetica $\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t$

Campo di variazione (*range*) $R = \text{Max}(y_t) - \text{min}(y_t)$

Varianza $S^2 = \frac{1}{n-1} \sum_{t=1}^n (y_t - \bar{y})^2$

Deviazione standard $S = +\sqrt{S^2} = +\sqrt{\frac{1}{n-1} \sum_{t=1}^n (y_t - \bar{y})^2}$

Sintesi numeriche proprie delle Serie Storiche

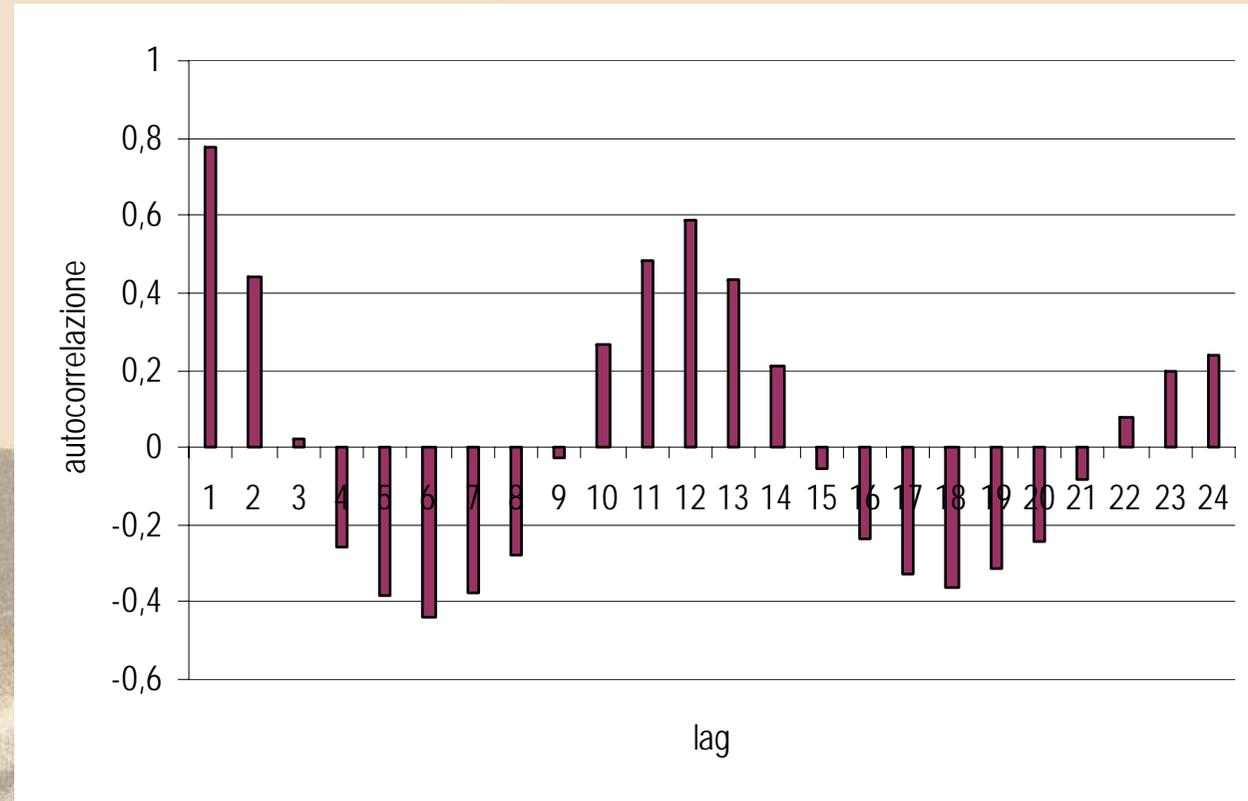
Autocovarianza e autocorrelazione

Autocovarianza
$$c_k = \frac{1}{n-1} \sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})$$

Autocorrelazione
$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} = \frac{c_k}{c_0}$$

Correlogramma

Vendite mensili di bottiglie di bibita



Dal correlogramma possiamo dedurre la ciclicità del legame di autocorrelazione dovuto al marcato fenomeno stagionale che influenza i dati. Tutto ciò è una conferma della presenza di una marcata stagionalità di ampiezza 12 periodi (mesi).

Misure di bontà di adattamento e di accuratezza della previsione

In molte analisi statistiche in cui un modello viene adattato a un set di dati (ad es. nell'analisi di regressione), il termine accuratezza si riferisce alla capacità del modello di riprodurre i dati sui quali è stato stimato e cioè della bontà di adattamento del modello (*goodness of fit*).

L'accuratezza nella previsione (*forecast accuracy*), invece, misura la capacità del modello a riprodurre i dati futuri della serie.

Accuratezza della Previsione

Se il metodo di previsione si propone un orizzonte temporale di più periodi, la verifica della sua capacità previsiva può avvenire come di seguito indicato:

- 1) si utilizzano i primi m dati della serie per la stima del modello;
- 2) si usano i successivi $m+1, \dots, n$ dati per la verifica dell'accuratezza previsiva.

Per capire quanto appena detto introduciamo un po' di simbologia:

- y_1, y_2, \dots, y_n indica la serie di dati disponibile;
- y_1, y_2, \dots, y_m $m < n$, indica la serie di dati che viene usata per la stima del modello di previsione (*training sample*);
- $y_{m+1}, y_{m+2}, \dots, y_n$, indica la serie di dati che viene usata per la verifica della capacità previsiva (*test sample*);
- $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m$ sono le stime dei valori del *training sample* ottenute tramite il modello stimato;
- $F_{m+1}, F_{m+2}, \dots, F_n$ sono le previsioni (*forecast*) riferite al periodo di tempo da $t=m+1$ a $t=n$ (*test sample*), ottenuto tramite il modello stimato sul *training sample*.

La differenza fra l'errore di stima e_t e l'errore di previsione f_t , può essere apprezzata dalle formule seguenti:

$$e_t = y_t - \hat{y}_t, t=1, \dots, m;$$

$$f_t = y_t - F_t, \quad t=m+1, \dots, n.$$

Misure della *Bontà* di *Adattamento/Previsione*

- errore medio (*mean error*: ME): media aritmetica degli errori;
- errore quadratico medio (*mean squared error*: MSE): media aritmetica dei quadrati degli errori;
- errore medio assoluto (*mean absolute error*: MAE): media aritmetica degli errori presi in valore assoluto;
- errore medio assoluto percentuale (*mean absolute percentage error*: MAPE): media aritmetica degli errori relativi, presi in valore assoluto e moltiplicati per 100.

Misure della *Bontà di Adattamento/Previsione*

Bontà di adattamento	Accuratezza della previsione
$ME = \frac{1}{m} \sum_{t=1}^m e_t$	$ME = \frac{1}{m} \sum_{t=1}^m f_t$
$MSE = \frac{1}{m} \sum_{t=1}^m e_t^2$	$MSE = \frac{1}{m} \sum_{t=1}^m f_t^2$
$MAE = \frac{1}{m} \sum_{t=1}^m e_t $	$MAE = \frac{1}{m} \sum_{t=1}^m f_t $
$MAPE = \frac{1}{m} \sum_{t=1}^m \frac{ e_t }{y_t} \cdot 100$	$MAPE = \frac{1}{m} \sum_{t=1}^m \frac{ f_t }{y_t} \cdot 100$

Misure della *Bontà di Adattamento/Previsione*

- ME è l'unico indice che può assumere anche valori negativi. Il ME avrà valori bassi perché elementi positivi e negativi, nella sommatoria, tenderanno a compensarsi. Il segno di ME ci dice se si tende, in media, a sopravvalutare ($ME < 0$) o sottovalutare ($ME > 0$) il fenomeno.
- MSE e MAE rendono positivi i singoli addendi della sommatoria. Per la presenza dell'esponente, MSE è meno agevole da interpretare da persone non specialiste.
- I tre indici ME, MSE e MAE forniscono un valore che dipende dall'unità di misura della serie. Il MAPE elimina questo problema in quanto l'errore viene relativizzato dividendolo per il valore osservato. Il MAPE può però essere usato solo se il fenomeno è misurabile su scala a rapporto. Per il calcolo del MAPE sorgono difficoltà quando la serie osservata contiene valori nulli o molto prossimi a zero.

Esempio metodo di previsione

- . I dati fino a $t=24$ (*training sample*) della vendita di birre, sono utilizzati per la stima del modello di previsione; i dati del 2001 (*test sample*) sono impiegati per la verifica della capacità previsiva del metodo che consiste nel prevedere il dato del mese j ($j=1, \dots, 12$) del 2001, mediante la media aritmetica semplice dei valori dello stesso mese j nei due anni precedenti.

Esempio: Vendite mensili di bottiglie di bibita

Mese	y_t	F_t	f_t	
1	298	217	82	
2	378	263	116	
3	373	284	89	
4	443	330	114	$ME=171,2$
5	374	287	88	$MSE=36614,2$
6	660	494	167	$MAE=171,2$
7	1004	746	259	$MAPE=25,5\%$
8	1153	869	285	
9	1388	1034	355	
10	904	686	218	
11	715	546	169	
12	441	324	117	

In questo caso $ME=MAE$ poiché tutti gli errori di previsioni sono positivi. Questi indici possono presentare delle difficoltà interpretative innanzi tutto perché non abbiamo un valore di riferimento per decidere se il metodo di previsione è soddisfacente o no. Non c'è dubbio qui che un MAPE pari a 25% è inaccettabile.

Metodo Naive

Un procedimento molto semplice consiste nel confrontare i valori ottenuti col metodo M1 con quelli derivanti dall'impiego di un metodo cosiddetto *naive*. Un metodo *naive* è, ad esempio, il seguente: prendere come previsione per il periodo t il dato del periodo $t-1$. La previsione per il mese di gennaio 2001 è uguale al valore della serie in corrispondenza di dicembre 2000, e così via.

Mese	y_t	F_t	f_t	
1	298	371	-73	
2	378	298	80	
3	373	378	-5	
4	443	373	70	$ME=5,8$
5	374	443	-69	$MSE=53668,8$
6	660	374	286	$MAE=188,2$
7	1004	660	344	$MAPE=27,6\%$
8	1153	1004	149	
9	1388	1153	235	
10	904	1388	-484	
11	715	904	-189	
12	441	715	-274	

Problemi di interpretazione delle misure di *Adattamento/Previsione*

- Un problema interpretativo concerne il fatto che questi indici esprimono funzioni di *distanza*, fra valori osservati e previsti, che sono molto diverse fra loro e che possono essere anche discordanti (nell'esempio sopra riportato abbiamo visto un ME più basso per il metodo *naive* rispetto a M1, e valori più elevati per gli altri indici). Occorre quindi decidere in via preliminare l'indice da usare per misurare l'accuratezza della previsione.
- Al fine di valutare le prestazioni del metodo di previsione, è buona regola condurre anche analisi grafiche degli errori, per esaminare il loro andamento rispetto al tempo.