

La previsione della domanda

Introduzione

La domanda di un particolare articolo, o famiglia di articoli, può essere modellizzata come una serie storica di particolari valori. Chiaramente ai fini della pianificazione e controllo delle operazioni è necessario avere una previsione della domanda nei periodi futuri. La previsione della domanda può essere definita come **«il tentativo di determinare oggettivamente la natura e l'entità di ciascuna richiesta che un'azienda può ragionevolmente attendersi in un prefissato orizzonte di tempo»** [1]. In particolare, se il tempo di approvvigionamento dei materiali unito a quello necessario per la fabbricazione e per la consegna risulta essere maggiore di quello che il consumatore è disposto ad aspettare, allora la previsione della domanda diviene essenziale.

La previsione è fondamentale ogni qual volta è necessario prendere una decisione inerente ad attività future: per esempio per pianificare l'investimento totale in scorte, per prevenire il bisogno di capacità produttiva addizionale, per scegliere tra diverse strategie gestionali, etc. Dopo che sono state prese tali decisioni, solo una cosa è certa: *la previsione sarà sbagliata*. Le procedure di pianificazione e controllo devono, perciò, tenere conto della presenza di questo errore, anche perché il costo di questi errori può essere molto alto.

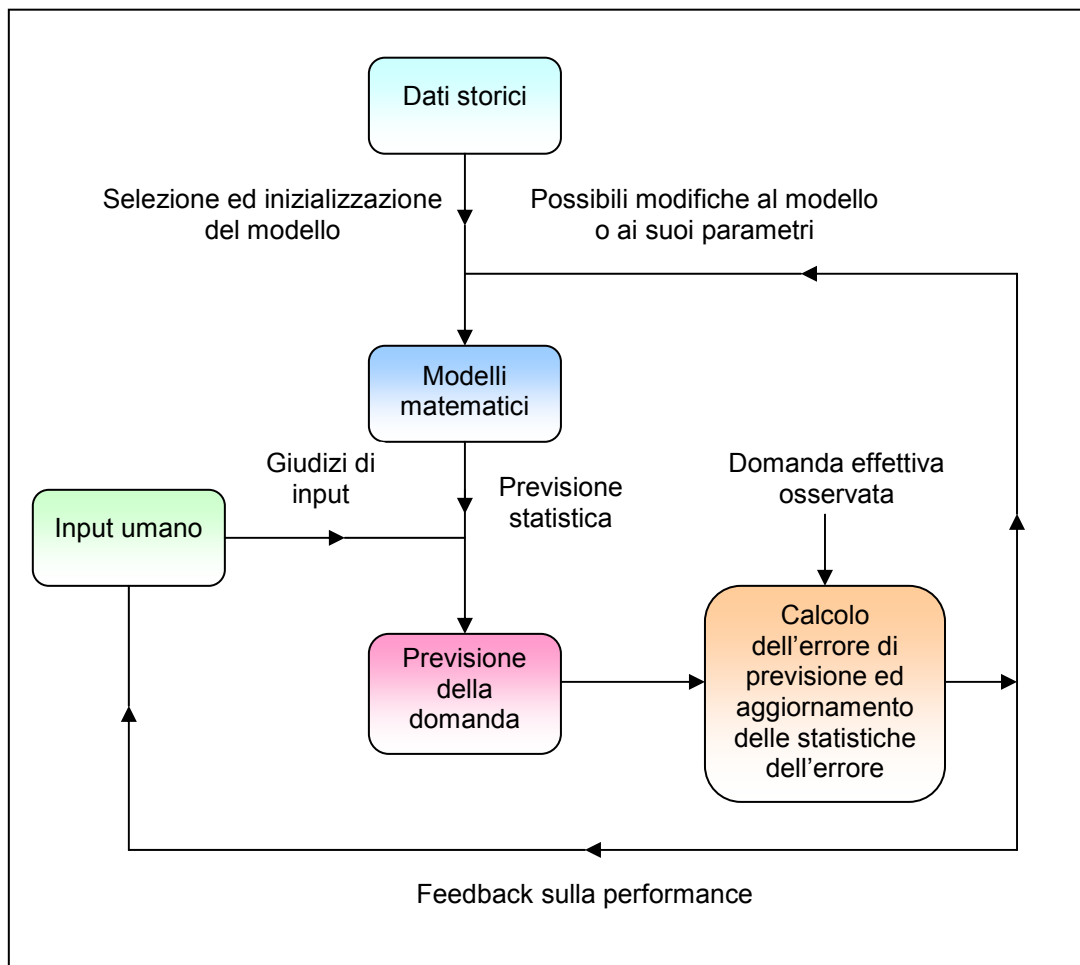
La previsione in genere si basa su di una combinazione tra un'estrapolazione di ciò che si è osservato in passato (chiamata *previsione statistica*) e giudizi "informati" sugli eventi futuri. I giudizi informati possono derivare dalla conoscenza di ordini futuri da parte di clienti esterni, delle condizioni economiche generali, etc., ma anche da giudizi di marketing come l'effetto di promozioni, di sconti, di campagne pubblicitarie, etc.

La struttura complessiva di un generico sistema di previsione è mostrata nella figura seguente.

Il giudizio umano è un fattore cruciale. Inoltre è da notare che la domanda effettiva in un dato periodo è confrontata con la previsione iniziale in maniera tale da poterne misurare l'errore. Monitorare questi errori è importante per almeno tre ragioni:

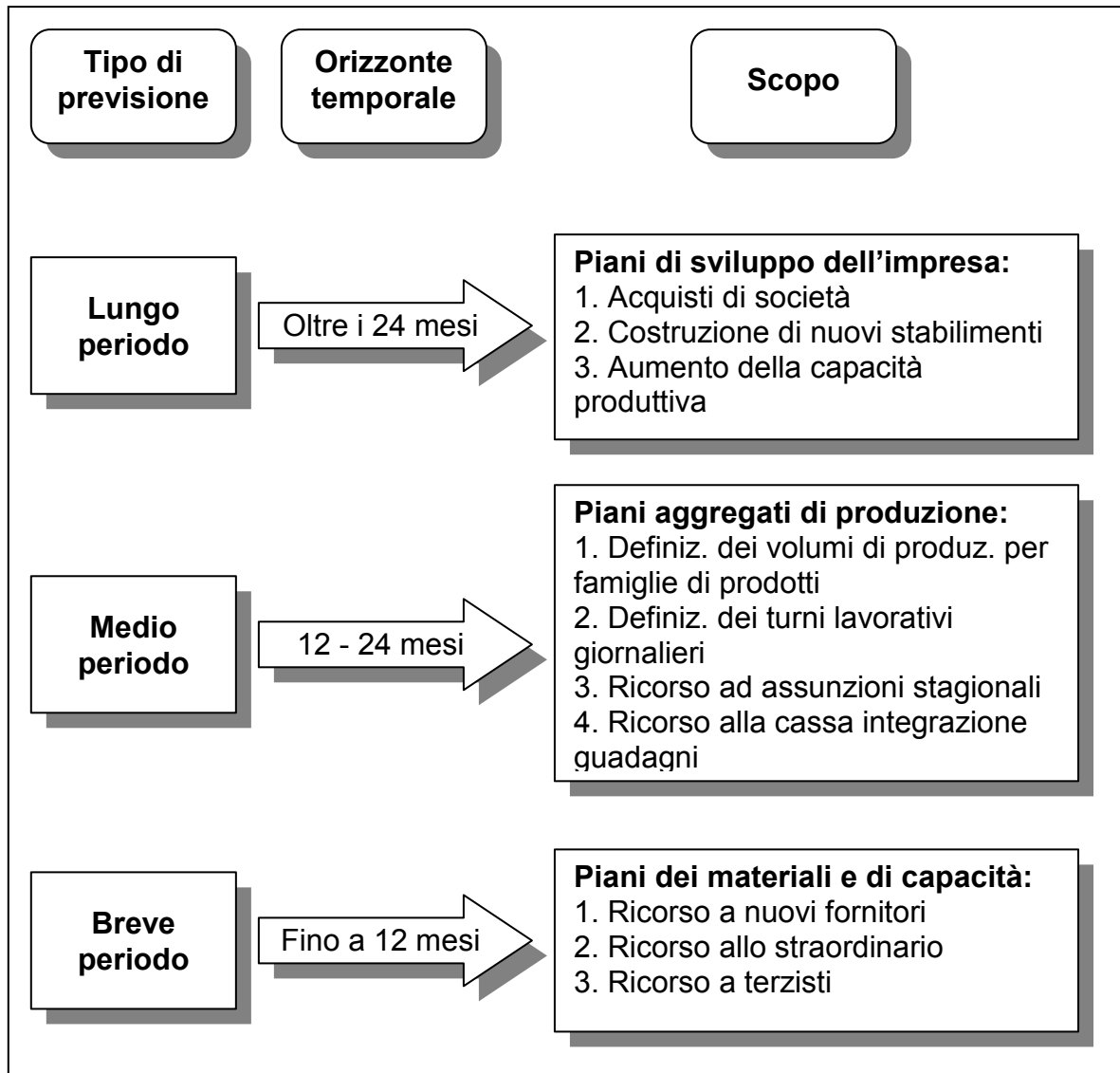
1. La quantità di scorte di sicurezza dipende dal livello di servizio al cliente che, a sua volta, dipende dall'entità dell'errore di previsione;

2. La previsione statistica si basa su specifici modelli caratterizzati da opportuni parametri. Il valore ed il segno dell'errore (positivo o negativo) può suggerire delle modifiche sia ai valori di questi parametri, sia addirittura alla forma del modello stesso, per esempio introducendo una componente stagionale che prima non era stata prevista;
3. L'errore può servire a monitorare l'influenza degli input soggettivi del modello previsionale.



È possibile classificare le diverse tipologie di previsione in base all'orizzonte temporale coperto [2]. Si possono definire tre orizzonti temporali. Le previsioni di lungo periodo (oltre i 24 mesi) vengono formulate come supporto alle decisioni manageriali relative ai piani di sviluppo dell'impresa: acquisti di società, costruzione di nuovi stabilimenti, aumento della capacità produttiva, ecc. Le previsioni di medio periodo (da 12 a 24 mesi) vengono formulate, invece, per prendere decisioni relative ai piani aggregati di produzione:

definizione dei volumi di produzione per famiglie di prodotti, definizione dei turni lavorativi giornalieri, ricorso ad assunzione stagionali, ricorso alla cassa integrazione guadagni, ecc. Sul breve periodo (fino a 12 mesi) le decisioni sono relative ai piani dei materiali e di capacità: ricorso a nuovi fornitori, allo straordinario, a terzisti.

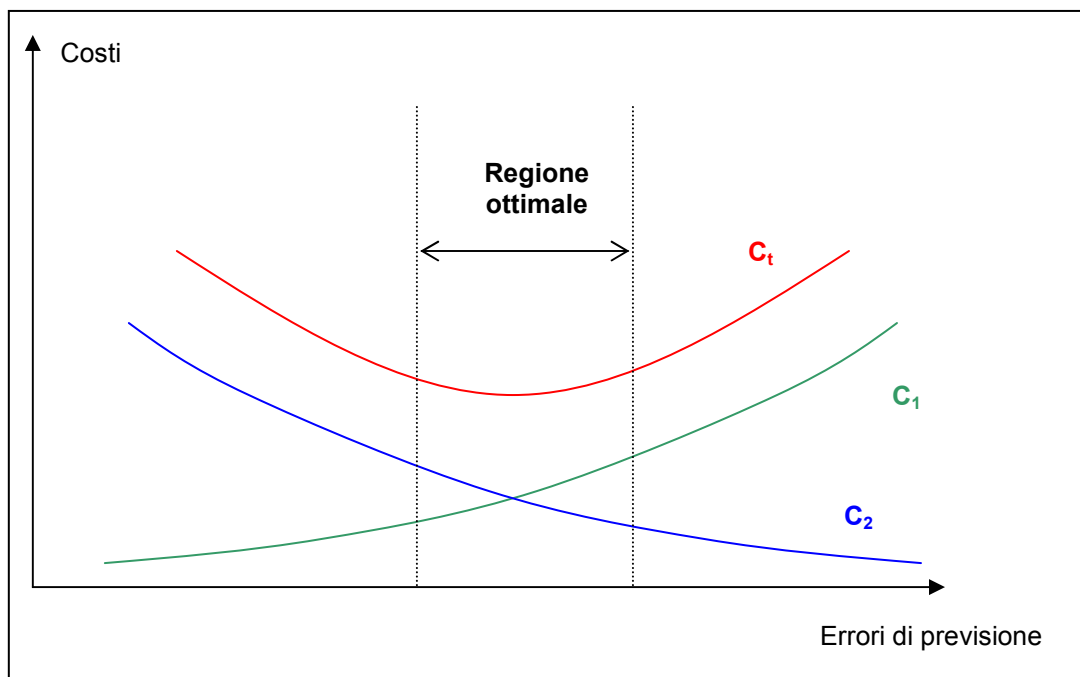


Le previsioni di lungo, medio e breve periodo si differenziano per il crescente grado di aggregazione degli oggetti della previsione. Sul lungo termine, oggetto di previsione sono i volumi complessivi di vendita, sul medio termine le famiglie di prodotti, mentre, sul breve termine i prodotti finiti, i moduli e i sottoassiemi.

Bisogna ora fare una puntualizzazione [6]: nel seguito la domanda verrà sempre trattata come una variabile *esogena*, cioè indipendente dalle azioni intraprese dal management. In

realtà, soprattutto nel lungo periodo, l'abilità nel soddisfare le richieste del cliente sicuramente influenzerà la domanda dei prodotti. Al momento non è possibile sviluppare un modello che esplicitamente tenga conto di questi due tipi di interazione, in particolar modo per quanto riguarda i singoli articoli. Per questo motivo la domanda verrà trattata generalmente come una variabile esogena, ma in alcuni casi si cercherà di incorporare gli effetti delle azioni intraprese dal management, come sconti, promozioni, etc.

Lo scopo della previsione è quello di ridurre il rischio nel processo decisionale in relazione al metodo previsionale adottato; è evidente che destinando maggiori risorse all'attività previsionale dovremmo essere in grado di migliorare l'accuratezza dei risultati riducendo gli errori legati all'incertezza insita nella stima.



Quindi è bene tener conto dei costi legati alla previsione della domanda; in linea di principio nello scegliere tra le varie procedure previsionali, la linea di condotta più corretta è quella di cercare di mantenere i costi totali attesi durante l'orizzonte decisionale il più basso possibile. I suddetti costi dovrebbero comprendere sia i costi di ottenimento della previsione che i costi derivanti dall'errore di previsione:

$$C_t = C_1 + C_2;$$

in cui:

C_t = costi totali relativi alla procedura previsionale;

C_1 = costi di ottenimento della previsione;

C_2 = costi derivanti dall'errore di previsione.

In pratica l'applicazione di questo criterio apparentemente semplice risulta essere piuttosto limitata in quanto è molto difficile misurare il costo dovuto all'errore di previsione. Comunque ciò non significa che almeno non si debba provare a trovare il punto di ottimo mostrato in figura.

Classificazione dei metodi previsionali

Con riguardo alla natura dei fenomeni, le previsioni possono riguardare i dati originari, i valori aggregati nello spazio e/o nel tempo di dati elementari, i tassi di variazione, il rapporto tra variabili.

Circa i differenti approcci statistici per la previsione, si possono distinguere i metodi *univariati* e *multivariati* (a seconda del numero di variabili utilizzate), metodi *statistici* ed *econometrici* (se il modello viene costruito dai dati, ovvero viene dedotto da ipotesi economiche), metodi *estrapolativi* e metodi *strutturali* (a seconda che l'enfasi della modellistica venga posta sulla capacità previsiva o interpretativa), metodi *deterministici* e metodi *stocastici* (a seconda che l'adattamento del modello per l'uso previsivo si basi su una funzione matematica o su un processo stocastico e così via).

Le precedenti classificazioni, come vedremo, si sovrappongono per più aspetti perché uno stesso modello può essere esaminato secondo ottiche diverse. Per esempio, una funzione di regressione è di per sé un metodo multivariato, ma se contiene solo un trend polinomiale rappresenta un approccio deterministico alle previsioni.

Le previsioni possono essere di diverso tipo (economico-general, sociali, tecnologiche, di mercato ecc.) e compongono il quadro delle ipotesi entro cui si formulano gli obiettivi e le politiche aziendali. Ci occuperemo specificamente delle previsioni di vendita, cioè delle stime (esprese in unità fisiche, monetarie o convenzionali) delle quantità di prodotto collocabili sul mercato in un periodo di tempo futuro. Con esse l'impresa intende conoscere in anticipo quale potrà essere l'assorbimento dei suoi prodotti, in modo da programmare gli investimenti, i cicli di lavorazione, l'approvvigionamento delle risorse, l'attività di distribuzione commerciale.

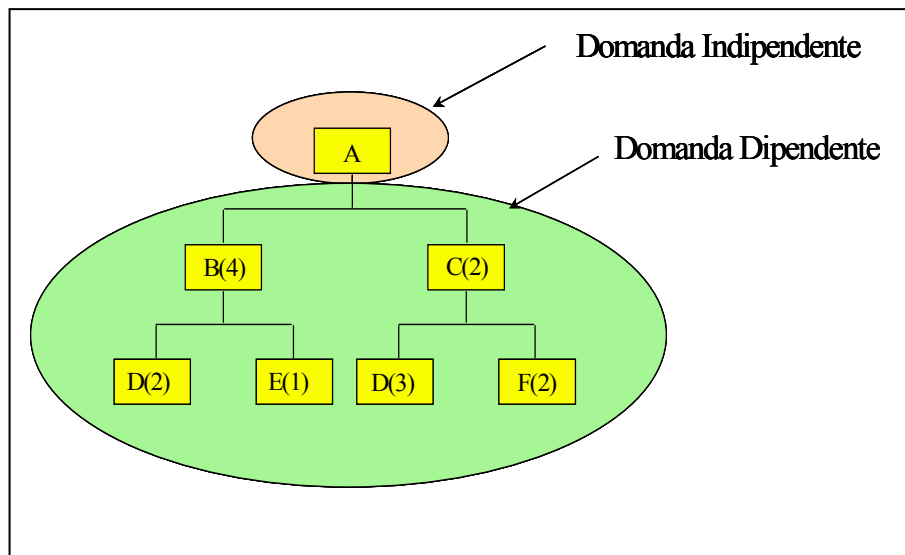
La previsione delle vendite è generalmente frutto di un processo di approssimazioni successive, che si sviluppa in tre fasi fondamentali:

- **Analisi della domanda** cioè valutazioni delle tendenze espansive o recessive dei consumi e quantificazione della domanda globale (previsioni del mercato);
- **Determinazione delle quota aziendale di vendita**, cioè individuazione di quella che, in base alle scelte correnti di marketing (prodotti, prezzi, promozione ecc.) potrà essere la fetta del mercato globale che l'impresa è in grado di soddisfare (previsione delle vendite);
- **Definizione**, in base alle nuove scelte di marketing, del volume di vendita effettivamente raggiungibile (obiettivo di vendita).

Alla previsione delle vendite si può dunque giungere per due vie e sulla base di due ipotesi alternative. La prima via è quella più lunga, che parte dalla stima della domanda globale e che, attraverso la definizione della quota di mercato spettante all'impresa, perviene a quantificare il volume.

Il materiale che si gestisce in un'azienda, qualunque sia la tipologia del prodotto, dalle materie prime al semilavorato, al componente generico, può essere essenzialmente classificato in domanda dipendente e indipendente.

I materiali a **domanda dipendente** sono quelli che rientrano nella distinta base di un qualsiasi prodotto, dipendono, appunto, dal piano generale di produzione che è stato formulato, ovvero la domanda di tali prodotti *dipende* dalla domanda di altri beni; tutti gli altri sono a **domanda indipendente** (olio da taglio, lubrificante carta refrigerante, etc...)



La distinzione fondamentale sta nel fatto che per i materiali a domanda dipendente, la gestione avviene in un ambiente deterministico (per esempio in sede di planning si decide il mixing produttivo di certi prodotti A,B,C,D..., e si utilizzano tecniche specifiche come la programmazione lineare), per questi materiali l'aleatorietà si concentra esclusivamente nella previsione della domanda del prodotto finito. Infatti quegli articoli che vengono utilizzati come componenti del prodotto finito non necessitano di tale metodo in quanto il loro fabbisogno può essere ricavato deterministicamente una volta nota la domanda dei beni di cui fanno parte.

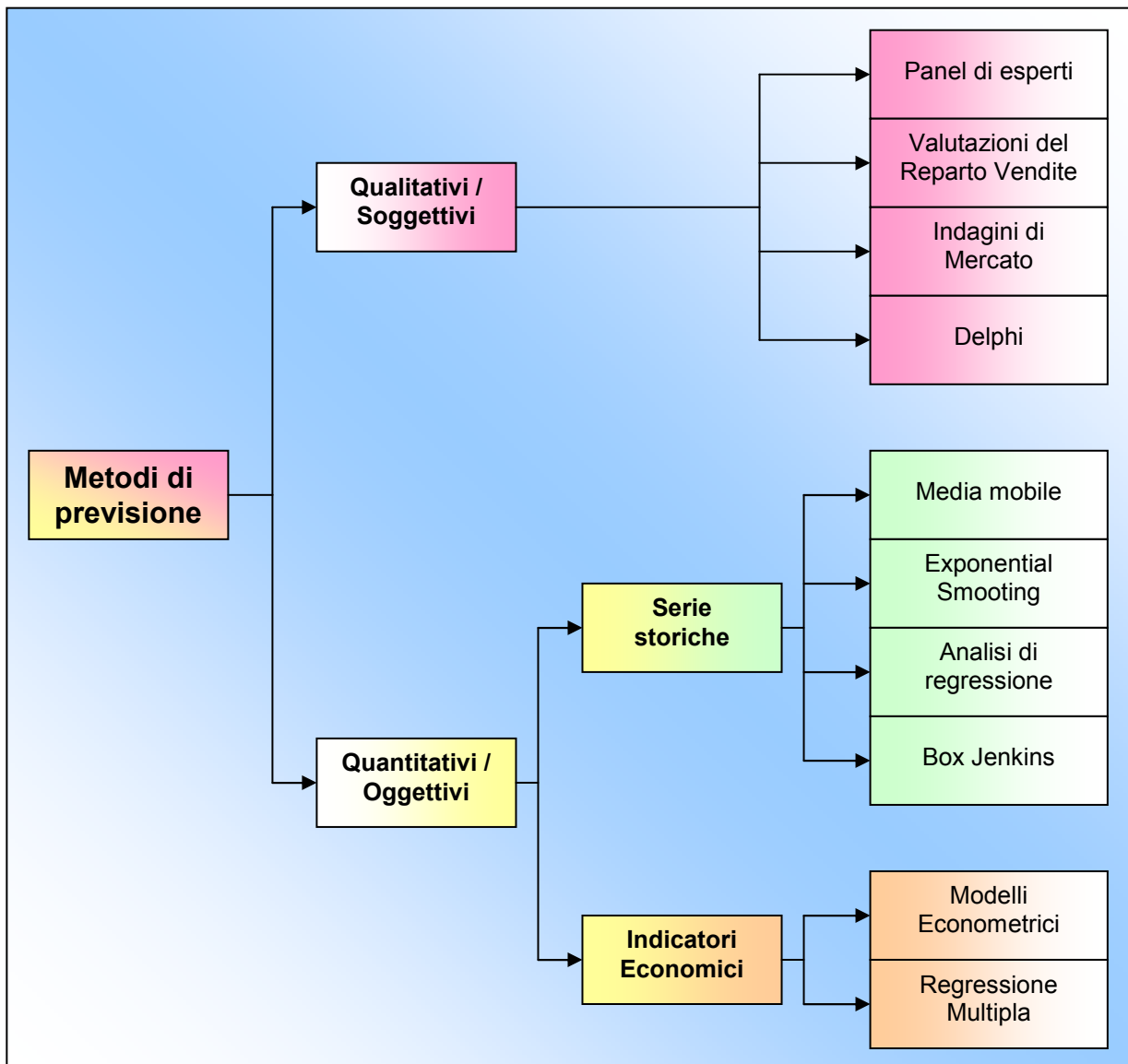
Per i materiali a domanda indipendente, invece, non si conosce la domanda (almeno in termini deterministici), tali materiali sono inoltre "indipendenti" dal piano generale di produzione.

La scelta dei più appropriati metodi di previsioni è influenzata da una serie di fattori:

- Tipo di previsione richiesta
- Orizzonte di previsione
- Dati in possesso
- Accuratezza richiesta
- Comportamento del processo di previsione
- Costi di sviluppo e installazione
- Facilità operativa
- Comprensione e cooperazione del management

In particolare considereremo i seguenti metodi di previsione che possono essere classificati

in *soggettivi* (o *qualitativi*) e *oggettivi* (o *quantitativi*) [4].



Un **metodo soggettivo** è basato sul giudizio umano. Tra questi ultimi ricordiamo [5]:

- **valutazioni del reparto vendite**: ciascun agente di vendita stima la domanda futura relativamente al proprio territorio per il prossimo periodo. L'ipotesi alla base di questo metodo, anche se non sempre vera, è che le **persone più vicine al cliente conoscono meglio di chiunque altro le sue necessità future**. Queste informazioni vengono successivamente aggregate per giungere a previsioni globali per ciascuna area geografica o famiglia di prodotti;
- **indagini di mercato**: le aziende spesso si rivolgono ad imprese specializzate nelle indagini di mercato per effettuare questo tipo di previsione. Le informazioni vengono ricavate direttamente dai clienti o più spesso da un campione rappresentativo di essi. Questo tipo di indagine, comunque, viene soprattutto utilizzata per cercare nuove

idee, cosa piace o non piace di prodotti già esistenti, quali sono le marche preferite di un determinato prodotto, etc;

- **panel di esperti:** le previsioni vengono sviluppate da un ristretto gruppo di esperti delle varie aree funzionali dell'azienda (marketing, finanzia e produzione) che interagiscono direttamente tra loro. La previsione viene sviluppata tramite incontri con scambi di idee ed informazioni tra managers di tutti i livelli; vi è però il problema che gli impiegati di più basso livello sono intimoriti da quelli di più alto livello, e per questo le loro opinioni non emergono spesso come dovuto.
- **metodo Delphi:** come si è visto, un'opinione di un impiegato di più alto livello, finisce col pesare di più di quella di un impiegato di basso livello. Il caso peggiore è che quest'ultimo non contribuisce alla discussione per non contrariare i suoi capi. Per prevenire questo tipo di problemi nel metodo Delphi è garantito l'anonimato di coloro che partecipano allo studio, in maniera tale che ognuno abbia lo stesso peso. Viene redatto un questionario che viene distribuito ai partecipanti. Le risposte vengono aggregate e viene preparato, in base a queste, un nuovo set di domande che vengono riproposte al gruppo.

La procedura può essere schematizzata in cinque fasi:

1. Scelta degli esperti. Ci devono essere impiegati appartenenti a più aree aziendali e a diversi livelli;
2. Tramite un questionario inviato a tutti i partecipanti, inviato anche tramite e-mail, si ricava la previsione;
3. Si aggregano i risultati e si ridistribuiscono ai partecipanti mediante un appropriato nuovo set di domande;
4. Si aggregano di nuovo i risultati, si affina la previsione e si sviluppa ancora un nuovo questionario;
5. Se necessario si ripete la fase 5 e si distribuisce il risultato finale ai partecipanti.

Il metodo Delphi generalmente raggiunge dei risultati accettabili in tre tornate ed il tempo richiesto è funzione del numero di partecipanti, quanto tempo e lavoro impiegano per sviluppare la loro previsione e la loro velocità nel rispondere al questionario.

I **metodi di previsione oggettivi** impiegano, invece, modelli matematici e dati storici per prevedere la domanda. L'ipotesi base è che il futuro si assume essere uguale al passato. Esistono due tipi di metodi oggettivi: il metodo delle *serie storiche* ed il metodo degli *indicatori economici* [3].

Il **metodo degli indicatori economici** consiste nel ricercare un'espressione funzionale che pone in correlazione l'entità della domanda di un prodotto finito, o una famiglia di prodotti finiti, ad alcuni indicatori economici. Questi indicatori sono delle variabili che descrivono le condizioni economiche prevalenti in un determinato periodo di tempo. Esempi di indicatori sono: reddito nazionale lordo, reddito procapite, reddito agricolo ed industriale, licenze edilizie concesse, produzione automobilistica, livello di occupazione, prezzi al consumo e all'ingrosso, depositi bancari, produzioni industriale, produzione di acciaio e cemento, ecc. Se si indica con Y la domanda di un prodotto che si vuole prevedere e con X_1, X_2, \dots, X_n le n variabili che si suppone sono collegate ad Y , allora il metodo asserisce che [3]:

$$Y = f(X_1, X_2, \dots, X_n).$$

Si può assumere, ad esempio, che tale relazione sia lineare, in questo caso:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n.$$

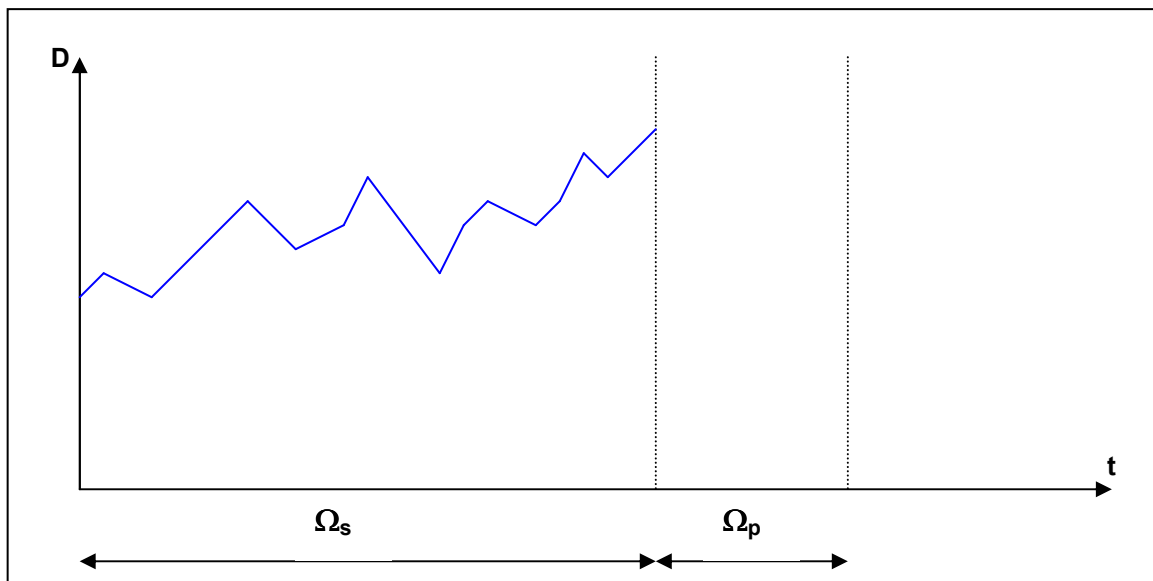
Normalmente, per la stima delle precedenti costanti si adopera il metodo dell'analisi di regressione multipla.

Metodi basati sulle serie storiche

Il metodo delle serie storiche è quello più comunemente impiegato per la programmazione della produzione. Una serie storica costituisce «**una rappresentazione numerica o grafica dell'andamento di una variabile in un prefissato intervallo di tempo**» [1].

Vediamo operativamente quali sono le ipotesi semplificative che si adottano.

Abbiamo a disposizione *un orizzonte storico* s per il quale sappiamo come si è comportata la domanda D in funzione del tempo.



La prima ipotesi è quella di ritenere che nel futuro la variabile si comporterà secondo la stessa legge che ha caratterizzato la serie storica nell'intervallo Ω_s . Questo significa ipotizzare che nell'orizzonte di previsione Ω_p la funzione f avrà la stessa espressione che nel passato;

La seconda ipotesi consiste nel porre :

$$D=\varphi(t)$$

abbiamo cioè escluso le n variabili e consideriamo la f funzione solo del tempo . E' un'ipotesi forte che considera costanti le variabili x_1, x_2, \dots, x_n , consentendo di passare da uno spazio (a $n+1$ dimensioni) ad un piano $D(t)$; ciò che rende accettabile questa ipotesi è l'orizzonte di previsione che si considera, infatti, nella scala dei tempi, teoricamente infinita, andiamo ad aprire una finestra di previsione abbastanza ristretta (qualche mese), che permette di ritenere pressoché costanti le n variabili.

La terza ipotesi impone che la $D=\varphi(t)$ sia una somma funzionale:

$$\varphi(t)=T(t)+S(t)+C(t)+A(t)$$

- $T(t)$: *Componente di trend*, che tiene conto dell'andamento a lungo termine del valore medio della variabile in oggetto;
- $S(t)$ *Componente di stagionalità*, che tiene conto delle variazioni della variabile in oggetto che si ripetono ad intervalli regolari attorno al valore medio di base;
- $C(t)$: *Componente ciclica*, che è simile alla stagionalità eccetto che per il fatto che la

lunghezza e la intensità del ciclo può variare. Si associano i cicli con le variazioni economiche di lungo termine;

- *A(t)*: *Componente di aleatorietà*, considera gli scostamenti della variabile dal valore che essa avrebbe dovuto assumere per effetto delle precedenti componenti.

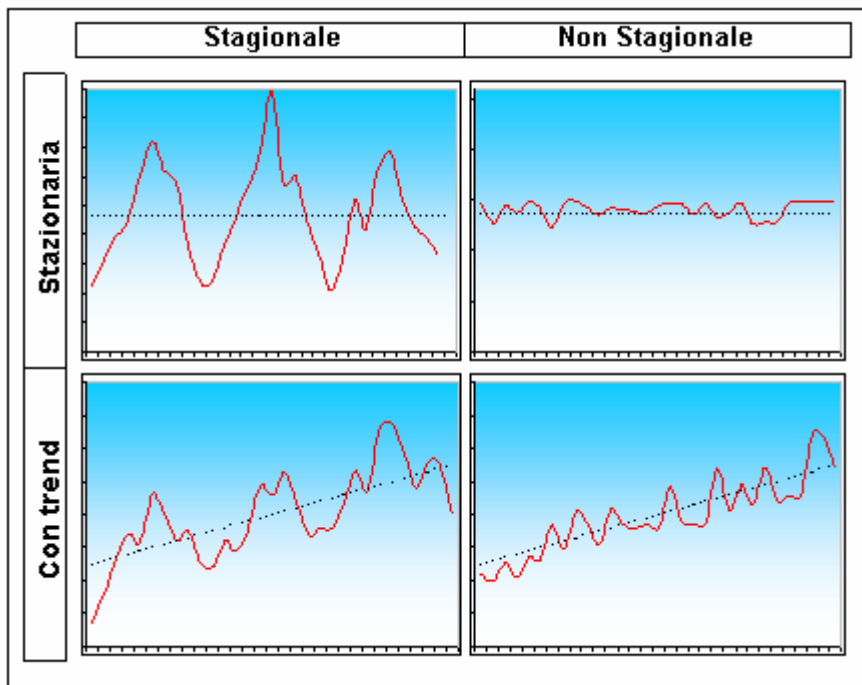
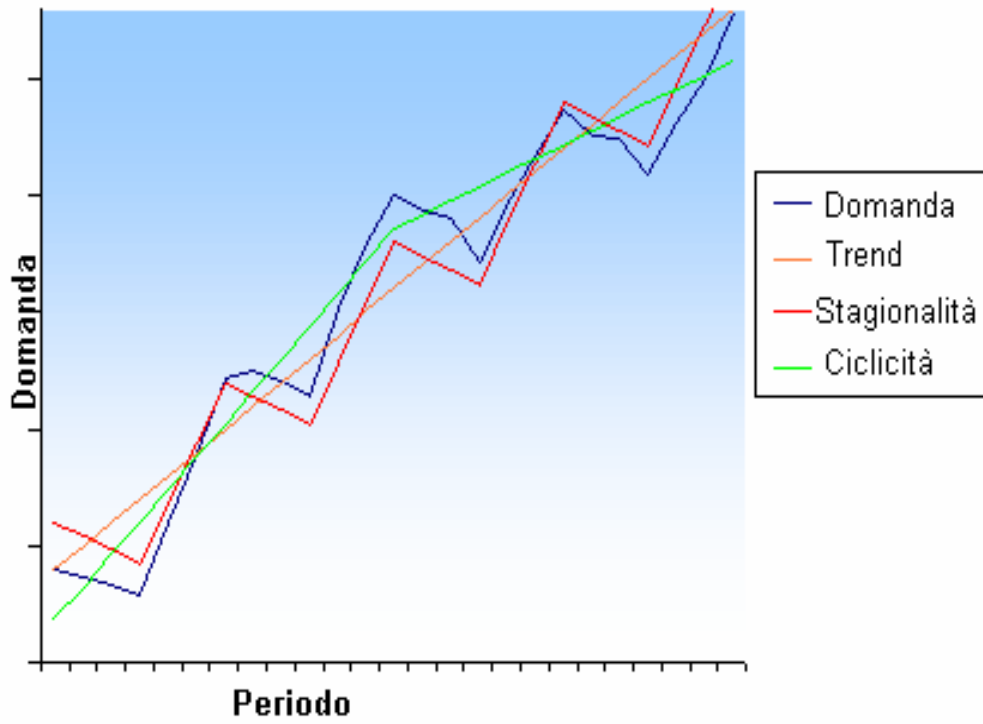
Quest'ultima, importante ipotesi presuppone un'indipendenza delle cause; nel caso in cui, invece, si ritiene che le cause siano **dipendenti**, si preferisce utilizzare il modello moltiplicativo:

$$I(t) = T(t) * S(t) * C(t) * A(t)$$

Non sempre tutte le componenti di cui sopra sono presenti in una serie storica, infatti, come mostrato in figura, si possono avere delle serie stazionarie, stagionali o con trend. Normalmente in sede di programmazione della produzione, la stima della futura richiesta viene effettuata, esclusivamente, in base alle componenti di trend e di stagionalità. In aderenza a tale assunzione, è agevole verificare che una serie storica può essere caratterizzata mediante quattro modelli:

1. *stazionario*: l'andamento della domanda appare distribuito intorno ad un valore medio ed ogni scostamento da tale valore può essere attribuito solo a cause aleatorie
2. *con trend*: l'andamento della domanda è continuamente crescente (o decrescente) al variare del tempo,
3. *stagionale*: la domanda presenta una variazione con periodicità accertata,
4. *stagionale con trend*: l'andamento della domanda è caratterizzato dalla presenza di una componente di trend ed una componente di stagionalità.

Componenti di una serie storica



Componente di trend

Per *trend* di una serie temporale s'intende la tendenza, negativa o positiva, che i dati manifestano ed è imputabile ad una causa sistematica che agisce nello stesso senso per un periodo più o meno lungo, da non confondere con quei fenomeni perturbatori, casuali, che producono delle oscillazioni non deterministicamente valutabili.

È quella funzione che ci dice come si comporta la variabile nel lungo periodo.

Esistono diversi andamenti possibili della $T(t)$:

$$T(t) = cost;$$

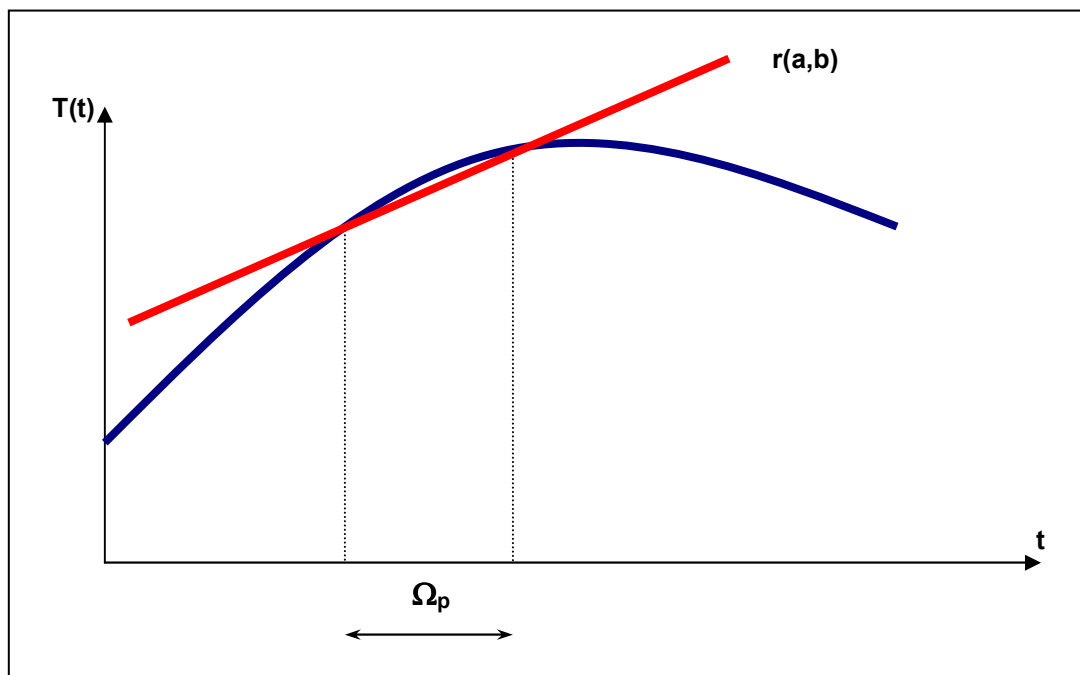
$$T(t) = a + bt;$$

$$T(t) = a + bt + ct^2.$$

Il modello più utilizzato per la componente di trend della domanda è quello analiticamente rappresentabile mediante un polinomio di primo grado:

$$T(t) = a + b*t$$

In realtà l'andamento è tutt'altro che lineare, ma visto che ci occupiamo di quello che succede su una "finestra" della curva di trend è più che plausibile l'approssimazione con una retta:

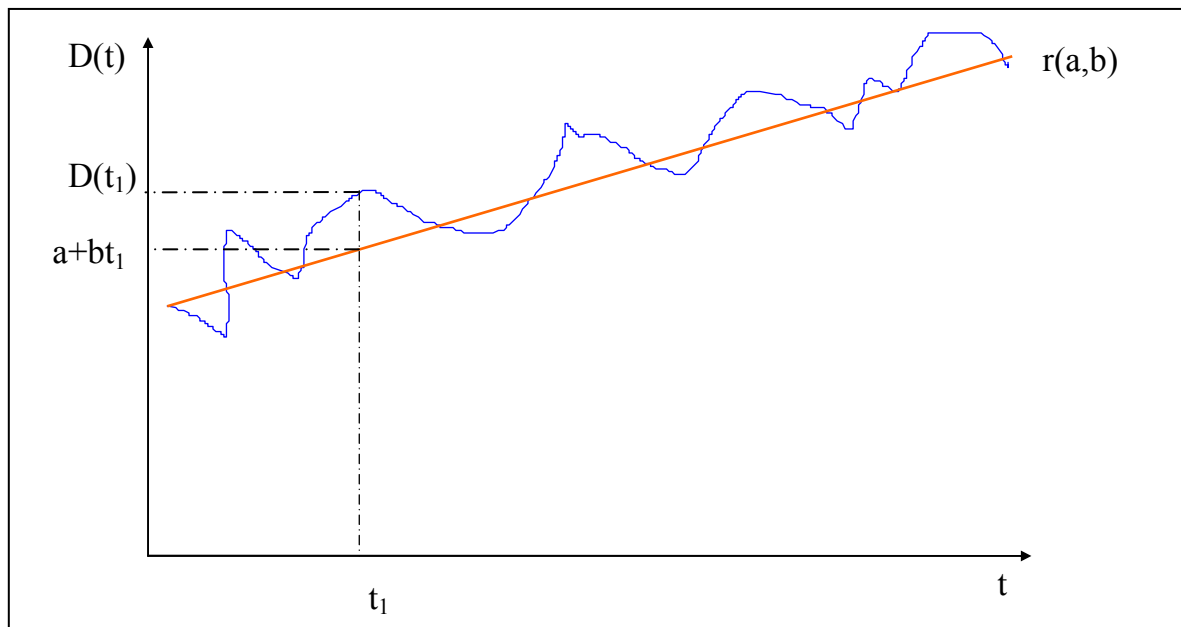


Cioè, se confondiamo la curva nel tratto *storico* Ω_p con una retta di parametri a, b non

commetteremo un errore significativo, semprechè i parametri, funzione del tempo, vengano ricalcolati man mano che ci si sposta e la pendenza della retta vari con continuità.

L'andamento della serie storica riportata su di un diagramma può essere rappresentato da un'insieme di punti qualsiasi; siamo interessati ad individuare la retta r , di parametri a e b , che approssimi al meglio questi punti (retta interpolare); per far ciò possiamo ricorrere al **metodo dei minimi quadrati**.

Consideriamo un istante t_1 generico: la domanda reale è definita da $D(t_1)$ ma, se in luogo della serie storica adoperassimo, per rappresentare il fenomeno, la retta r individueremmo per lo stesso punto un'ordinata $a+bt_1$, commettendo con tale rappresentazione lineare, un errore quadratico nel punto pari:



$[D(t_1)-(a+bt_1)]^2$ (il quadrato, o anche il valore assoluto, ha la funzione di evitare fenomeni di compensazione).

In definitiva, per n punti si ricerca quella retta di parametri a e b che renda minima la sommatoria degli scarti al quadrato:

$$\sum_{i=1}^n [D(t_i) - (a + bt_i)]^2 = \text{Min!}$$

In tal senso, si assumeranno come parametri caratteristici della componente di *trend* i coefficienti della retta che meglio approssima l'insieme dei dati a disposizione.

Tali coefficienti si determinano, com'è noto, risolvendo il sistema:

$$\frac{\partial}{\partial a} \left(\sum_{i=1}^n [D(t_i) - (a + bt_i)]^2 \right) = 0$$

$$\frac{\partial}{\partial b} \left(\sum_{i=1}^n [D(t_i) - (a + bt_i)]^2 \right) = 0$$

sviluppando i calcoli e risolvendo in funzione di a e b si ottengono i seguenti valori dei parametri:

$$\hat{a} = \frac{\sum_{i=1}^n D(t_i) - b \sum_{i=1}^n t_i}{n}$$

$$\hat{b} = \frac{n \cdot \sum_{i=1}^n t_i \cdot D(t_i) - \left(\sum_{i=1}^n t_i \right) \cdot \left(\sum_{i=1}^n D(t_i) \right)}{n \cdot \sum_{i=1}^n t_i^2 - \left(\sum_{i=1}^n t_i \right)^2}$$

Nel caso in cui si scelga di porre l'origine dell'asse dei tempi sul valore medio si ha:

$$\sum_{i=1}^n t_i = 0$$

il che comporta per i coefficienti della retta di regressione la forma:

$$\hat{a} = \frac{\sum_{i=1}^n D(t_i)}{n}$$

$$\hat{b} = \frac{\sum_{i=1}^n t_i \times D(t_i)}{\sum_{i=1}^n t_i^2}$$

Componente di stagionalità

Alla tendenza di lungo termine (trend) si possono sovrapporre fenomeni ciclici che si ripetono in maniera periodica e che vengono attribuiti agli effetti della componente stagionale sull'andamento di base della domanda (pensiamo al caso di un'industria di gelati o di ombrelli).

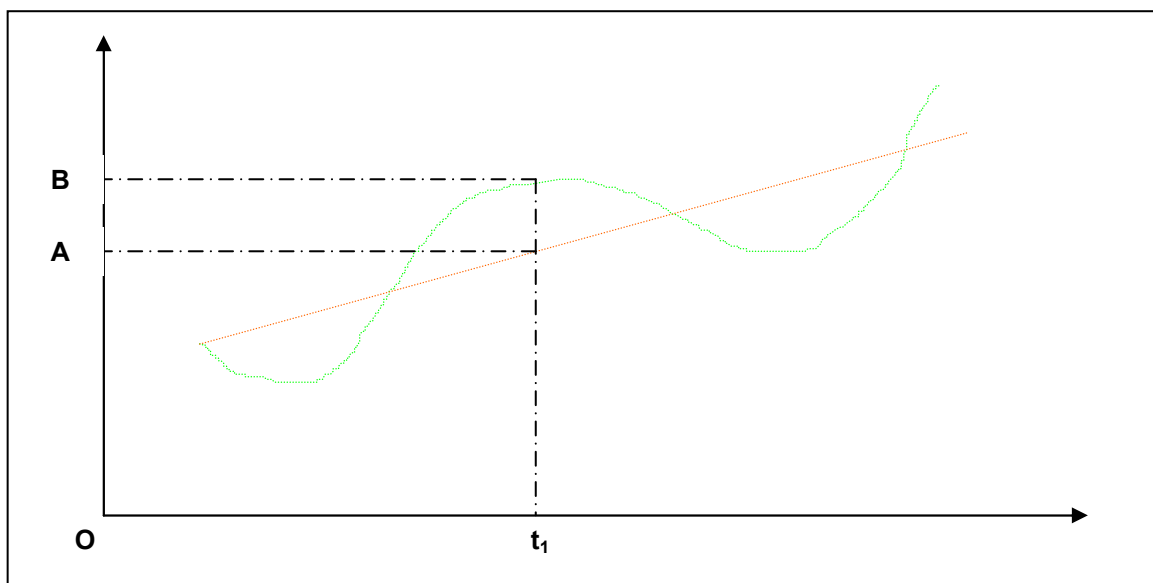
È evidente che le ciclicità non necessariamente coincidono con le stagioni meteorologiche, inoltre esse possono presentarsi sovrapposte, cioè all'interno di un fenomeno possono esserci più cicli con frequenze ed ampiezze diverse.

Al fine di identificare analiticamente tali effetti, occorre innanzitutto suddividere gli N_p periodi che costituiscono la serie storica in cicli di uguale lunghezza k . Alla funzione $S(t)$ verrà quindi assegnato un valore S_j , $j=1,2,\dots,k$ (S_j rappresenta il rapporto tra la media della domanda di tutti i periodi j e la media della domanda di tutti i periodi N , *metodo degli indici stagionali*).

La stagionalità può essere trattata secondo due modelli:

additivo:
$$D(t) = T(t) + S_j$$

moltiplicativo:
$$D(t) = S_j * T(t)$$



Qualora il fenomeno abbia una tendenza di lungo periodo (trend), descritta dalla retta $r(a,b)$, il caso più semplice di stagionalità si ha rappresentando il ciclo con una sinusoide. Pertanto, fissato un istante t , ad \overline{OA} , che rappresenta la componente di *trend*, dobbiamo sommare la componente *stagionale* \overline{AB} ; il modello additivo si complica nel caso in cui la componente di *trend* cresca linearmente con t , infatti il valore S_j rimane costante sicché l'incidenza percentuale della componente periodica diminuisce proporzionalmente con il trascorrere del tempo, converrà, in tal caso, correlare direttamente l'andamento della componente *stagionale* con quello della componente di base e ciò si ottiene esprimendo $D(t)$ mediante:

$$D(t) = S_j * T(t)$$

Componente aleatoria

Indica gli scostamenti dalla tendenza di base che possono variare la domanda con ricorrenza occasionale e a differenza della componente *stagionale*, tali deviazioni, non presentano una ricorrenza ciclica ma assolutamente casuale. Tuttavia un evento casuale appare quasi sempre correlato con una legge di probabilità di occorrenza ben definita; pertanto, una volta individuata l'influenza della componente di *trend* e di quella *stagionale* sul valore della domanda, è agevole determinare la probabilità che quest'ultima subisca scostamenti, dal valore deterministicamente calcolato, per effetto di fenomeni aleatori (sarà sufficiente applicare i noti principi di statistica e calcolo della probabilità).

Esempio:

Dato un andamento della domanda riassunto in tabella, eseguiamo una previsione della domanda.

Anno	Domanda nei periodi					Domanda tot annua
	1	2	3	4	5	
1	100	120	150	130	107	607
2	110	130	165	135	113	653
3	120	134	174	140	125	693
Domanda totale per periodo	330	384	589	405	345	
Media per periodo	110	128	163	135	115	=Σmedia domanda nei periodi (1, 2, 3)/3=Σ(medie per periodo)/5=130.2

Supponiamo di considerare solo il trend e la stagionalità, la domanda è quindi:

$$D(t) = S_j * T(t)$$

Supponendo inoltre il trend lineare, possiamo stimare i valori di a e b mediante il metodo dei minimi quadrati semplificato.

Riassumiamo il procedimento nella seguente tabella:

Anni	Periodo	t	D _t	t ²	tD _t
Primo anno	1	-7	100	49	-700
	2	-6	120	36	-720
	3	-5	150	25	-750
	4	-4	130	16	-520
	5	-3	107	9	-321
Secondo anno	6	-2	110	4	-220
	7	-1	130	1	-130
	8	0	165	0	0
	9	1	135	1	135
	10	2	113	4	226
Terzo anno	11	3	120	9	360
	12	4	134	16	536
	13	5	174	25	870
	14	6	140	36	840
	15	7	125	49	875
Periodi successivi	16		?		
	17		?		
TOTALE		0	1953	280	481

Da cui si ricava :

$$\begin{cases} \hat{a} = \frac{1935}{15} = 130.2 \\ \hat{b} = \frac{481}{280} = 1.72 \end{cases}$$

Quindi $D_t = (130.2 + 1.72 t) \cdot S_j$ con $j = 1, 2, 3, 4, 5$ periodi.

Per quanto riguarda la componente stagionale applichiamo il metodo degli indici stagionali: si otterranno cinque componenti stagionali poiché per ogni anno i periodi in esame sono cinque.

Bisogna calcolare la media di ogni periodo e dividerla per la media totale. Le medie del periodo si rilevano dalla tabella, mentre la media totale si ottiene facendo la media delle medie del periodo, cioè:

$$mediatot. = \frac{110 + 128 + 163 + 135 + 115}{5} = 130.2$$

Quindi passando al calcolo dei coefficienti stagionali:

Coeff.	S ₁	S ₂	S ₃	S ₄	S ₅
Formula	110/130.2	128/130.2	163/130.2	135/130.2	115/130.2
Risultato	0.84	0.98	1.25	1.04	0.88

In definitiva, la previsione al quarto anno è:

Primo periodo	$D(8) = 0.84 \cdot (130.2 + 1.36 \cdot 8) = 118.5$
Secondo periodo	$D(9) = 0.98 \cdot (130.2 + 1.36 \cdot 9) = 139.6$
Terzo periodo	$D(10) = 1.25 \cdot (130.2 + 1.36 \cdot 10) = 179.8$
Quarto periodo	$D(11) = 1.04 \cdot (130.2 + 1.36 \cdot 11) = 151$
Quinto periodo	$D(12) = 0.88 \cdot (130.2 + 1.36 \cdot 12) = 128.9$

Metodo della media mobile

La previsione del fabbisogno per un periodo futuro viene determinata dalla media aritmetica delle richieste osservate durante n periodi (ad esempio mesi) anteriori. La media si dice «mobile» perché viene costantemente aggiornata sostituendo via via l'ultimo dato disponibile al più lontano nel tempo. Analiticamente il calcolo della media mobile applicata alla previsione per un periodo futuro indicato con t risulta dall'espressione:

$$D_t = (d_{t-1} + d_{t-2} + \dots + d_{t-n})/n,$$

dove: D_t è la previsione per il periodo t , d_{t-1} è la domanda effettiva nel periodo precedente, n è il numero di periodi di ampiezza uguale a t .

Da ciò consegue che per elevati valori di n , la variabilità dei tempi anteriori risulterà più attenuata; per valori piccoli risulterà enfatizzata.

Questo metodo si utilizza in genere quando la domanda non cresce o decresce rapidamente e non si ha componente stagionale.

Esempio

Date le seguenti informazioni si vuole una previsione al dodicesimo mese:

Mesi	1	2	3	4	5	6	7	8	9	10	11	12
Domanda passata	420	410	430	380	350	370	340	330	380	420	440	?

Si ha:

Mesi	N°mese	D. passata	n=3	n=4	n=5
Gennaio	1	420			
Febbraio	2	410			
Marzo	3	430			
Aprile	4	380	420		
Maggio	5	350	407	410	
Giugno	6	370	387	392	398
Luglio	7	340	367	382	388
Agosto	8	330	353	360	374
Settembre	9	380	347	347	354
Ottobre	10	420	350	355	354
Novembre	11	440	377	367	368
Dicembre	12		413	392	382

Si può notare che per diversi valori di n si ottengono diverse previsioni:

- n=3 $D_{12} = 413$
- n=4 $D_{12} = 392$
- n=5 $D_{12} = 382$

Quindi appare evidente che si deve opportunamente scegliere il valore di n. Di solito tale valore è compreso tra tre e cinque, ma per determinarlo in maniera univoca si ricorre al seguente indice:

$$MAD = \frac{\sum_{t=1}^k |d_t - D_t|}{k}, \quad k = t - 1 - n$$

(MAD sta per *Mean Absolute Deviation* e come si vedrà meglio in seguito, è una delle più usate misure dell'errore della previsione).

Nell'esempio si ha:

Mese	Vendite previste	Vendite effettive	Differenza	MAD ₃
4	420	380	40	41.25
5	407	350	57	
6	387	370	17	
7	367	340	27	
8	353	330	23	
9	347	380	33	
10	350	420	70	
11	377	440	63	

Mese	Vendite previste	Vendite effettive	Differenza	MAD ₄
5	410	350	60	46.43
6	392	370	22	
7	382	340	42	
8	360	330	30	
9	347	380	33	
10	355	420	65	
11	367	440	73	

Mese	Vendite previste	Vendite effettive	Differenza	MAD ₅
6	398	370	28	47.33
7	388	340	48	
8	374	330	44	
9	354	380	26	
10	354	420	66	
11	368	440	72	

Quindi, in questo esempio, la scelta corretta è $n=3$ e $D_{12} = 413$.

Media mobile pesata

Con la tecnica precedente viene dato lo stesso peso a tutti gli elementi della serie storica, con una *media mobile pesata*, invece, è possibile attribuire un peso a ciascun elemento. In questo caso la formula è la seguente:

$$D_t = p_1 * d_{t-1} + p_2 * d_{t-2} + \dots + p_n * d_{t-n},$$

con lo stesso significato dei simboli precedenti e dove $p_1 \dots p_n$ sono i pesi attribuiti in base all'esperienza ai rispettivi periodi tali che:

$$\sum_{i=1}^n p_i = 1.$$

Più lontane saranno le informazioni e meno peso esse avranno.

Esempio

Si abbiano i seguenti dati:

Mese	Domanda	Peso
1	100	0.10
2	90	0.20
3	105	0.30
4	95	0.40

Si vuole una previsione della domanda nel mese 5:

$$D_5 = 0.10 \cdot 100 + 0.20 \cdot 90 + 0.30 \cdot 105 + 0.40 \cdot 95 = 97.5$$

Exponential smoothing

Un'altra versione più raffinata, affidabile e facilmente automatizzabile, è quella che prevede l'attenuazione esponenziale (*exponential smoothing*) del peso dei dati più lontani. Si tratta di un particolare tipo di media mobile, ove i fattori sono ponderati in base ad un parametro sottoposto a continua revisione, che decresce geometricamente in funzione della loro anzianità, offrendo il modo, quindi, di agire sull'apporto informativo di ciascuno, alterandone il peso in funzione del grado di aggiornamento e della conseguente significatività. In altri termini alla fine di ogni fissato intervallo di tempo t il valore del parametro che si sta osservando, cioè la domanda D_t , viene modificato alla luce dei mutamenti rilevati. Se in sede previsionale all'istante $t-1$ si è prevista una domanda D_{t-1} , mentre la domanda reale successivamente si è rivelata essere d_{t-1} , allora si sarà commesso un errore pari a $[d_{t-1} - D_{t-1}]$; pertanto la previsione successiva si effettuerà sommando all'ultima previsione, D_{t-1} , un'aliquota correttiva α dello scostamento $[d_{t-1} - D_{t-1}]$.

Analiticamente la domanda media stimata di un periodo futuro t , viene espressa dalla funzione:

$$D_t = D_{t-1} + \alpha (d_{t-1} - D_{t-1}),$$

dove: D_t è, come al solito, la previsione per il periodo t , D_{t-1} la previsione effettuata per il precedente periodo, d_{t-1} è la domanda che si è verificata nel precedente periodo e α è il *coefficiente di attenuazione esponenziale* (o *smoothing constant*), che è una misura di quanto si vuole pesare l'errore. Reiterando la precedente formula, si ottiene:

$$D_t = \alpha * d_{t-1} + \alpha * (1 - \alpha) * d_{t-2} + \alpha * (1 - \alpha)^2 * d_{t-3} + \dots + \alpha * (1 - \alpha)^{t-1} * d_0,$$

dove: $d_{t-1} \dots d_0$ rappresentano la domanda che si è verificata nei periodi precedenti.

Da notare che per

- $\alpha \rightarrow 1$, modello ipersensibile, la previsione diventa più reattiva;
- $\alpha \rightarrow 0$, modello insensibile, in pratica diventa una semplice media mobile.

È facile intuire che al di là di un certo valore il coefficiente è così piccolo da diventare

trascurabile. Si può mostrare che il numero di periodi N significativi, cioè necessari per individuare il nuovo valore della domanda prevista, dipende da α secondo la relazione:

$$N = \frac{2 - \alpha}{\alpha}$$

Quindi già per $\alpha = 0.2$ solo gli ultimi nove termini della serie sono significativi per la previsione; ovviamente il numero di periodi da prendere in considerazione aumenta al diminuire di α .

In definitiva si tratta di un metodo di ponderazione dei dati non equiponderale che attribuisce pesi diversi a dati di età diversa che consente di adattare meglio e più velocemente i parametri caratteristici del modello ai mutamenti che di volta in volta intervengono a modificare l'andamento della domanda; infatti con questa tecnica assumono rilievo solo i valori più recenti della richiesta. Un'altra facilitazione è dovuta alla scarsa numerosità di dati da dover gestire; bisogna, infatti, conservare solo la vecchia previsione D_{t-1} più il nuovo valore reale d_{t-1} e, ovviamente, α e non più tutti gli N elementi della serie.

Questa tecnica può essere efficacemente impiegata quando la domanda è stazionaria. Nel caso, invece, in cui è presente un trend è necessario introdurre un secondo coefficiente di attenuazione esponenziale denominato β [5]. Questo nuovo coefficiente riduce l'impatto dell'errore che si verifica tra la previsione e il valore attuale. Per il calcolo della previsione della domanda con correzione di trend si procede calcolando inizialmente il livello della domanda prevista:

$$D_t = \alpha \cdot d_{t-1} + (1 - \alpha) \cdot (D_{t-1} + T_{t-1})$$

essendo T il valore del trend.

Successivamente viene calcolato il trend corrente come media pesata tra il trend apparente, valutato sulla base della differenza rilevata tra i livelli previsti della domanda nei due periodi adiacenti ed il trend calcolato al periodo precedente:

$$T_t = \beta \cdot (D_t - D_{t-1}) + (1 - \beta) \cdot T_{t-1}$$

per poi, infine, rivalutare la domanda prevista con la correzione di trend:

$$DT_t = D_t + T_t$$

essendo DT_t è la previsione con il trend per il periodo t , e β è il secondo coefficiente di attenuazione esponenziale da utilizzare per il calcolo del trend.

L'analisi di regressione

Scopo di chi effettua la previsione, come abbiamo visto, è prevedere il valore di una variabile Y (per esempio le vendite), avendo a disposizione una variabile esogena, cioè non legata da una relazione funzionale con Y , X (per esempio le spese pubblicitarie). L'obiettivo diventa, quindi, sviluppare un modello che relazioni la variabile Y alla X . In altre situazioni vi è una variabile Y da prevedere ed una serie di variabili esogene (X_1, X_2 , etc.) conosciute e l'obiettivo è trovare una funzione che relazioni Y a tutte le altre variabili [7]. Nel primo caso si parla di *regressione semplice*, e sarà oggetto di questo paragrafo, nel secondo di *regressione multipla*, che per brevità non verrà trattato.

Nella situazione più generale vi è un set di n campioni appaiati (X_i, Y_i), con $i = 1, 2, \dots, n$.

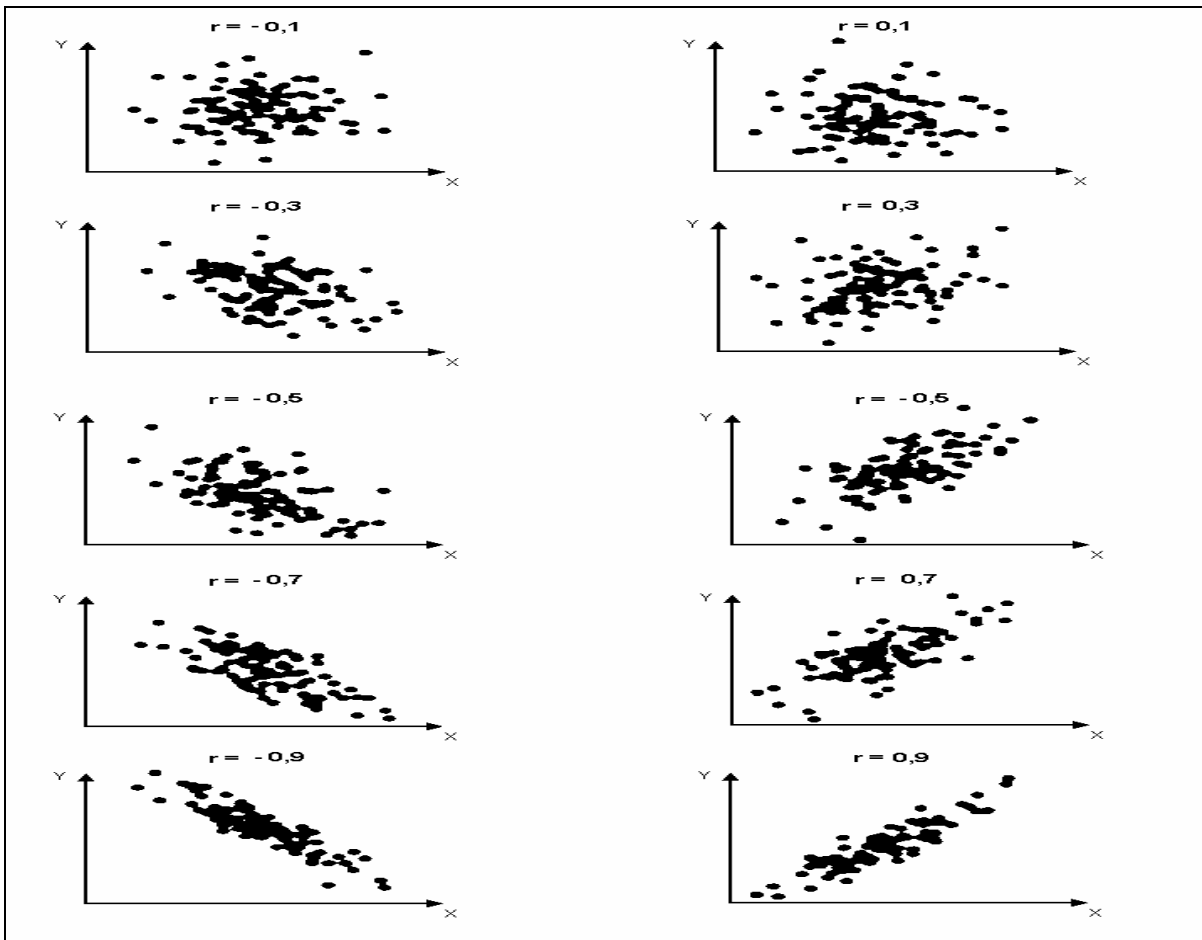
In genere si considera una relazione lineare tra X e Y rappresentata dall'equazione:

$$Y = a + bX + e$$

Dove a è l'intercetta, b l'inclinazione della retta ed e denota l'errore, cioè la deviazione dell'osservazione dalla relazione lineare.

L'obiettivo è di trovare i valori di a e b tali che la retta $\hat{Y} = a + bX$ presenti il miglior adattamento ai dati; a tale scopo si può utilizzare il metodo dei minimi quadrati già visto in precedenza.

Spesso accade che due variabili siano correlate l'una all'altra, ma risulta non corretto affermare che il valore di una variabile dipende dal valore dell'altra variabile. Per questo è stato introdotto il **coefficiente di correlazione** r che fornisce una misura della correlazione lineare tra due variabili. r può variare tra 0, nel caso in cui non vi è correlazione, ed $r = \pm 1$, nel caso di perfetta correlazione lineare. Quando $r > 0$, si dice che le due variabili sono correlate positivamente (al crescere di una cresce anche l'altra); quando $r < 0$, negativamente (al crescere di una decresce l'altra).



Il coefficiente di correlazione r_{XY} viene calcolato mediante la formula matematica

$$r_{XY} = \frac{Cov_{XY}}{S_x S_Y}$$

In cui

- Cov_{XY} è la covarianza tra X e Y , ovvero $\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$;
- S_X, S_Y è la deviazione standard rispettivamente di X e di Y .

La correlazione ci fornisce fondamentalmente due informazioni:

1. Il segno di r (+ o -), indica la *direzione* della relazione tra le due variabili. se è positivo, esse tendono ad crescere o decrescere insieme; se è negativo, una cresce mentre l'altra decresce; se è vicino a zero, esse crescono o decrescono separatamente.
2. Il valore di r è una misura della forza della correlazione, cioè quanto più il valore assoluto

di r è diverso da zero, tanto più le due variabili sono fortemente associate.

L'approccio Box-Jenkins

Box e Jenkins (1976) hanno concepito un modello più complesso di quelli visti precedentemente, ma comunque rappresenta un'opportunità per una previsione più accurata che è sicuramente interessante nel medio periodo e per prodotti aggregati.

La cosiddetta *media mobile autoregressiva* (AutoRegressive Moving Average, ARMA) valuta la domanda nel periodo attuale, d_t , attraverso una somma pesata di domande passate e di componenti casuali non prevedibili [6]. In termini matematici:

$$d_t = \phi_1 d_{t-1} + \phi_2 d_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

dove le ε sono le cosiddette *white noise*, vale a dire variabili casuali, stocasticamente indipendenti, distribuite normalmente con media zero e varianza costante; ϕ , θ , p e q sono costanti.

In genere la suddetta equazione può essere semplificata. Anderson (1976) afferma che molte serie storiche *stazionarie* sono adeguatamente rappresentate da modelli ARMA con $p+q \leq 2$, e, quindi, nell'equazione precedente vi sono al più tre termini al secondo membro. Per esempio, nel caso in cui $p=q=1$ si ha:

$$d_t = \phi_1 d_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Quest'ultima equazione, infatti, implica che la serie d_t sia stazionaria. Vi possono essere due casi in cui questa ipotesi non è valida: ci può essere un trend, lineare o di ordine superiore, nell'andamento di base; ci può essere una componente di stagionalità. In questi casi il modello prevede l'introduzione di termini addizionali nell'equazione di base.

Misura dell'errore

La domanda di un prodotto si genera attraverso l'interazione di un notevole numero di fattori, troppo complessa perché sia descritta completamente in un modello matematico. Di conseguenza, come è già stato accennato, gli errori di previsione non possono essere completamente evitati.

Come misurare l'accuratezza di un determinato metodo di previsione è quindi un aspetto molto importante che deve essere preso in seria considerazione.

Per illustrare meglio i concetti, ci riferiremo ad una serie di dati riguardanti la produzione mensile di birra in Australia [7]. La tavola seguente mostra i dati riguardanti gli ultimi otto mesi di osservazione (Gennaio – Agosto 1995).

Periodo t	Osservazione d_t	Previsione D_t
1	138	150.25
2	136	139.50
3	152	157.25
4	127	143.50
5	151	138.00
6	130	127.50
7	119	138.25
8	153	141.50
9	—	140.50
10	—	167.25

Nella terza colonna vi è la previsione ottenuta usando un metodo molto semplice: è stata fatta la media per ogni mese degli ultimi quattro anni; per esempio la previsione per Gennaio 1995 è la media della produzione di Gennaio 1991, Gennaio 1992, Gennaio 1993 e Gennaio 1994.

Se indichiamo con d_t la domanda effettiva per il periodo t e con D_t la previsione per lo stesso periodo, allora possiamo definire l'errore come:

$$E_t = d_t - D_t$$

Se vi sono osservazioni e previsioni per n periodi, allora ci saranno n valori dell'errore. Possiamo quindi definire:

- **Errore Medio (Mean Error):**

$$ME = \frac{1}{N} \sum_{t=1}^N E_t$$

- **Errore Medio Assoluto (Mean Absolute Error o Deviation):**

$$MAD = \frac{1}{N} \sum_{t=1}^N |E_t|$$

- **Errore Quadratico Medio (Mean Squared Error):**

$$MSE = \frac{1}{N} \sum_{t=1}^N (E_t)^2$$

Il ME dà, ovviamente, l'errore medio su più periodi, però c'è da osservare che gli errori di segno opposto tendono a compensarsi; quindi il ME fornisce semplicemente un'informazione circa l'esistenza o meno di un eventuale errore sistematico nella previsione.

Per questo motivo il MAD rende prima di tutto ogni errore positivo, prendendone il valore assoluto, e poi ne fa la media.

Un ragionamento analogo è alla base del MSE, in cui gli errori sono resi tutti positivi mediante elevazione al quadrato.

Il MAD ha il vantaggio di essere più facilmente interpretabile e più facile da spiegare ai non specialisti. Inoltre quando l'errore di previsione è distribuito secondo una gaussiana (il caso più comune), vi è una relazione tra deviazione standard σ e MAD:

$$\sigma = \sqrt{\frac{\pi}{2}} \times MAD \cong 1.25 \cdot MAD$$

Il MSE, d'altro canto, ha il vantaggio di essere più facile da maneggiare matematicamente e per questo è spesso usato nelle ottimizzazioni statistiche.

La misura dell'errore ottenuta attraverso il ME, il MAD o il MSE dipende dalla scala, dall'ordine di grandezza dei dati. Per questo motivo, le suddette misure non facilitano il confronto tra diverse serie temporali e per differenti intervalli di tempo. Un errore di 10 MI nella previsione mensile della domanda di birra è sicuramente differente da un errore di 10 MI nella previsione annuale della domanda di birra, o quando si prevede il consumo d'acqua di una città.

Per fare confronti come questi c'è bisogno di misure dell'errore *relative* o *percentuali*. Prima di tutto si definisce **errore percentuale** (*Percentage Error*):

$$PE_t = \left(\frac{d_t - D_t}{d_t} \right) \times 100$$

Le due misure relative dell'errore più usate sono:

- **Errore Medio Percentuale** (*Mean Percentage Error*):

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

- **Errore Medio Percentuale Assoluto** (*Mean Absolute Percentage Error*):

$$MAPE = \frac{1}{n} \sum_{t=1}^n |PE_t|$$

Il PE_t può essere utilizzato per calcolare l'errore percentuale per ogni periodo. La media dei PE_t su più periodi fornisce il MPE. Anche ora, come per il ME, nel MPE gli errori di segno opposto tendono a bilanciarsi e per questo si utilizza il MAPE.

La tavola seguente mostra come calcolare ME, MAD, MSE, PE, MPE e MAPE.

Periodo t	Osservazione d_t	Previsione D_t	Errore $d_t - D_t$	Errore assoluto $ d_t - D_t $	Errore al quadrato $(d_t - D_t)^2$	Errore percentuale $\left(\frac{d_t - D_t}{d_t}\right) \times 100$	Errore percent.assol. $\left \frac{d_t - D_t}{d_t}\right \times 100$
1	138	150.25	-12.25	12.25	150.06	-8.9	8.9
2	136	139.50	-3.50	3.50	12.25	-2.6	2.6
3	152	157.25	-5.25	5.25	27.56	-3.5	3.5
4	127	143.50	-16.50	16.50	272.25	-13.0	13.0
5	151	138.00	13.00	13.00	169.00	8.6	8.6
6	130	127.50	2.50	2.50	6.25	1.9	1.9
7	119	138.25	-19.25	19.25	370.56	-16.2	16.2
8	153	141.50	11.50	11.50	132.25	7.5	7.5
Totale			-29.75	83.75	1140.20	-26.0	62.1
ME= $-29.75 / 8 = -3.72$							
MAD= $83.75 / 8 = 10.47$							
MSE= $1140.20 / 8 = 142.52$							
MPE= $-26.0 / 8 = -3.3\%$							
MAPE= $62.1 / 8 = 7.8\%$							

Dal punto di vista dell'utilizzatore finale della previsione, sapere che il MAPE di un metodo di previsione è il 7.8% può essere molto più significativo piuttosto che sapere che il MSE è 142.52.

Infine, vale la pena accennare un utile, ma poco conosciuto indicatore:

Indicatore di Durbin-Watson:

$$DW = \frac{\sum_{t=1}^N (E_t - E_{t-1})^2}{\sum_{t=1}^N (E_t)^2}$$

E' un ottimo indicatore della correlazione degli errori, la cui conoscenza può rivelarsi molto utile ai fini di un'accurata e corretta previsione.

L'esperienza mostra che $0 < DW < 4$;

In particolare:

- $DW \cong 2$: gli errori sono sostanzialmente casuali e si compensano a vicenda;
- $DW < 2$: gli errori sono autocorrelati positivamente¹;
- $DW > 2$: gli errori sono autocorrelati negativamente.

¹ Autocorrelazione positiva significa che se si commette un errore per eccesso, nel periodo successivo si verificherà un errore dello stesso segno.