

Università degli Studi di Padova



Facoltà di Scienze Statistiche
Corso di Laurea in Statistica e Gestione delle Imprese

**PREVISIONI DELLE VENDITE:
IL CASO LOTTO SPORT ITALIA S.P.A.**

Relatore: Prof.ssa Luisa Bisaglia

*Laureanda: Federica Franzoia
Matricola n° 599217*

Anno Accademico 2010-2011

INDICE

INTRODUZIONE.....	5
CAPITOLO I – L’AZIENDA E IL MERCATO	7
1.1 - Il distretto della calzatura sportiva di Montebelluna	7
1.2 - Introduzione all’azienda Lotto Sport Italia	9
1.2.1 - La nascita dell’azienda	9
1.2.2 - Il risanamento e la riorganizzazione	11
1.2.3 – L’azienda Lotto Sport Italia attuale	13
1.2.4 - La delocalizzazione produttiva	16
1.3 - Le tipologie di prodotto e i canali distributivi	20
1.4 - Le Collezioni	25
1.5 - La Business Unit EMEA	26
1.5.1 – Ripartizione Fatturato B.U. EMEA	26
1.5.2 – Planning & Controlling Analysis	27
CAPITOLO II – METODI E OBIETTIVI DELLE PREVISIONI STATISTICHE	29
2.1 – La previsione in azienda.....	29
2.2 – Le serie storiche	30
2.2.1 – Approccio classico e approccio moderno	31
2.2.2 – L’approccio classico	32
2.2.3 – Il lisciamento esponenziale (Di Fonzo, Lisi, 2005).....	34
2.2.4 – L’approccio moderno	36
2.2.5 – Processi Stazionari	38
2.2.6 – Processi invertibili	39
2.2.7 – Le funzioni di autocorrelazione empiriche	39
2.2.8 – I principali processi stocastici (Di Fonzo, Lisi, 2005)	40
2.2.8.1 – Processi non stazionari	42
2.2.8.2 – Processi stagionali.....	43
2.2.9 – La procedura Box e Jenkins	45
2.2.10 – Misura dell’accuratezza delle previsioni	47
CAPITOLO III – APPLICAZIONE DEI MODELLI	49
3.1 – Il progetto di analisi.....	49

3.2 – I dati.....	49
3.3 – L'applicazione ai dati dei modelli SARIMA	52
3.3.1 – Analisi della serie di vendita di Tracksuits.....	54
3.3.2 – Le serie rimanenti.....	69
CONCLUSIONI.....	75
BIBLIOGRAFIA	77
PROGRAMMI STATISTICI.....	77

INTRODUZIONE

Questo elaborato nasce dall'esperienza di stage in un'azienda di notevoli dimensioni come Lotto Sport Italia e dalla volontà di vedere applicati al mondo reale del lavoro dei concetti per me molto interessanti ma visti finora solo negli studi come l'analisi delle serie storiche. Questi tre mesi in azienda mi hanno permesso, inoltre, di conoscere il mondo aziendale e in particolar modo alcuni dei processi aziendali più vicini al mio campo di interessi nonché agli studi. Lo stage si è svolto all'interno dell'ufficio *Planning and Controlling Analysis*, organo di supporto nonché fulcro dell'analisi statistica per l'intera azienda; con questo elaborato si è avuto modo di affiancare uno dei processi aziendali che vengono svolti nell'ufficio, il processo di *forecasting*, che costituisce un'importante fase di programmazione economica aziendale.

Verrà inizialmente descritto il settore dello *Sport System* montebellunese, dove si colloca Lotto Sport Italia, e illustrata la storia dell'azienda dal 1973, anno della nascita, passando per la crisi e il risanamento fino ad arrivare alla situazione attuale.

Dopo aver illustrato brevemente le Collezioni, le tipologie di prodotto e i canali distributivi, verranno messe in risalto le difficoltà che l'azienda deve affrontare e le decisioni che deve prendere per riuscire a gestire il rapporto con il cliente rispettando le tempistiche ma soprattutto realizzando l'obiettivo che si era prefissata. Questo meccanismo risulta ancor più valido se, come nel caso di Lotto Sport Italia, siamo in presenza di una delocalizzazione produttiva totale. Per raggiungere determinati obiettivi l'azienda deve rispettare dei vincoli di costo che si concretizzano in due punti principali: innanzitutto produrre in giusta quantità senza eccedere né essere insufficienti e, in secondo luogo, produrre tali quantità nelle tempistiche richieste dal cliente e in modo che rimangano in magazzino il minor tempo possibile. Per riuscire a soddisfare tali esigenze è necessario prevedere il prima possibile il dato che dovrà essere fornito alla Produzione ma per far ciò sono molti i fattori che bisognerà conoscere per formulare un'attenta analisi: l'andamento del mercato, le sue evoluzioni ma soprattutto i dati storici di vendita.

Nel secondo capitolo verranno riprese le tecniche statistiche di modellazione lineare per serie storiche, dal Lisciamento Esponenziale ai modelli SARIMA, e successivamente saranno proposte le fasi operative per la scelta del miglior modello per ognuna delle serie

e il calcolo delle previsioni. In particolare, sarà analizzata nel dettaglio la serie di vendita delle *Tracksuit*, macrocategoria che comprende tutti i modelli di tute sportive, descrivendone i passaggi e producendo infine delle previsioni mensili, utili per il forecasting.

È stato ovviamente più complicato rispetto alle analisi svolte a lezione, essendo queste serie reali e dunque con alcune "imperfezioni" rispetto agli esempi teorici dei libri o alle esercitazioni proposte durante i corsi universitari ma proprio per questo l'analisi è stata più stimolante. Si è potuto comprendere inoltre molte informazioni sugli andamenti dell'azienda, utili per integrare tutte le altre analisi. Inoltre, aspetto a mio avviso di importanza principale, ho avuto la possibilità di applicare alla realtà le conoscenze apprese durante gli studi che sono state così ampliate e mi hanno permesso di comprendere più a fondo l'importanza di tali strumenti.

CAPITOLO I – L’AZIENDA E IL MERCATO

1.1 - Il distretto della calzatura sportiva di Montebelluna

I distretti industriali (DI) sono una forma particolare di organizzazione della produzione che ha raggiunto un elevato livello di presenza e sviluppo in Italia.

Un DI è un agglomerato di imprese, in genere di piccola e media dimensione, ubicate in un ambito territoriale circoscritto e storicamente determinato, specializzate in una o più fasi di un processo produttivo e integrate mediante una rete complessa di interrelazioni di carattere economico e sociale (Grandinetti, 2005).

Un tipico DI funziona come un concentrato di singoli contesti aziendali in cui si producono conoscenze e da un’altrettanto elevata densità di canali interni di trasferimento della conoscenza.

A loro volta i processi di trasferimento alimentano la produzione di nuova conoscenza, in quanto le conoscenze trasferite possono entrare in combinazione con altre, generando rielaborazioni e sintesi originali.

All’origine di un distretto industriale c’è la presenza di una comunità locale ben integrata, con frequenti rapporti interpersonali tra i soggetti locali, arricchita ed integrata da un’accentuata divisione del lavoro tra le imprese, scomponibilità del processo produttivo, legami tra piccoli produttori e mercati esterni di sbocco e continue innovazioni tecnologico-organizzative. Sono questi alcuni dei punti di forza che hanno determinato il successo del distretto Montebellunese (provincia di Treviso), che negli anni ha saputo trasformarsi e crescere per rispondere a esigenze sempre più complesse e diversificate, richieste da mercato e consumatori.

La nascita del distretto risale agli inizi del ’900, quando sorsero le prime industrie quali Tecnica, Dolomite e Nordica, che inizialmente realizzavano calzature destinate soprattutto al mondo del lavoro. La svolta si ebbe nel primo dopoguerra, quando si riscoprì la passione per la montagna. Fu così che le imprese trevigiane, grazie alle conoscenze di base che avevano nella fabbricazione della calzatura, colsero l’occasione decidendo di progettare una scarpa adatta alle escursioni. Negli anni Trenta poi, con l’avvento dello sci nel nostro Paese, lo scarpone da sci divenne il vero punto di forza delle imprese del distretto, essendo inizialmente una variante della pedula da montagna.

Alla fine degli anni Sessanta si introdusse l'uso di materie plastiche nello scarpone da sci e questo rappresentò al contempo la fortuna e la sfortuna di molte aziende che, non possedendo il *knowhow* e non riuscendo a restare al passo con i *leader* di settore, decisero di dedicarsi ad altri tipi di calzatura sportiva come i pattini da ghiaccio, le scarpe da fondo, le scarpette da calcio, le scarpe da tennis, ecc. (www.montebellunadistrict.com). Nei primi anni Settanta, con l'idea di produrre scarpe da tennis, nasce Lotto.

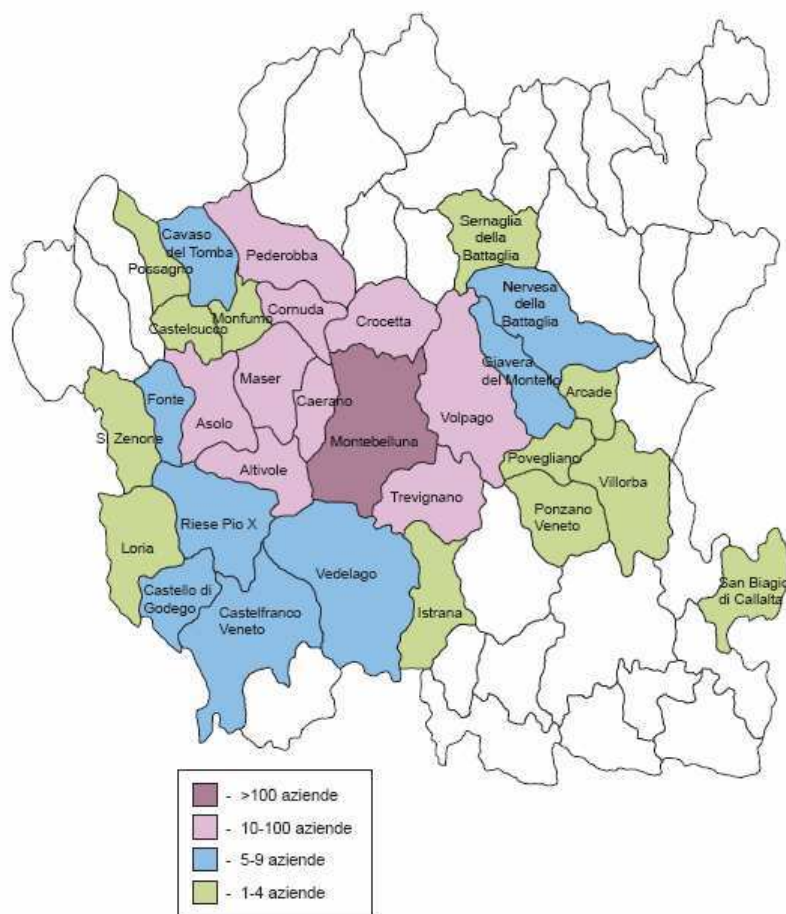
Attualmente il distretto della calzatura Montebellunese è composto da 499 aziende (dati al 2009, www.osservatoriodistretti.org), divisi tra calzaturifici e imprese legate al mondo della scarpa come studi di *design*, suolettifici, servizi di modellistica, ecc.

Il settore industriale delle imprese che producono articoli per lo sport è una delle principali realtà economiche dell'Italia e, sebbene sia costituito da molte imprese di dimensioni contenute rispetto ad alcuni grandi colossi di livello mondiale, contribuisce grazie all'identificazione del *made in Italy*, ad esportare nel mondo valori quali *design*, qualità ed innovazione tecnologica e di processi che fanno riconoscere ed apprezzare il nostro Paese all'estero.

Tale settore comprende in Italia imprese di medie e piccole dimensioni che sono equamente ripartite a seconda che producano calzature, abbigliamento o materiale sportivo in generale, e molte di esse trovano il proprio mercato operando in più di un comparto.

Il Veneto, in cui risiede la sede centrale di Lotto Sport Italia, è la regione italiana con il maggior numero di aziende operanti nello *sportsystem*, per lo più concentrate nella provincia di Treviso. Secondo l'ultimo rapporto OSEM (2009) le aziende operanti sono 390 per un totale di circa 7629 addetti e la maggior parte di tali imprese si dedica alla produzione di calzature, abbigliamento o accessori; altre invece coprono tutti i segmenti del mercato sportivo, come l'azienda in esame. Tra queste, la totalità delle imprese con un fatturato superiore a 25 milioni di euro attua una politica di decentramento mentre se analizziamo la tendenza alla delocalizzazione delle aziende che operano nei diversi comparti, abbiamo una percentuale del 100% per i produttori di scarpe da calcio, ciclismo, tennis, running; la percentuale scende un po' (80-90%) con le scarpe di sicurezza e da città mentre troviamo tra il 70 e l'80% i produttori di doposci, pattini, scarpe da montagna e abbigliamento.

Il fatturato complessivo del 2008 del distretto dello *sportsystem* è pari a poco più di 2 miliardi 127 milioni di euro cui il 44% è rappresentato dalla scarpa da città, il 20% dall'abbigliamento, il 15% dalla calzatura tecnica (www.osservatoriodistretti.org).



1.2 - Introduzione all'azienda Lotto Sport Italia

1.2.1 - La nascita dell'azienda

Lotto viene fondata nel 1973 dai fratelli Caberlotto con l'intenzione di avviare un'attività imprenditoriale nel mercato della calzatura sportiva. Iniziarono la loro avventura puntando sulla scarpa da tennis, ma sin da subito intuirono che il settore che sarebbe divenuto poi il vero *core business* dell'azienda era quello della calzatura da calcio.

I tre fratelli Caberlotto avevano già esperienza imprenditoriale essendo stati proprietari della società Caber, secondo leader mondiale nella produzione di scarponi da sci, ceduta poi alla multinazionale americana a cui fa capo la Spalding. I tre imprenditori fondano la

sede principale della nuova azienda nei pressi di Montebelluna, in provincia di Treviso, nello storico distretto Veneto della calzatura. La scelta del nuovo marchio è un elemento peculiare, che sta a sottolineare un proseguimento della loro attività: i fratelli Caberlotto adoperano, infatti, le restanti lettere del loro cognome che non erano state utilizzate per il nome della precedente azienda appena ceduta e nacque così l'idea "Lotto".

Già dopo il primo decennio Lotto era divenuta uno dei marchi di riferimento nel settore dell'articolo sportivo, essendosi specializzata, oltre che nelle calzature, anche nell'abbigliamento da gioco (in particolare tennis e calcio, a cui presto si affiancarono basket, atletica e pallavolo) e, più di recente, anche in quello d'ispirazione sportiva destinato al tempo libero.

La diversificazione produttiva, unita all'innovazione tecnologica e alle campagne di *marketing*, portarono Lotto a diventare uno dei marchi di riferimento nel settore dall'articolo sportivo in Italia. Così avvenne successivamente un'ulteriore espansione nel mercato della calzatura e seguendo l'evoluzione del mondo legato allo sport ed al tempo libero, approdando prima nel mondo della scarpa da calcetto e da *running* e successivamente in quello della scarpa da passeggio, con la creazione della linea *leisure*.

A partire dagli anni '60 fino alla fine degli anni '80, a seguito del processo di diffusione della pratica sportiva, la domanda era caratterizzata da una forte espansione e da un conseguente aumento dei produttori che tentavano l'entrata nel settore, favoriti dalla presenza di basse barriere all'entrata, soprattutto a causa dei bassi investimenti necessari. Per cercare di far fronte alla crescente concorrenza, all'inizio degli anni '80 l'azienda decide di tentare una strategia di distribuzione sui mercati esteri e di puntare, per il successo, a grandi politiche di sponsorizzazione e di promozione del marchio grazie anche alla collaborazione di alcuni atleti ai vertici delle classifiche internazionali e di campioni conosciuti in tutto il mondo (del calibro di: Ruud Gullit, Zvonimir Boban e Dejan Savicevic nel calcio; John David Newcombe, Thomas Muster e Boris Becker nel Tennis; squadre come il Milan prima e la Juventus poi). Comincia in questo modo la lunga e proficua tradizione di importanti sponsorizzazioni che caratterizzeranno l'azienda in tutta la sua storia e che le procureranno notorietà e credibilità internazionale.

La seconda metà degli anni '80, fino all'inizio dei '90, segna la fase, breve ma intensa, dell'iperbole tecnologica, dell'innovazione estrema nella ricerca di materiali e tecniche costruttive. Proprio l'innovazione tecnologica e la ricerca di materiali all'avanguardia

permettono all'azienda di dimostrarsi all'altezza di gestire questi nuovi presupposti, essenziali per rimanere sul mercato.

Nel frattempo avvengono rilevanti modifiche nel sistema distributivo: si affermano le catene di negozi specializzati che determinano un progressivo cambiamento dei rapporti di forza tra produttori e distributori, favorendo il potere contrattuale di questi ultimi.

Verso la fine degli anni '90 però l'azienda entra in una forte crisi, dovuta principalmente ad un rapido calo di partecipazione in alcune attività sportive (ad esempio il tennis) e al cambiamento della domanda, prima basata sul nuovo acquisto e successivamente legata al ricambio di prodotto: il cosiddetto mercato di sostituzione, infatti, richiede prodotti e servizi differenti ed una differenziazione sempre più marcata. A peggiorare questa situazione ci fu in contemporanea una crisi finanziaria a livello internazionale che incise profondamente sui consumi. Come se non bastasse nello stesso periodo si ebbe anche la morte di due dei fratelli fondatori, che portò l'ultimo dei Caberlotto a cercare di rimediare alla situazione negativa che si andava profilando.

Così nell'aprile del 1999, a seguito di un concordato preventivo, i soci con un gruppo di manager ed imprenditori locali già attivissimi nel settore dello *sport-equipment* ricostituirono l'azienda creando l'odierna Lotto Sport Italia S.p.A.

Con il passaggio, la guida dell'azienda venne assunta da Andrea Tomat, ex manager Lotto, e con lui il pacchetto di maggioranza era tenuto da G. Lorenzato, direttore della Lotto, A. Sartor, partner di Tomat in Stonefly, e R. Danieli, proprietario in passato di un'altra famosa azienda Montebellunese del settore, Diadora; l'azienda italiana è capogruppo di altre numerose società operanti nei Paesi esteri. (Rampin V., 2007-2008)

1.2.2 - Il risanamento e la riorganizzazione

Obiettivo della nuova proprietà è di valorizzare i punti di forza del marchio (dinamismo, innovazione, qualità, design made in Italy e passione per lo sport), uniti ad un servizio sempre più attento ed efficace alla clientela. (www.lottosport.com).

I fattori che hanno permesso all'azienda di risanare il *gap* creatosi sono stati una serie di nuove e più indicate strategie competitive in termini di mercati e di prodotti, che hanno portato ad una profonda modifica all'assetto aziendale e al suo posizionamento competitivo.

La rinascita ebbe inizio con la scelta di focalizzarsi nei settori più importanti e strategici per l'azienda, scegliendo di reindirizzare le proprie attività per avere maggiori opportunità di successo e migliorare lo sfruttamento delle proprie risorse.

Il nuovo core business viene definito con precisione ed identificato nella calzatura performance, concentrando così le risorse per far sì di poter sfruttare le economie di scala. Solo in seguito sarà ampliata l'offerta, ad esempio con abbigliamento perform e, sfruttando il *know-how* accumulato, si espanderà in settori affini come l'abbigliamento e la calzatura leisure, potendo così sfruttare le economie di esperienza.

Contemporaneamente si dette l'avvio ad una serie di misure che portarono l'azienda a modificare la struttura organizzativa ed operativa.

Per quanto riguarda la struttura distributiva, dal 2000 il mercato europeo viene servito direttamente dalla sede centrale di Montebelluna; tale scelta è dovuta alla volontà di avere sotto controllo tutto l'operato dell'azienda nei maggiori mercati esteri e, di conseguenza, di accentrare tutte le attività logistiche e di marketing, in maniera tale da ridurre i costi della complessa logistica preesistente, grazie all'eliminazione dei vari magazzini sparsi per l'Europa.

In Italia e nei Paesi che rappresentano un importante mercato per Lotto, l'azienda si compone di una rete di agenzie più o meno strutturate le quali, gestite direttamente dalla sede madre, si compongono solitamente di un ufficio commerciale con un responsabile che gestisce la rete di agenti.

L'obiettivo di accentrare tutta una serie di attività critiche nella sede madre è connesso a quello di mantenere alto il livello di servizi al cliente tramite la creazione di un *customer service* internazionale, che ha consentito una più efficace risposta nei confronti dell'acquirente ed un più tempestivo controllo rispetto l'evolversi del mercato.

Inizialmente la struttura organizzativa era una classica struttura matriciale, che seguiva una doppia specializzazione, quella funzionale e quella per area di business. Dal giugno del 2006 è divenuta una vera e propria struttura divisionale grazie alla separazione delle varie Business Unit per area geografica e per marchio (Lotto Works e Lotto Leggenda, quest'ultimo in pista solamente dalla stagione invernale del 2006) ed all'interno delle varie B.U. vi è un'ulteriore separazione a seconda della tipologia distributiva. La *strategic business unit* più importante, com'è facilmente intuibile, è l'EMEA (*Europe, Middle-East e Africa*) (Figura 1.2).

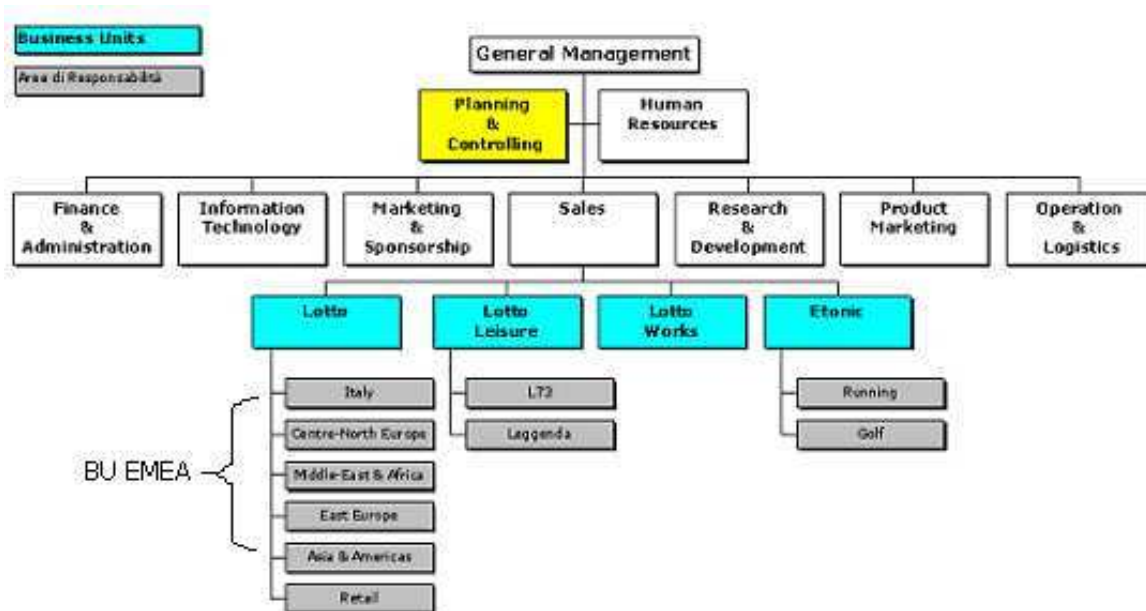


Figura 1.2 – Attuale organigramma di Lotto Sport Italia SpA

1.2.3 – L'azienda Lotto Sport Italia attuale

Oggi Lotto Sport Italia produce e distribuisce calzature ed abbigliamento tecnico per uomo, donna e bambino dedicati alla pratica del calcio, del calcetto e del tennis, con lo scopo di soddisfare sia le esigenze del giocatore professionista sia di quello amatoriale. A fianco di tale settore, denominato in ambito aziendale *performance*, ne è stato sviluppato un altro chiamato *leisure*, che si compone di abbigliamento e calzature per il tempo libero di chiaro stampo *lifestyle*, per uomo, donna e bambino, con un'immagine ed un gusto tutto italiano nella scelta di tessuti, forme e colori. In questo segmento si possono trovare prodotti della linea *training*, per una pratica sportiva non strettamente legata ad una disciplina specifica, e *sportswear* per il tempo libero. Tali prodotti esprimono quasi sempre un'ispirazione al *design* del mondo del calcio e del tennis, discipline che rappresentano il cuore dell'intera attività aziendale.

Sulla base delle competenze acquisite da Lotto nell'area della calzatura sportiva, l'azienda ha creato una nuova area di business denominata Lotto Works, che si occupa di gestire una collezione di abbigliamento e scarpe antinfortunistiche, conformi alle normative vigenti per la sicurezza negli ambienti di lavoro e di alta qualità. Dalla collezione autunno-inverno 2006 Lotto ha inoltre voluto allargare il proprio raggio d'azione espandendosi in un settore in piena evoluzione, quello della scarpa da città. Per realizzare ciò crea due nuovi marchi:

Lotto, LeDd e Lotto Leggenda. Leggenda, *griffe* rivolta ad una fascia di prezzo più alta, presenta modelli di tendenza e dal look accattivante, abbinando stili e materiali alternativi per un pubblico giovane e sempre attento alla moda. Vengono infine prodotti e distribuiti su licenza, da aziende terze, linee di occhiali da sole, intimo sportivo, camiceria e cartotecnica.

Il 2006 è stato un anno di svolta per l'azienda che, dopo un periodo di stallo di tre anni, ha ricompensato le decisioni strategiche e i grandi investimenti in Ricerca & Sviluppo del prodotto, nell'espansione commerciale e nelle acquisizioni, con un importante incremento del fatturato di otto punti percentuali rispetto all'anno precedente.



Figura 1.3 – Fatturato Lotto Sport Italia nel mondo

Il valore dei prodotti Lotto dal 2000 (anno della nascita del nuovo marchio) al 2006 ha avuto una variazione positiva del 60% raggiungendo un fatturato di 277 milioni di euro.

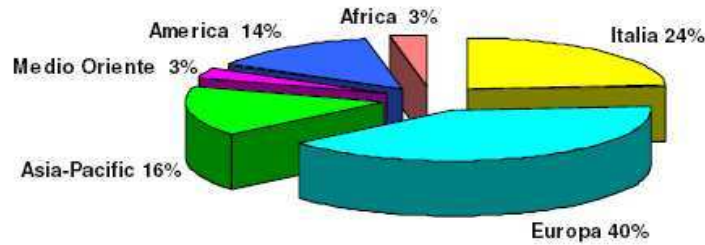


Fig 1.4 Ripartizione geografica del fatturato LSI nel 2007.

Fonte: Il Gazzettino di Treviso, estratto 14/07/2008

La scelta aziendale di espandersi all'estero risulta una soluzione vincente, confortata anche dagli incrementi dei ricavi nei vari mercati, ma resta sempre l'Italia il vero fulcro aziendale che apporta alla società ben un quarto del fatturato mondiale (2007).

“Il risultato è particolarmente soddisfacente – dichiara Andrea Tomat, Presidente dell'azienda – perché vede l'affermazione del marchio non solo nell'area tecnica di prodotto, dove abbiamo registrato una crescita del +5%, ma anche nell'offerta per il tempo libero con un incremento dell'11%, grazie anche al lancio della linea Lotto Leggenda. A valle della filiera grande propulsione è stata data alla presenza diretta del marchio sul territorio con l'inizio di un'articolata strategia di *retail* in Italia ed all'estero così come si è perfezionata la presenza di Lotto nei principali mercati esteri di Asia ed America in particolare. In prospettiva questo scenario schiude delle ampie opportunità per il futuro dove rafforzeremo ancora di più la ricetta che con dedizione proponiamo al mercato: innovazione italiana e *design* italiano.”

Ad oggi le vendite Lotto vedono una leggera flessione, dovuta in prima battuta alla crisi economica che ha colpito il mercato mondiale dal 2008 in avanti. Nonostante questa inversione di rotta, i cui effetti sono visibili dal 2009, si mantengono le percentuali dei quantitativi di vendita per tipo prodotto, con l'abbigliamento che rimane la categoria con il fatturato più elevato, addirittura più della metà dell'importo annuo.

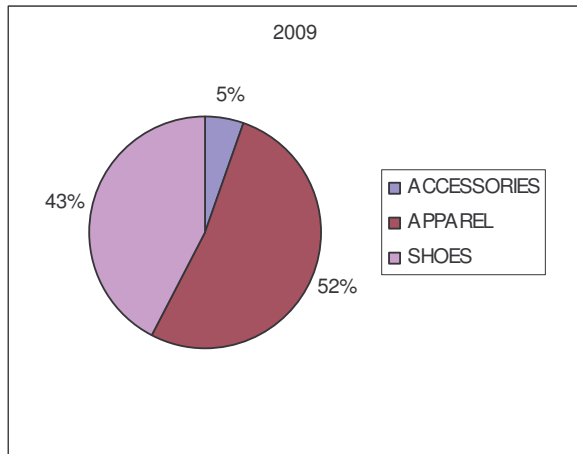


Figura 1.5 – Fatturato 2009 per tipo prodotto

1.2.4 - La delocalizzazione produttiva

La *delocalizzazione internazionale* delle imprese è un aspetto di un fenomeno più ampio definito *internazionalizzazione*, ovvero tutto un insieme di processi di sviluppo delle imprese nei mercati esteri. È dopo gli anni '70 che l'aumento della complessità della domanda, delle tecnologie e del mercato ha creato nelle imprese l'esigenza di godere di una maggiore flessibilità produttiva e si sviluppò quindi una nuova corrente che spinse ad esternalizzare fasi del processo produttivo avvalendosi di sub-fornitori o produttori esteri, localizzati in luoghi in cui era possibile godere di vantaggi particolari non disponibili nel Paese d'origine, come ad esempio una manodopera a basso costo. Il processo di delocalizzazione presenta diverse modalità di attuazione: dagli investimenti diretti fino alle più diverse forme di collaborazione con imprese autonome estere.

Anche la realtà italiana è stata interessata dalla delocalizzazione, ma la sua struttura particolare, data dalla presenza di imprese per lo più di piccole-medie dimensioni e da distretti industriali, fa sì che queste fasi di apertura debbano essere affrontate in modo prudente, cercando di individuare i pro e i contro, prevedendo eventuali soluzioni a tutela delle specificità che hanno reso unico il sistema produttivo italiano. All'interno dei distretti sono state per lo più le imprese produttrici dei prodotti finiti, che in genere sono anche le imprese *leader*, ad aver introdotto e sviluppato maggiormente il processo di internazionalizzazione. Grazie alla spinta di queste grandi aziende anche le imprese minori presenti nel distretto hanno avviato attività all'estero, principalmente di tipo mercantile ma

ultimamente hanno anche dovuto affrontare la concorrenza delle imprese estere presenti nei Paesi di delocalizzazione.

Ciò ha comportato il fallimento di molte aziende dovuta alla concorrenza di costo imbattibile, che però ha spinto il distretto e le sue piccole e medie imprese a mettersi in discussione e a riconsiderare la loro attività e il loro ruolo nella filiera produttiva e nella catena del valore.

Questa esperienza ha interessato anche il distretto mondiale della calzatura di Montebelluna e in particolare Lotto Sport Italia che è stata interessata da questo processo fin dalla propria nascita: ha delocalizzato per la prima volta la produzione di calzature nel 1976, e quella dell'abbigliamento a partire dal 1986. I Paesi cui puntò inizialmente sono quelli dell'Est Europa: Romania, Bulgaria, Ungheria; tuttavia essi non furono in grado di sviluppare una propria conoscenza sul prodotto tale da creare un *know-how* di base. Per questo Lotto puntò in seguito ai Paesi del Far East, ampliando quindi i propri orizzonti verso Taiwan, Cina, Corea, Indonesia e Thailandia, ove i vantaggi per le aziende estere erano rappresentati dal costo del lavoro ancor più basso dei Paesi dell'Est Europa e dalla maggiore disponibilità di risorse (materie prime, risorse energetiche), senza contare la vicinanza con i grandi mercati di sbocco come Giappone e Cina. Oltre a queste situazioni di base, si deve aggiungere la capacità da parte di queste popolazioni di acquisire rapidamente conoscenze sulle lavorazioni effettuate e quindi di creare un *know-how*. Con questi fornitori Lotto intrattiene un rapporto di *global sourcing*: questi produttori sono scelti su scala globale anziché locale e sono autonomi nell'acquisizione di risorse e nella strutturazione del processo produttivo; praticamente la delocalizzazione avviene come se fosse una semplice acquisizione di prodotti finiti da produttori esteri. Il rapporto tra Lotto e il fornitore estero è regolato attraverso delle *Trading* che consentono comunque all'impresa leader un monitoraggio costante sui prodotti in lavorazione. Dalla prima fase di delocalizzazione ad oggi Lotto ha continuato ad esternalizzare sempre più fasi legate alla produzione; oggi non attua più al suo interno alcun tipo di lavorazione: tutto è interamente delocalizzato nei Paesi esteri e solo in parte in Italia presso fornitori specializzati in particolari lavorazioni, per i prodotti che sono *top di gamma* e che richiedono la certificazione *Made in Italy*.

Lotto, nella definizione delle tempistiche e del *time to market*, deve tener conto che il ciclo di vita di questo tipo di prodotti è al massimo di 6 mesi e deve quindi progettare il lancio di

un prodotto con un anno di anticipo. Durante la fase di progettazione si definisce già la capacità produttiva dei fornitori disponibili, la si confronta con la capacità produttiva necessaria stimata in base al *budget* di vendita, e si individuano le fonti utilizzabili. I produttori esteri impongono dei lotti minimi di produzione al di sotto dei quali rifiutano l'ordine, in quanto non raggiungono le economie di scala necessarie per rendere conveniente la produzione. È in questa fase che gli analisti e le altre figure interessate devono valutare la convenienza a lanciare o meno un prodotto, valutazione fatta grazie al processo di *forecasting*.

La produzione di Lotto oggi è localizzata in Cina, Indonesia, Vietnam, Cambogia, Bangladesh e altre nazioni minori, come in figura:

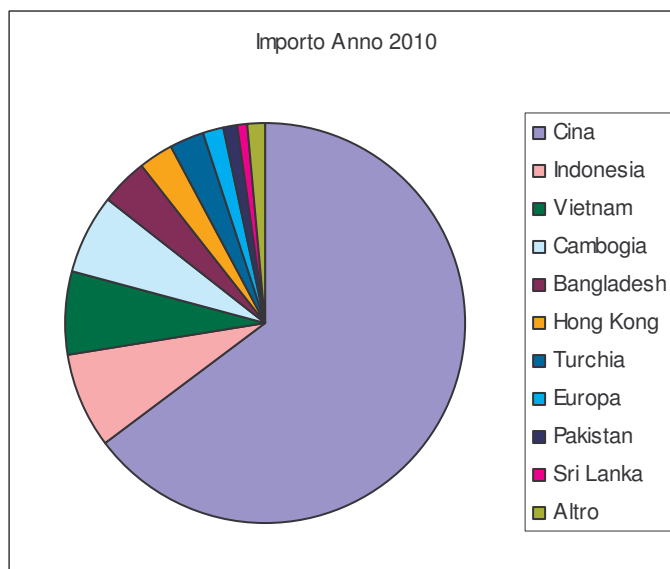


Figura 1.6 – Importo fornitori 2010

L'azienda controlla poi due *trading*: Lotto Sport Hong Kong Ltd e Lotto Sport Taiwan Ltd, di totale proprietà aziendale e con un organico di circa 100 persone.

La prima si occupa di coordinare le attività logistiche e di approvvigionamento delle fonti produttive che realizzano i capi di abbigliamento e gli accessori a marchio Lotto; la seconda invece segue le attività legate alla produzione di calzature.

Le due società svolgono quindi il compito di intermediari tra la Lotto Sport Italia e i propri clienti localizzati in Asia ed America: esse controllano la qualità dei prodotti, ricevono gli

ordini, e coordinano tutte le attività amministrative, finanziarie e quelle relative alla spedizione delle merci per i prodotti realizzati.

Questa scelta strategica ha portato molti benefici in termini di snellimento, accuratezza e velocità dei processi, riduzione dei costi e miglioramento dei servizi.

Ad oggi la produzione degli articoli marchiati Lotto è così strutturata:

- I prodotti d'abbigliamento arrivano per più del 40% dalla Cina, seguita dall'Indonesia con quasi il 30%, mentre il restante è prodotto negli altri Paesi in cui sono localizzate le aziende fornitrici di Lotto.

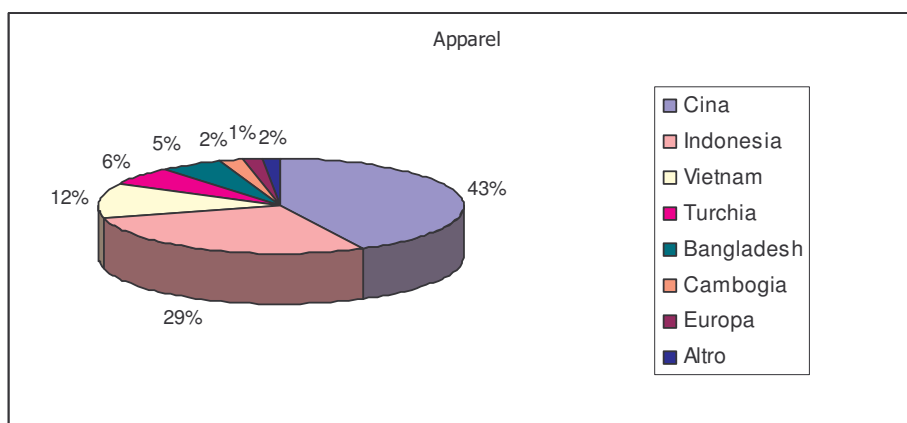


Fig. 1.7 – Ripartizione geografica dell'attività produttiva di LSI per l'abbigliamento

- Anche le calzature sono in maggior parte *Made in China*, per quasi un'80%, seguita da Vietnam e Cambogia. Solo le calzature top di gamma altamente tecniche e con una lavorazione altamente specializzata, che necessitano del marchio Made in Italy, vengono prodotte in Italia da produttori specializzati.

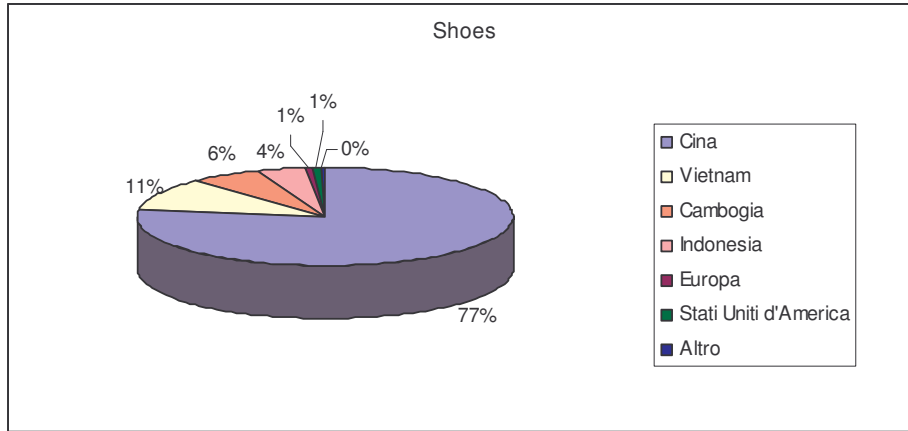


Fig. 1.8 – Ripartizione geografica dell'attività produttiva di LSI per le scarpe

- Per quanto riguarda gli accessori, invece, la maggior parte della produzione proviene da Hong Kong, seguita comunque sempre dalla Cina con un 22%.

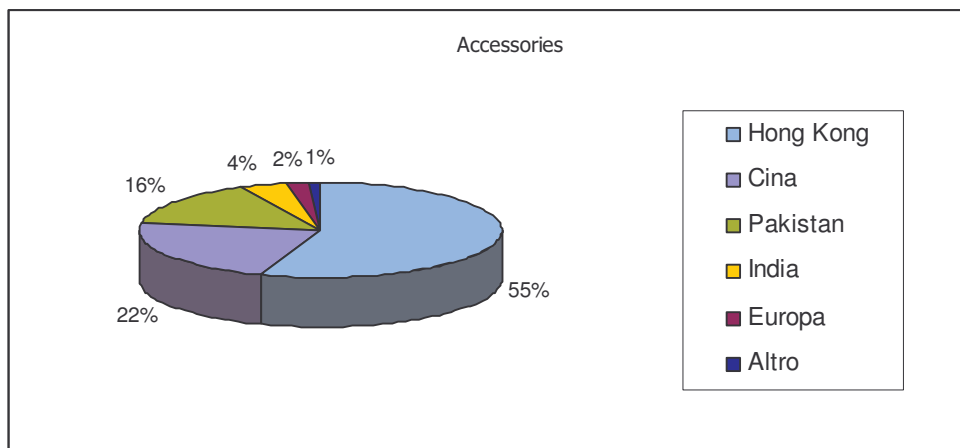


Fig. 1.9 – Ripartizione geografica dell'attività produttiva di LSI per gli accessori

1.3 - Le tipologie di prodotto e i canali distributivi

Per meglio comprendere la natura e la classificazione dei dati analizzati risulta molto importante chiarire quale sia il prodotto venduto da Lotto e come sia strutturata la rete distributiva dell'azienda.

Come già accennato, tre sono le tipologie di prodotto commercializzate (scarpe, abbigliamento ed accessori) e diverse le discipline, ovvero linee commerciali, quali calcio, calcetto, volley, ecc.

Con riferimento alla rete distributiva di Lotto risulta opportuno spendere prima due parole in merito all'organizzazione adottata in generale dalle aziende. All'interno del *marketing mix* i canali di distribuzione rivestono un ruolo importante in quanto, fornendo al consumatore l'accesso ai prodotti o servizi, lo mettono nelle condizioni di poter effettuare l'acquisto. Per politica distributiva si intende quindi l'insieme delle decisioni che l'impresa deve prendere per far arrivare il prodotto al consumatore finale: decisioni relative ai canali di distribuzione, alle scelte degli intermediari commerciali, all'organizzazione e gestione della forza di vendita (Rispoli, 1984).

Come è facilmente intuibile sono tutte decisioni di grande rilievo per l'impresa industriale. Infatti, dopo aver venduto il prodotto al distributore non si può certo ritenere di aver esaurito l'attività commerciale, ignorando che il prodotto è solo temporaneamente "depositato" presso il distributore e che molte volte agli sforzi commerciali sostenuti per farlo entrare nei negozi (*sell in*) si devono successivamente aggiungere altri sforzi per far sì che venga comprato, come le azioni promozionali (*sell out*).

Un canale di distribuzione consiste in un insieme di operatori che partecipano al trasferimento della proprietà dei beni dal produttore al consumatore, e l'attività degli operatori è favorita dalla presenza di numerosi ausiliari che non agendo in forma autonoma fanno parte del canale; sono quindi membri di un canale il produttore, il consumatore e tutti gli intermediari commerciali che intervengono (Rispoli, 1984).

I canali si distinguono in:

- *canali diretti*, con i quali il produttore è in contatto immediato col consumatore, per esempio attraverso delle proprie filiali di vendita;
- *canali indiretti*, nei quali il produttore si appoggia ad intermediari commerciali.

I canali indiretti sono a loro volta distinti tra:

- *canali brevi*, dove si richiede l'intervento di un solo intermediario fra il produttore e l'acquirente finale attraverso dei dettaglianti;
- *canali lunghi*, quando viene previsto il passaggio attraverso due o più intermediari come grossisti e dettaglianti.

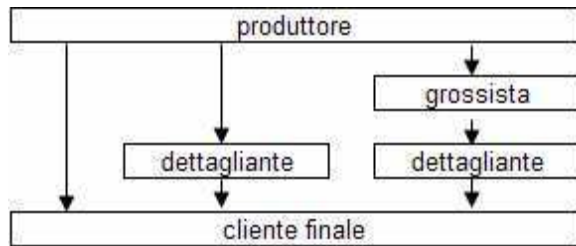


Figura 1.5 – I canali di distribuzione (Rispoli, 1984)

I fattori condizionanti la scelta dei canali di distribuzione possono essere distinti in fattori interni e fattori esterni all'azienda. Tra i fattori esterni più importanti vanno ricordate le caratteristiche di prodotto come ad esempio il grado di deperibilità degli stessi (che porta a ridurre o meno i passaggi della merce) e le caratteristiche della clientela, ad esempio il numero dei clienti potenziali da raggiungere e la loro concentrazione o dispersione geografica.

I fattori interni principali sono invece rappresentati dalla gamma dei prodotti, la capacità finanziaria aziendale e gli obiettivi di vendita. In merito agli obiettivi di vendita in termini di volume l'impresa deve decidere come vuole realizzarli, deve cioè decidere se agire su un maggior numero di punti vendita, cercando di assicurarsi un'elevata probabilità di contatto con il maggior numero di consumatori finali o su un minor numero di intermediari, ma altamente qualificati, per raggiungere un gruppo selezionato di consumatori.

La prima soluzione, che è di fatto quella adottata da Lotto Sport, viene definita *politica di vendita intensiva* e in genere viene realizzata utilizzando più canali. Lotto si avvale di tre tipi principali di intermediari commerciali, quindi i canali distributivi si articolano in tre gruppi: le Agenzie, i Distributori e Licenziatari ed il *Retail*.

Le Agenzie

In Italia e nei Paesi più importanti d'Europa che costituiscono i mercati di riferimento per l'azienda, la rete distributiva si compone di sette Agenzie più o meno strutturate; esse sono in Italia, Francia, Spagna, Germania, Austria, Grecia e Regno Unito.

L'agenzia di vendita è una forma di contratto con cui l'agente assume stabilmente l'incarico di promuovere, per conto del produttore, la conclusione di contratti di vendita in una determinata zona (Rispoli, 1984). In sostanza l'agente, o responsabile di mercato, si

assume l'incarico stabile e quindi continuativo di promuovere i prodotti o servizi venduti dall'azienda; il compenso per questa attività, chiamato provvigione, consiste di solito in una percentuale sul valore di ogni affare concluso.

Tale strategia distributiva permette di avere un ottimo controllo del mercato nel quale si va ad operare grazie alle Agenzie che rappresentano l'azienda nel territorio ed è quindi adatta ai Paesi di rilevanza strategica per l'impresa.

I Distributori e Licenziatari

In genere per distribuire i prodotti nei Paesi più lontani Lotto si avvale del supporto di Distributori e Licenziatari. I Distributori sono dei clienti veri e propri, cioè un'azienda terza che compra da Lotto e rivende poi la merce al dettaglio, gestendo quindi un magazzino proprio. Quindi mentre le Agenzie sono presiedute direttamente dalla sede madre ed è direttamente Lotto Sport che si accolla i rischi vendendo ai negozi finali tramite agenti situati nelle nazioni di riferimento, con i Distributori la catena distributiva richiede un passaggio ulteriore, a causa della presenza di un'impresa in loco che si incarica di distribuire ai negozi finali gli articoli marchiati Lotto. Con questa tipologia distributiva il carico di rischio non è più sulle spalle della casa madre, ma su quelle del rivenditore che si mette in gioco e si autogestisce.¹

I Licenziatari hanno la licenza di produrre e distribuire prodotti Lotto pagando all'azienda, come controprestazione per l'uso del marchio, una *royalty*, ossia un compenso pesato secondo il valore del venduto. Essi possono acquistare dei prodotti da Lotto al fine di distribuirli, però a differenza dei Distributori possono anche sviluppare alcuni modelli diversi dalla collezione della casa madre in modo da soddisfare al meglio il loro mercato di riferimento decidendo se richiederne la produzione a Lotto o scegliere altri fornitori, previa autorizzazione della casa madre. Lotto ha concesso licenze ad aziende distribuite in più di 50 Paesi nel mondo lontani dalla realtà Europea.

Questa tipologia di contratto permette di ottenere una reazione rapida ai cambiamenti del mercato, soprattutto grazie alla vicinanza dei partner al consumatore finale (sia in termini

¹ I distributori si trovano in: Arabia Saudita, Bolivia, Bosnia, Bulgaria, Cipro, Croazia, Cuba, Danimarca, Emirati Arabi, Ecuador, Estonia, Ghana, Irlanda, Islanda, Kuwait, Lettonia, Lituania, Malta, Norvegia, Paraguay, Perù, Polonia, Portogallo, Qatar, Repubblica Ceca, Romania, Russia Europea, Serbia, Siberia, Slovacchia, Slovenia, Suriname, Svezia, Svizzera, Ucraina, Ungheria, Uruguay.

geografici che culturali). Questi paesi possono rivolgersi per i loro acquisti alle due *Tradings* di proprietà Lotto.²

Il Retail

Oggi Lotto è presente in più di 80 Paesi ed è in continua evoluzione cercando di affermarsi come realtà importante a livello mondiale nel settore dell'abbigliamento e della calzatura sportiva. L'azienda ha sviluppato negli anni una rete distributiva mirata e ben studiata che comprende, oltre alle tipologie sopra indicate, anche una serie di *corner* Lotto e negozi monomarca in Italia e all'estero (in tutto oltre 300) che le permettono di essere presente in modo ancor più capillare. I *corner* sono anche chiamati *shop in shop*, letteralmente sono degli "angoli" di esposizione prodotti solitamente localizzati all'interno di punti vendita trattanti beni durevoli, i negozi monomarca sono invece punti vendita di prodotti facenti capo ad un unico *brand*.

Di fatto le Agenzie costituiscono un canale indiretto medio lungo, i Distributori un canale indiretto lungo, i Licenziatari il canale indiretto più lungo in assoluto, mentre il *Retail* è un canale diretto.



Figura 1.10 – I Paesi serviti da Lotto nel Mondo

² I licenziatari si trovano in: Argentina, Australia, Benelux, Brasile, Brunei, Canada, Cile, Cina, Colombia, Egitto, Giappone, Corea, Hong Kong, India, Indonesia, Israele, Malaya, Messico, Nuova Zelanda, Sebah, Sarawak, Singapore, Taiwan, Thailandia, Tunisia, Turchia, USA, Venezuela, Vietnam.

Importante è anche la distinzione tra tipologie di cliente. I dettaglianti serviti direttamente da Lotto attraverso le Agenzie vengono classificati in una determinata Tipologia Cliente a seconda dell'importanza strategica che rivestono e del tipo di distribuzione, in particolare:

- *Key Accounts*: clienti strategicamente più importanti vengono chiamati *Key Account* e sono considerati le colonne portanti della divulgazione del marchio Lotto; sono coloro che, acquistando un elevato quantitativo di merce, portano all'azienda un maggior fatturato. Essi hanno, inoltre, un trattamento particolare e sono regolarmente seguiti da un *customer service* che si occupa esclusivamente di loro. Sono generalmente costituiti da gruppi di acquisto (ad esempio: Intersport) o da catene della grande distribuzione organizzata.

- *Retailers*: clienti medio-piccoli, costituiti da uno o alcuni punti vendita, acquistano solo pochi articoli e magari solo alcune tipologie di prodotto in base alle competenze del negozio. Si tratta di clienti di dimensioni solitamente più piccole, perlopiù indipendenti cioè non legati a gruppi di acquisto.

1.4 - Le Collezioni

Una considerazione molto importante da tener presente ai fini della comprensione delle seguenti analisi è che nel mondo della calzatura e dell'abbigliamento il periodo di riferimento è da considerarsi semestrale e non annuale, data la naturale stagionalità del prodotto moda. Perciò si riferiscono ad un arco di tempo semestrale il processo di *forecasting* ed in parte anche il processo di *budgeting* che in seguito verranno illustrati.

Dato che il periodo di riferimento è semestrale, due sono le collezioni realizzate e due le conseguenti campagne vendita seguite dall'azienda nell'arco dei dodici mesi di un anno:

- la campagna vendite *Spring-Summer* (codificata SS), la cui raccolta ordini parte a giugno e prosegue fino a novembre, riguarda tutti gli articoli che vengono presentati nei negozi da gennaio a maggio;
- la campagna vendite *Fall-Winter* invece (codificata FW), la cui raccolta ordini parte a dicembre e si protrae fino a maggio, fa riferimento a tutti quei prodotti che vengono venduti da giugno a novembre.

Una *collezione* è per definizione una gamma di prodotti che l'azienda produttrice studia, realizza e propone al mercato per una determinata stagione e che verrà eventualmente

rinnovata nelle stagioni successive in risposta alle evoluzioni della moda, alle tendenze del mercato e ai progressi nella tecnologia di prodotto.

Una *campagna vendite* consiste invece nell'arco temporale durante il quale l'azienda produttrice svolge determinate azioni commerciali per presentare e promuovere la collezione al mercato e durante il quale avvengono la raccolta ordini dei clienti e successivamente la consegna della merce venduta.

Il *lead time* complessivo per lo sviluppo di una collezione è un anno (52 settimane), durante il quale avviene una continua attività trasversale di raccolta informazioni sul mercato, il tutto strutturato in diverse fasi.

1.5 - La Business Unit EMEA

L'analisi è stata effettuata considerando l'andamento delle vendite della Business Unit EMEA (*Europe, Middle-East e Africa*), l'area strategica d'affari più grande ed importante dell'azienda sia dal punto di vista delle vendite che del fatturato con circa l'83% del totale dell'azienda (anno 2010).

<u>Business Unit</u>	<u>Fatturato</u>	<u>Fatturato %</u>
BU EMEA	131.215.084,74	83,70%
BU RETAIL	12.186.826,06	7,77%
BU WORKS	6.241.713,05	3,98%
BU LEISURE	3.216.151,85	2,05%
BU ETONIC	2.022.346,78	1,29%
BU AMERICAS	847.104,58	0,54%
BU LE DD	591.815,01	0,38%
BU MARKETING	357.369,34	0,23%
BU ASIA/PACIFIC	93.948,45	0,06%
BU OPERATION	4.370,40	0,00%

1.5.1 – Ripartizione Fatturato B.U. EMEA

La ripartizione del fatturato di questa divisione per le tre tipologie di prodotto è così strutturata: le scarpe rappresentano il 43% del totale, gli accessori solamente il 5%, mentre è l'abbigliamento il vero punto di forza dell'azienda anche nella Business Unit EMEA, che fattura ben il 52% del totale. È quindi notevole la trasformazione che nel tempo ha subito l'offerta dell'azienda: nata puntando sulla scarpa da tennis, e quindi

partendo dal settore *performance*, è cresciuta sviluppando importanti linee di abbigliamento e prodotti del settore *leisure*.

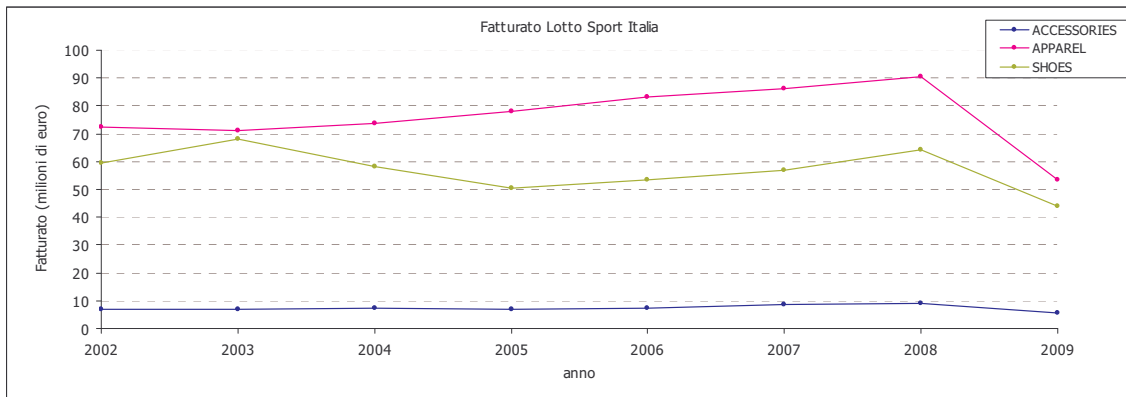


Figura 1.11 – Fatturato Lotto nel tempo per tipo prodotto

1.5.2 – Planning & Controlling Analysis

L'esperienza di stage in Lotto si è realizzata presso l'ufficio *Planning & Controlling Analysis*, organo di supporto statistico per l'intera azienda, che si occupa tra le varie cose anche dell'effettuazione delle analisi statistiche sui dati di ordinato e di fatturato dell'azienda, i quali costituiscono la base informativa necessaria per mettere in atto gli svariati processi di cui esso direttamente si occupa e che sono:

- Il processo di *forecasting*, ossia di previsione delle vendite relative ad ogni collezione, utile per pianificare la produzione con il giusto anticipo ed evitare il rischio di un accumulo eccessivo di merce nel magazzino.
- Il processo di *targeting*, il quale porta alla fissazione degli obiettivi di vendita (coerenti con il *budget*) da assegnare ad ogni responsabile di BU, il quale poi distribuirà il proprio *target* tra le varie aree geografiche che gestisce.
- Il processo di *reporting*, il quale consente la predisposizione di documenti prospettici contenenti informazioni di vario tipo riguardanti le attività svolte dall'azienda (come l'andamento delle vendite, la gestione del magazzino, ecc) utili al management in fase decisionale.

- Il processo di *pricing* il quale porta alla definizione dei prezzi di vendita da applicare a ciascun prodotto di ogni collezione e che coinvolge anche la gestione attenta del complesso universo dei listini prodotto dell'azienda.

Essendo un organo preposto al supporto informativo risulta evidente che il parco dati a disposizione degli analisti che vi lavorano deve essere in assoluto il più ampio possibile.

CAPITOLO II – METODI E OBIETTIVI DELLE PREVISIONI STATISTICHE

2.1 – La previsione in azienda

Per le aziende le previsioni sono uno strumento indispensabile al fine di prevederne l'andamento e progettare uno sviluppo che permetta una maggiore efficienza. In particolare le previsioni delle vendite hanno una rilevanza operativa nelle decisioni razionali quali l'organizzazione delle risorse e delle funzioni aziendali come gli acquisti, il personale, la produzione. Per quanto riguarda invece un'ottica di lungo periodo le previsioni delle vendite permettono di organizzare le risorse e i piani di investimento.

Nello specifico le previsioni in azienda valgono come strumento che supporta le decisioni e indirizzano i responsabili verso le scelte più indicate in base ai dati di vendita passati. Infatti questo strumento permette un'inferenza sui valori futuri del fenomeno sotto studio che si basa sull'andamento della sua storia passata; si andrà quindi cercando una regola che interpreti l'andamento del fenomeno (popolazione) in base ai dati a nostra disposizione (campione) e con questa si ipotizzano dei valori futuri. Il problema principale sta dunque nel dedurre una funzione che tenga conto delle variabili più rilevanti per il fenomeno in questione senza però comprometterne la fruibilità e l'aggiornamento ogniqualvolta si dispongano di nuovi e più recenti dati. La previsione ha quindi fini di orientamento e di decisione strategica ma bisogna però tener presente che in un qualunque modello costruito per l'analisi di un fenomeno, sia esso economico, sociale o di altra natura, è impossibile includere la totalità delle variabili e delle relazioni necessarie per una completa spiegazione o interpretazione dei dati. È però possibile supporre che le informazioni riguardanti il passato, adeguatamente elaborate, forniscano un valido strumento che permetta di ridurre l'incertezza delle previsioni, sempre consci del fatto che non sarà possibile una fedele ed esatta ripetizione degli accadimenti passati.

Nelle aziende le previsioni sono uno strumento che influenza in maniera evidente lo sviluppo di un buon *management* e un efficiente apparato di controllo di gestione perché devono garantire delle politiche di produzione e commercializzazione adeguate agli obiettivi posti dall'azienda. Perciò le previsioni consistono nella stima dell'evoluzione futura di un fenomeno a partire dallo stato attuale e tenendo presente la sua evoluzione precedente. Proprio per queste loro caratteristiche, i piani di previsione permettono di coordinare le varie attività interne all'azienda e mediarne gli obiettivi, dalla produzione al

marketing, dalla direzione finanziaria a quella amministrativa. La previsione è dunque un elemento dinamico che collega e riflette le tendenze in atto e ha quindi un ruolo fondamentale nel processo strategico perché permette una gestione dinamica e continua di tutta l'azienda fornendo delle indicazioni sul più probabile andamento nell'immediato futuro. Questo processo di previsione interagisce, come detto, con tutte le funzioni aziendali e in particolare con i responsabili commerciali i quali vengono a conoscenza dell'andamento prossimo delle vendite, i responsabili di produzione per ottimizzare i processi di produzione e gli addetti all'amministrazione che gestiranno così in modo ottimale il *budget*; grazie alla possibilità di aggiornamento continuo, permettono inoltre un controllo strategico di ogni settore aziendale.

Una corretta determinazione delle vendite future riduce inutili investimenti nella produzione di quantità non vendibili o, viceversa, il rischio di esaurimenti delle scorte con conseguente perdita di competitività nel mercato e quindi di profitto. La fluidità delle disponibilità all'interno dell'azienda rappresenta la sua forza e la sua vitalità e la successiva trasformazione in prodotto finito, se non successivamente venduto, rappresenta un'inevitabile perdita di profitto. Questa semplice situazione basta a rappresentare l'importanza della previsione delle vendite in azienda, ancor più in un'azienda come Lotto Sport Italia dove, come abbiamo già visto, la merce che viene immessa nel mercato viene acquistata già sotto forma di prodotto finito.

I piani di previsione vengono elaborati da più membri all'interno dell'azienda, poiché i soli metodi statistici di previsione, scientifici ed oggettivi, non sono sufficienti in quanto tali, ma vanno integrati da fattori soggettivi quali l'esperienza e l'intuito dei manager che adattano la situazione ad eventuali cambiamenti ambientali. Quindi, la quantificazione di giudizi qualitativi ci consente di aggiungere l'indispensabile elemento di "novità" che un'impostazione puramente statistica ha maggiore difficoltà ad accogliere. Ecco perché le previsioni statistiche, pur essendo un buon punto di partenza a supporto dei processi decisionali, necessitano di un'integrazione da parte di manager e figure aziendali preposte.

2.2 – Le serie storiche

Per serie storica (o temporale) si intende un insieme di dati ordinati secondo il tempo, in

altre parole una successione di dati numerici nella quale ogni dato è associato ad un particolare istante od intervallo temporale (Di Fonzo, Lisi, 2005). Se in ciascun punto o intervallo del tempo viene osservato un solo fenomeno allora la serie che ne deriva viene detta univariata, se invece le variabili osservate sono più d'una si ha una serie multivariata.

Una serie storica univariata Y_t ($t=1,2, \dots, n$) è quindi una successione ordinata di osservazioni relative ad un certo fenomeno Y_t , esaminato con riferimento alla sua evoluzione rispetto la variabile t , il tempo. L'ordinamento temporale, nello studio delle serie storiche, assume un ruolo fondamentale, infatti è lecito presumere che vi sia dipendenza tra osservazioni successive e che tale dipendenza sia legata alla posizione dell'osservazione nella sequenza.

Una serie storica può essere di tipo deterministico, nel senso che può essere prevista in modo esatto sulla base della propria storia passata, ma la maggior parte delle serie storiche è di tipo stocastico, per queste ultime il futuro viene determinato dal passato solo in misura parziale, rendendo quindi impossibile l'elaborazione di previsioni del tutto prive d'errore.

Un modello statistico rappresenta quindi un tentativo di descrivere la legge statistica sottostante i dati e generatrice degli stessi ed ha una duplice funzione: sia quella di generare i valori per i momenti passati e presenti (la cosiddetta funzione di *adattamento* del modello ai dati esistenti della serie) sia quella di generare dei valori per il futuro (la funzione *previsiva*).

Si presenta dunque il concetto di previsione: la previsione di un dato fenomeno è la stima, puntuale o intervallare, dei valori futuri che il fenomeno può assumere, costruita sulla base dei valori passati e presenti esplicitati sotto forma di modello. La bontà delle previsioni è dunque strettamente dipendente dal modello scelto.

2.2.1 – Approccio classico e approccio moderno

Un modello stocastico può essere generalmente descritto da un processo generatore dei dati di una serie storica Y_t ($t=1,2,\dots,n$) relativa ad una variabile Y_t è dato da $Y_t = f(t) + u$ dove si assume che la serie osservata sia la composizione di:

- Una componente totalmente deterministica $f(t)$ che costituisce la parte sistematica della serie
- Una sequenza di variabili casuali u_t che rappresenta la parte stocastica della serie ed obbedisce ad una data legge di probabilità.

Nell'approccio classico allo studio delle serie storiche si suppone esista una legge di evoluzione temporale del fenomeno, rappresentata da $f(t)$. La componente casuale u_t viene invece assunta di entità trascurabile e non viene considerata all'interno di $f(t)$. I residui di Y_t , che non saranno spiegati dal modello, vengono quindi assimilati come errori accidentali. Questo equivale ad ipotizzare che la componente stocastica del modello sia generata da un processo *white noise*, ossia una successione di variabili casuali incorrelate a media zero e varianza costante: $\varepsilon \sim WN(0, \sigma^2)$. In questo approccio si concentra quindi tutta l'attenzione su $f(t)$, avendo considerato u_t trascurabile.

Per alcuni tipi di serie però questo tipo di approccio classico non è adeguato perché può non essere presente nessuna delle componenti di $f(t)$ che descriveremo nella prossima sezione, oppure perché, qualora presenti, tali componenti potrebbero avere andamenti tali da non poter essere modellati in questo modo. In questi casi è più utile cercare di modellare la componente stocastica ipotizzando che il meccanismo che genera i dati sia governato da regole probabilistiche (Di Fonzo, Lisi, 2005). Questo modo di approcciarsi alla serie si definisce approccio moderno all'analisi delle serie storiche secondo il quale viene trascurata la parte deterministica $f(t)$ e con un opportuno modello della componente stocastica u_t è possibile estrarre dalle serie ulteriori informazioni, soprattutto a fini previsivi.

2.2.2 – L'approccio classico

Con *l'approccio classico* all'analisi delle serie storiche si assume che la parte sistematica $f(t)$ sia il risultato dell'azione congiunta delle seguenti quattro componenti, non direttamente osservabili, ma chiaramente definibili:

- il *trend*, o componente tendenziale, è la tendenza di fondo del fenomeno in osservazione, che si nota nel lungo periodo.

- il *ciclo* è costituito dalle fluttuazioni attorno al trend attribuibili alle espansioni e contrazioni congiunturali dovute all'attività economica; spesso il ciclo viene assimilato al trend e si studiano in modo congiunto come un'unica componente.
- la *stagionalità* è composta da tutti i movimenti del fenomeno che nel corso dell'anno tendono a ripetersi in modo quasi analogo nello stesso periodo in anni successivi.
- l'*errore* è invece la componente irregolare del fenomeno in questione, che raccoglie tutte le variazioni dovute alle cause più particolari, che non fanno parte di quelle elencate in precedenza. Tutte le serie storiche che non siano di tipo puramente deterministico presentano delle irregolarità di segno positivo e negativo, prodotte da un comportamento di tipo aleatorio.

Queste componenti possono essere combinate in diversi modi di aggregazione, i più semplici sono:

modello additivo

$$Y_t = T_t + C_t + S_t + \varepsilon_t$$

modello moltiplicativo

$$Y_t = T_t * C_t * S_t + \varepsilon_t$$

modello misto

$$Y_t = T_t * C_t * S_t + \varepsilon_t$$

Il *modello moltiplicativo* può essere ricondotto a quello additivo con l'applicazione di una trasformazione logaritmica: $\log Y_t = \log T_t + \log C_t + \log S_t + \log \varepsilon_t$.

Occorre quindi saper stabilire il numero delle componenti, il metodo con cui queste agiscono e si combinano e il metodo per l'enucleazione di ciascuna componente; scopo dell'analisi è infatti quello di isolare ognuna di esse per poi poterla studiare con più precisione. Tra i metodi che cercano di isolare queste componenti il più semplice è rappresentato dalle medie mobili, ovvero trasformazioni lineari delle osservazioni. Le medie mobili vengono solitamente scelte per annullare il trend e conservare la stagionalità o viceversa ma in ogni caso devono eliminare o perlomeno ridurre la componente erratica (Di Fonzo, Lisi, 2005). Se, ad esempio, si applica una *media mobile* semplice di ordine S ad una serie osservata contenente un'onda di periodo S , l'onda viene eliminata. Allo stesso modo, considerando una serie storica di osservazioni mensili e applicandovi una *media mobile* centrata di 13 termini, si ottiene una stima della componente di trend-ciclo. Le medie mobili però non permettono di fare previsioni e comportano inoltre una perdita di osservazioni alle estremità della serie.

Un altro metodo, il metodo del lisciamiento esponenziale (*exponential smoothing*) è una particolare applicazione del metodo delle medie mobili che risulta però maggiormente interessante e completo visto che sfrutta un sistema di pesi da assegnare alle singole osservazioni che permettono di sfruttare tutti i dati del passato, anche quelli più lontani.

2.2.3 – Il lisciamiento esponenziale (Di Fonzo, Lisi, 2005)

Il lisciamiento esponenziale è un metodo basato su medie mobili ponderate (vengono applicati dei pesi, che decrescono man mano che ci si allontana nel tempo) di tutti i dati del passato e non solo ai k valori come per le medie mobili semplici. Viene chiamato esponenziale perché i pesi sono definiti dalla successione esponenziale: $c_j = \alpha \cdot (1 - \alpha)^j$,

$j=0, 1, 2, \dots$, (con vincolo $\sum_{j=0}^{\infty} c_j = 1$) con α costante di lisciamiento ($0 < \alpha < 1$). Se per

esempio si vuole stimare una serie non stagionale e senza trend sistematico, scopo del lisciamiento è quello di stimare il livello della serie e di utilizzare la stima per prevedere i valori futuri.

$$\hat{y}_{t+1|t} = c_0 y_t + c_1 y_{t-1} + c_2 y_{t-2} + \dots = \alpha \cdot \sum_{j=0}^{t-1} (1 - \alpha)^j y_{t-j}$$

Poiché si può scrivere $\hat{y}_{t+1|t} = \alpha \cdot y_t + (1 - \alpha) \cdot \hat{y}_{t|t-1}$ la nuova previsione può essere pensata come la media pesata tra l'ultima osservazione disponibile e la vecchia previsione.

Il valore di α determina quanto influisce l'osservazione corrente sulla previsione futura quindi la sua scelta è molto rilevante:

- se $\alpha \approx 1$ allora la previsione sarà maggiormente influenzata dai dati recenti mentre allontanandoci nel tempo i pesi decrescono, dunque la previsione sarà molto simile alla precedente osservazione
- se $\alpha \approx 0$ allora la memoria del processo è molto lunga e la previsione dipenderà da tutto il processo osservato, anche dalle osservazioni molto remote.

Se l'obiettivo è una previsione basata su una serie con dati molto variabili, si darà più peso a ciò che è appena accaduto mentre se si vuole una previsione stabile con variazioni casuali che vengano lisce si darà peso a tutta la serie.

Una generalizzazione del metodo di lisciamento esponenziale sono i metodi di Holt-Winters, che possono essere stagionali o non stagionali. Il metodo non stagionale si fonda sull'ipotesi che un aggiustamento in prossimità di n mediante una retta di equazione $L_n + T_n(t-n)$ sia preferibile alla perequazione fornita da una costante. Indicando con \hat{L}_n e \hat{T}_n le stime, rispettivamente, dell'intercetta e del coefficiente angolare della retta, assunti a rappresentare il livello e il trend della serie nel periodo n , le formule di aggiornamento sono date da

$$\begin{aligned}\hat{L}_n &= \alpha(\hat{L}_{n-1} + \hat{T}_{n-1}) + (1-\alpha)y_n \\ \hat{T}_n &= \beta\hat{T}_{n-1} + (1-\beta)(\hat{L}_n - \hat{L}_{n-1})\end{aligned}\quad \text{con } 0 < \alpha < 1, 0 < \beta < 1$$

La previsione è quindi data da $\hat{y}_{n+k} = \hat{L}_n + \hat{T}_n k$, $k \geq 1$.

Il metodo di Holt-Winters stagionale può essere invece additivo o moltiplicativo, da scegliere solitamente secondo i casi illustrati in figura 2.1. Per quanto riguarda il metodo additivo, la serie, in prossimità di n , è esprimibile come $L_n + T_n(t-n) + S_n$ dove S_n è un fattore stagionale di periodo pari a S . Indicata con \hat{S}_n la stima di tale fattore, le formule di aggiornamento diventano

$$\begin{aligned}\hat{L}_n &= \alpha(\hat{L}_{n-1} + \hat{T}_{n-1}) + (1-\alpha)(y_n - \hat{S}_{n-s}) \\ \hat{T}_n &= \beta\hat{T}_{n-1} + (1-\beta)(\hat{L}_n - \hat{L}_{n-1}) \\ \hat{S}_n &= \gamma\hat{S}_{n-s} + (1-\gamma)(y_n - \hat{L}_n)\end{aligned}\quad \text{con } 0 < \alpha < 1, 0 < \beta < 1, 0 < \gamma < 1$$

e la previsione è data da $\hat{y}_{n+k} = \begin{cases} \hat{L}_n + \hat{T}_n k + \hat{S}_{n+k-s} & \text{se } 1 \leq k \leq s \\ \hat{L}_n + \hat{T}_n k + \hat{S}_{n+k-2s} & \text{se } s+1 \leq k \leq 2s \end{cases}$.

Il metodo moltiplicativo prevede invece che la serie sia data, in prossimità di n , da $[L_n + T_n(t-n)] \cdot S_n$ e le formule di aggiornamento sono

$$\begin{aligned}\hat{L}_n &= \alpha(\hat{L}_{n-1} + \hat{T}_{n-1}) + (1-\alpha)\frac{y_n}{\hat{S}_{n-s}} \\ \hat{T}_n &= \beta\hat{T}_{n-1} + (1-\beta)(\hat{L}_n - \hat{L}_{n-1}) \\ \hat{S}_n &= \gamma\hat{S}_{n-s} + (1-\gamma)\frac{y_n}{\hat{L}_n}\end{aligned}\quad \text{con } 0 < \alpha < 1, 0 < \beta < 1, 0 < \gamma < 1$$

quindi la previsione è data da $\hat{y}_{n+k} = \begin{cases} (\hat{L}_n + \hat{T}_n k) \cdot \hat{S}_{n+k-s} & \text{se } 1 \leq k \leq s \\ (\hat{L}_n + \hat{T}_n k) \cdot \hat{S}_{n+k-2s} & \text{se } s+1 \leq k \leq 2s \end{cases}$

Il coefficiente β è la costante di lisciamento relativa al trend, mentre γ quella relativa alla stagionalità; ognuno di essi va ad influenzare la rispettiva componente (tendenziale e stagionale) come fa α con il livello: rappresentano il peso che si vuole dare all'ultima osservazione disponibile nella determinazione della previsione. Tutti i parametri hanno dominio $[0,1]$.

Per una giusta scelta del modello è fondamentale quindi una prima analisi grafica della serie dei dati per individuare la presenza o meno delle varie componenti e possibilmente del modo con cui queste interagiscono tra loro. La figura 2.1 illustra quali sono i modelli principali.

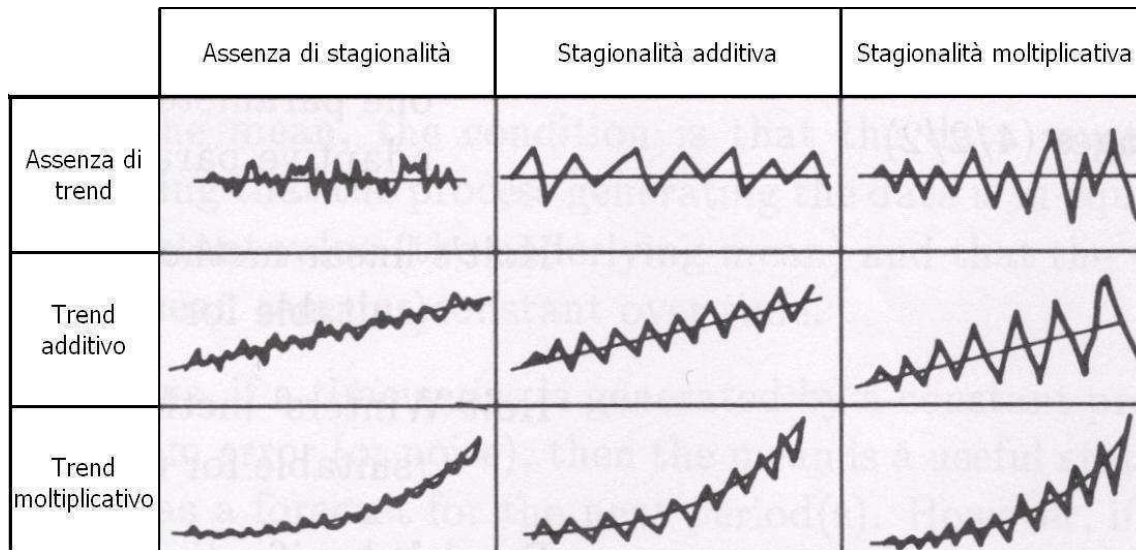


Figura 2.1 – I principali modelli di lisciamento esponenziale (Makridakis, Hyndman, 1998)

2.2.4 – L'approccio moderno

Nell'approccio moderno all'analisi delle serie storiche si sposta l'interesse da $f(t)$ a u_t supponendo che sia possibile estrarre un qualche tipo di informazione dalla componente stocastica. Proprio per questo motivo u_t non viene più assunto come processo *white noise* ma, piuttosto, come processo a componenti correlate, per il quale cioè $Cov(u_r, u_s) \neq 0$ per qualche $r \neq s$. Ne deriva che l'obiettivo non è più quello di stimare le componenti di una serie ma individuare un modello probabilistico che descriva l'evoluzione del fenomeno, modello che può essere usato sia per fini descrittivi che previsivi (Di Fonzo, Lisi, 2005). Per

definire un processo stocastico bisogna considerare uno spazio di probabilità (Ω, F, P) , dove Ω è lo spazio campionario, F una σ -algebra su Ω e P una misura di probabilità. Dato uno spazio parametrico T , si definisce processo stocastico una funzione finita e a valori reali di $\omega \in \Omega$ e $t \in T$ tale che, per ogni t , $Y_t(\omega)$ è una funzione misurabile di ω . Inoltre:

- per ogni $\omega = \omega_0$ fissato, $Y_t(\omega_0)$ è una funzione di t detta realizzazione (o traiettoria) del processo stocastico;
- per ogni $t = t_0$ fissato, $Y_{t_0}(\omega)$ è una funzione misurabile di $\omega \in \Omega$ e, quindi, una variabile casuale;
- per $t = t_0$ e $\omega = \omega_0$, $Y_{t_0}(\omega_0)$ è un numero reale.

Una serie storica è dunque una parte finita di una singola realizzazione del processo, quindi non solo costituisce un campione unico della famiglia di variabili casuali che costituiscono il processo, ma è anche un campione troncato perché si osserva per $t = 1, 2, \dots, n$. Come per l'approccio classico, anche in questo caso l'obiettivo dell'analisi è individuare il processo generatore, cioè l'ipotetico processo stocastico che genera i dati della serie, attraverso lo studio della serie stessa, fino ad arrivare alla formulazione del modello statistico.

Per la descrizione completa di un processo stocastico occorrerebbe specificare la distribuzione di probabilità congiunta di $(Y_{t_1}, \dots, Y_{t_n})$ per ogni insieme di istanti t_1, \dots, t_n ed ogni valore di n , ma data la difficoltà di ciò, si preferisce definire un processo stocastico sulla base dei suoi momenti, ed in particolare i primi due momenti della v.c. Y_t :

Funzione Media: $\mu = E(Y_t)$

Funzione Varianza: $\sigma^2 = \text{Var}(Y_t) = E(Y_t - \mu_t)^2$

Funzione di Autocovarianza: $\gamma(t_1, t_2) = E[(Y_{t_1} - \mu_{t_1})(Y_{t_2} - \mu_{t_2})]$

Mentre il significato delle funzioni media e varianza di una v.c. è già ben noto ed estendibile al caso di processi stocastici, è doveroso soffermarsi sul senso della funzione autocovarianza: non è altro che la covarianza tra v.c. del medesimo processo stocastico spaziate tra loro di uno sfasamento temporale pari a $k = t_1 - t_2$ (Di Fonzo, Lisi, 2005). Affinché risulti meglio interpretabile, di solito viene normalizzata e si ottiene così la *funzione di Autocorrelazione (ACF, AutoCorrelation Function)*. È definita per ogni coppia

t_1, t_2 come $\rho = \frac{\gamma(t_1, t_2)}{\sigma_{t_1} \cdot \sigma_{t_2}}$. Essa rappresenta una misura della correlazione tra Y_{t_1} e Y_{t_2} ed ha il vantaggio, rispetto all'autocovarianza, di essere indipendente dall'unità di misura di Y_t .

2.2.5 – Processi Stazionari

Un processo stocastico è detto stazionario quando la sua media e la sua varianza non presentano cambiamenti sistematici con il passare del tempo. Più formalmente un processo stocastico si dice stazionario in senso stretto se la distribuzione di probabilità di $(Y_{t_1}, \dots, Y_{t_n})$ è uguale alla distribuzione di probabilità di $(Y_{t_1+k}, \dots, Y_{t_n+k})$ per ogni n e k interi. La stazionarietà in senso stretto è però una condizione troppo forte quindi ci si basa solamente (e lo si farà da qui in poi) solo sulla stazionarietà in senso debole, detta anche stazionarietà di secondo ordine, la quale richiede che per qualsiasi t e k :

$$E(Y_t) = \mu < \infty$$

$$Var(Y_t) = \sigma^2 < \infty$$

$$\gamma(Y_t, Y_{t+k}) = \gamma(k)$$

$$\rho(Y_t, Y_{t+k}) = \rho(k)$$

L'autocovarianza assume un ruolo fondamentale come indicatore delle relazioni lineari esistenti tra v.c., in particolare nel caso dei processi stocastici stazionari, ha le seguenti proprietà (Wei, 1990):

$$1) \rho(0)=1$$

$$2) -1 < \rho(k) < 1$$

$$3) \rho(k) = \rho(-k)$$

$$4) \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \rho(|t_i - t_j|) \geq 0$$

la prima proprietà evidenzia che ciò che accade oggi è correlato con ciò che accade oggi (quindi autocorrelato), mentre la seconda indica l'intervallo di valori che l'autocorrelazione può assumere. Grazie alla terza proprietà, secondo cui l'autocorrelazione è una funzione pari, si può limitare il calcolo ai soli valori positivi di k ; l'ultima proprietà invece dice che l'ACF è una funzione semidefinita positiva per ogni valore di t_1, t_2, \dots, t_n e per ogni $\alpha \in \mathbb{R}$. La

correlazione tra due variabili può essere dovuta al fatto che esiste effettivamente un legame lineare diretto tra loro o al fatto che queste sono correlate con una terza variabile. Per poter quindi tener conto di ciò si può prendere in considerazione, a fianco alla ACF globale, la *funzione di autocorrelazione parziale (PACF, Partial AutoCorrelation Function)*, che misura l'autocorrelazione tra Y_t e Y_{t-k} al netto delle variabili intermedie: $P(k) = \text{Corr}(Y_t, Y_{t-k} | Y_{t-1}, \dots, Y_{t-k+1})$.

2.2.6 – Processi invertibili

Dato un processo stazionario $\{Y_t\}$ è possibile calcolare in modo univoco la sua funzione di autocovarianza $\gamma(k)$ e quindi autocorrelazione $\rho(k)$; data però $\gamma(k)$ il processo $\{Y_t\}$ che la possiede è unico (si ha cioè una relazione biunivoca) solo se oltre alla proprietà di stazionarietà il processo stocastico è anche invertibile. Infatti tra i processi con la stessa funzione di autocovarianza uno solo è anche invertibile. La condizione di invertibilità riguarda la possibilità di esprimere un processo $\{Y_t\}$ come funzione delle v.c. del passato. Più formalmente, un processo $\{Y_t\}$ è detto invertibile se esistono una funzione lineare $h(\cdot)$ ed un processo *white noise* $\{\varepsilon_t\}$ tali che, per ogni t , si ha $Y_t = h(Y_{t-1}, Y_{t-2}, \dots) + \varepsilon_t$

2.2.7 – Le funzioni di autocorrelazione empiriche

Dopo aver definito i processi stocastici e le relative funzioni di autocorrelazione (*ACF*) ed autocorrelazione parziale (*PACF*), verrà ora definito il correlogramma, utile strumento per lo studio dei processi stocastici stazionari: il correlogramma è il grafico dei valori $\rho(k)$ per $k=0,1,2,..$ e riproduce graficamente l'ACF di un processo stocastico mentre quando riporta i valori di P_k ne riproduce la PACF.

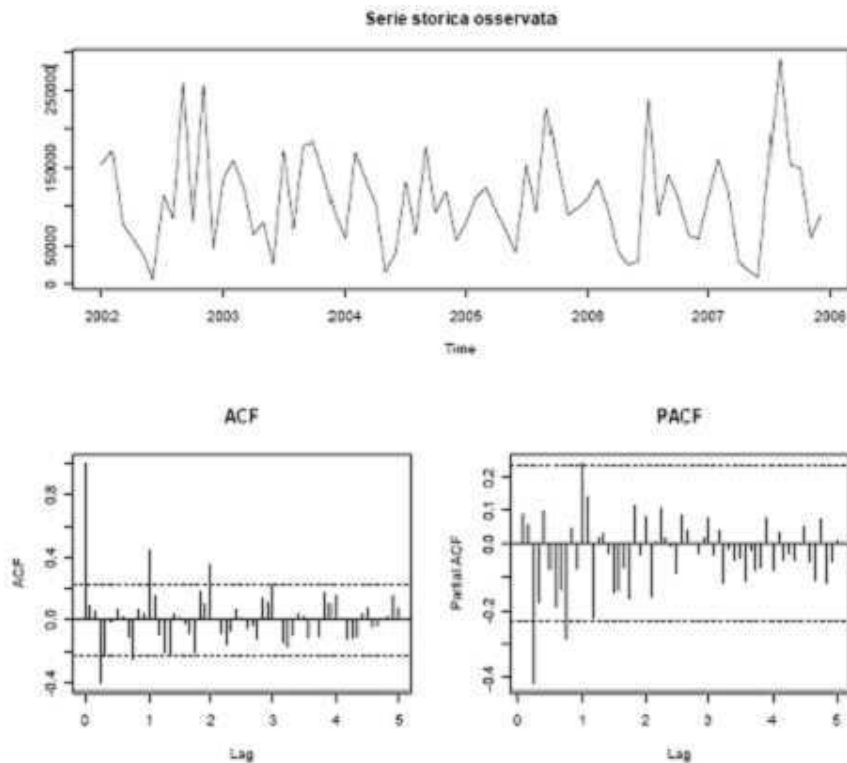


Figura 2.2 - Serie osservata e relativo correlogramma di ACF e PACF

2.2.8 – I principali processi stocastici (Di Fonzo, Lisi, 2005)

Verranno di seguito illustrati brevemente i principali processi stocastici che porteranno alla definizione della più ampia classe di modelli SARIMA (*Seasonal AutoRegressive Integrated Moving Average*).

PROCESSO WHITE-NOISE

Il processo *white noise*, già accennato precedentemente, è uno dei più importanti processi stocastici. Si tratta di un processo $\{\varepsilon_t\}$ che consiste in una successione di v.c. incorrelate a media nulla e varianza costante. Dato che le componenti di questo processo sono serialmente incorrelate, la sua funzione di autocorrelazione parziale, P_k , è identicamente uguale a $\rho(k)$, quest'ultima pari a :

$$\rho = \begin{cases} 1 \rightarrow k = 0 \\ 0 \rightarrow k = \pm 1, \pm 2, \dots \end{cases}$$

Dato che, da risultati noti, per l'autocorrelazione empirica di un processo *white noise* vale il risultato $\sqrt{n} \cdot \hat{\rho} \sim N(0,1)$ e che quindi si può considerare $\hat{\rho} \sim N\left(0, \frac{1}{n}\right)$, ciò può essere utilizzato a fini pratici per verificare se i residui di un modello stimato sono incorrelati. Il rifiuto di tale ipotesi, ad un opportuno livello di significatività, comporta solitamente l'abbandono dell'intero modello in quanto i residui non possono ritenersi imputabili solamente al caso.

PROCESSO A MEDIA MOBILE MA(q)

Sia $\{\varepsilon_t\}$ un processo *white noise* a media zero e varianza σ_ε^2 . Un processo a *media mobile* (*Moving Average*) di ordine q è definito come $Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$ con θ_j ($j=1, 2, \dots, q$) parametri costanti. Y_t è dunque il risultato di una somma pesata di impulsi casuali presenti e passati. Utilizzando l'operatore di ritardo B , tale che $B^h Y_t = Y_{t-h}$, $h=0, 1, 2, \dots$ e indicando l'operatore polinomiale a *media mobile* con $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ il modello può essere brevemente critto come $Y_t = \theta(B) \varepsilon_t$.

Poiché media, varianza e covarianza sono costanti e non dipendono dal tempo, un processo MA(q) è sempre stazionario.

Esso è invertibile se tutte le q radici dell'equazione caratteristica $\theta(B) = 0$ sono, in modulo, maggiori di 1.

Per un processo MA(q) l'ACF è diversa da zero per $k \leq q$ e si annulla per $k \geq q+1$ mentre l'autocorrelazione parziale tende a zero con una modalità che dipende dal valore dei parametri θ_j .

PROCESSO AUTOREGRESSIVO AR(p)

Sia $\{\varepsilon_t\}$ un processo *white noise* a media zero e varianza σ_ε^2 . Un processo autoregressivo (*AutoRegressive*) di ordine p è definito come $Y_t = \phi_0 + \phi_1 \cdot Y_{t-1} + \dots + \phi_p \cdot Y_{t-p} + \varepsilon$ con

Φ_i ($i=0, 1, \dots, p$) parametri costanti. Y_t è quindi il risultato di una somma pesata di valori passati e di uno shock casuale contemporaneo. Il polinomio autoregressivo è $\Phi(B)=(1-\Phi_1(B)-\dots-\Phi_p(B^p))$ e quindi $\Phi(B)Y_t = \Phi_0 + \varepsilon_t$.

Per un processo $AR(p)$ l'ACF tende a zero con un comportamento che dipende dal valore dei parametri Φ_i mentre la PACF è diversa da zero per $k \leq p$ e si annulla per $k \geq p+1$.

Il processo $AR(p)$ è sempre invertibile, è invece stazionario se tutte le p radici dell'equazione caratteristica $\Phi(B)=0$ sono, in modulo, maggiori di uno.

PROCESSO AUTOREGRESSIVO A MEDIA MOBILE ARMA(p,q)

Sia $\{\varepsilon_t\}$ un processo *white noise* a media zero e varianza σ_ε^2 . Un processo autoregressivo a *media mobile* (*AutoRegressive Moving Average*) di ordine (p,q) è definito come:

$$\begin{aligned} \phi(B) \cdot Y_t &= \phi_0 + \theta(B) \cdot \varepsilon_t \\ (1 - \phi_1 \cdot B - \phi_2 \cdot B^2 - \dots - \phi_p \cdot B^p) \cdot Y_t &= \phi_0 + (1 - \theta_1 \cdot B - \dots - \theta_q \cdot B^q) \cdot \varepsilon_t \\ Y_t - \phi_1 \cdot Y_{t-1} - \dots - \phi_p \cdot Y_{t-p} &= \phi_0 + \varepsilon_t - \theta_1 \cdot Y_{t-1} - \dots - \theta_q \cdot Y_{t-q} \end{aligned}$$

Un processo $ARMA(p,q)$ è stazionario se tutte le p radici dell'equazione caratteristica $\phi(B)=0$ sono in modulo maggiori di uno ed è invertibile se le q radici dell'equazione caratteristica $\theta(B)=0$ sono in modulo maggiori di uno. Per un processo $ARMA(p,q)$, come per un $AR(p)$, l'ACF tende ad annullarsi per $k > q$. La sua PACF ha invece una forma piuttosto complicata ma è sufficiente notare che, per $k > q$, la PACF tende a zero in maniera esponenziale come per un processo $MA(q)$.

2.2.8.1 – Processi non stazionari

I modelli di serie storiche fin qui considerati sono tutti stazionari. Tuttavia, soprattutto per variabili economiche, è piuttosto frequente il caso di serie per le quali è ragionevole ipotizzare un processo generatore non stazionario. Limitandoci ai primi due momenti, la non stazionarietà può dipendere dal fatto che la media non è costante e/o che non sono

costanti i momenti secondi. Nel primo caso si parla di non stazionarietà in media, che tipicamente si verifica in presenza di un trend, nel secondo caso si parla di non stazionarietà in varianza (o covarianza). Alcuni tipi di non stazionarietà possono essere eliminati mediante differenziazione ma è possibile che processi stazionari in media non lo siano in varianza. Questo tipo di problema viene generalmente risolto mediante altri tipi di trasformazione, volte a stabilizzare la varianza, in particolare si può far uso delle trasformazioni introdotte da Box e Cox (1964):

$$Y_t = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln Y_t & \lambda = 0 \end{cases}$$

di cui la radice quadrata ($\lambda=1/2$) e la trasformazione logaritmica ($\lambda=0$) rappresentano due casi particolari.

PROCESSO AUTOREGRESSIVO INTEGRATO A MEDIA MOBILE ARIMA(p,d,q)

Sia $\{\varepsilon_t\}$ un processo *white noise* a media zero e varianza σ_ε^2 . Si indichi con X_t la d-esima differenza di Y_t , $X_t=(1-B^d)Y_t$. Si dice che $\{Y_t\}$ è un processo auto regressivo integrato a *media mobile (Auto Regressive Integrated Moving Average)* di ordine (p,d,q) se $\{X_t\}$ è un processo ARMA(p,q). In sintesi: $\phi(B) \cdot (1-B)^d \cdot Y_t = \phi_0 + \theta(B) \cdot \varepsilon_t$.

2.2.8.2 – Processi stagionali

I processi fino a qui presentati non permettono di modellare andamenti periodici di tipo stagionale, se non appesantendo il modello con un gran numero di parametri. La gran parte delle serie storiche però, soprattutto quelle di natura economica, come le serie di vendite che sono soggette ad analisi in questo elaborato, presentano una componente di questo tipo. Spesso succede che la stagionalità sia stocastica e correlata con le

componenti non stagionali, per questo motivo Box e Jenkins (1976) hanno esteso la classe dei processi ARIMA in modo da trattare anche comportamenti di tipo periodico.

PROCESSO STAGIONALE ARIMA(p,d,q)x(P,D,Q)_s

I modelli SARIMA hanno l'obiettivo di poter descrivere due tipi di relazioni all'interno della serie osservata: la correlazione tra valori consecutivi, che è quella modellata dagli usuali modelli ARIMA, e la correlazione tra osservazioni che distano tra loro di un multiplo del periodo. Combinando tra loro due modelli ARIMA(p,d,q), in particolare il primo $\phi(B)(1-B)^d Y_t = \theta(B)b_{i,t}$, dove $b_{i,t}$ non costituisce un *white noise* perché contiene ancora della correlazione non spiegata, e poi $\Phi(B)(1-B)^D Y_t = \Theta(B)\varepsilon_t$ si ottiene il modello (S)ARIMA(p,d,q)x(P,D,Q)_s così definito:

$$\phi(B)\Phi(B)(1-B)^d(1-B)^D Y_t = \phi_0 + \theta(B)\Theta(B)\varepsilon_t$$

dove:

- S indica la lunghezza del periodo stagionale (12 nel caso di dati mensili);
- $\Phi(B)$ è l'operatore autoregressivo non stagionale di ordine p ;
- $\Phi(B^S)$ è l'operatore autoregressivo stagionale di ordine P ;
- $\theta(B)$ è l'operatore a media mobile non stagionale di ordine q ;
- $\Theta(B^S)$ è l'operatore a media mobile stagionale di ordine Q ;
- $(1-B)^d$ è l'operatore differenza non stagionale di ordine d ;
- $(1-B)^D$ è l'operatore differenza stagionale di ordine D .

Un processo SARMA(p,q)x(P,Q)_s è stazionario se le radici dell'equazione caratteristica $\phi(B)\Phi(B)=0$ sono tutte in modulo maggiore di uno, è invertibile se le radici dell'equazione $\theta(B)\Theta(B)=0$ sono tutte esterne al cerchio di raggio unitario.

2.2.9 – La procedura Box e Jenkins

Box e Jenkins (1976) hanno proposto una procedura per costruire, a partire dai dati, un modello ARIMA che approssimi adeguatamente il processo generatore di una serie storica.

Essa consiste nelle seguenti fasi fondamentali:

1. identificazione
2. stima dei parametri
3. controllo diagnostico

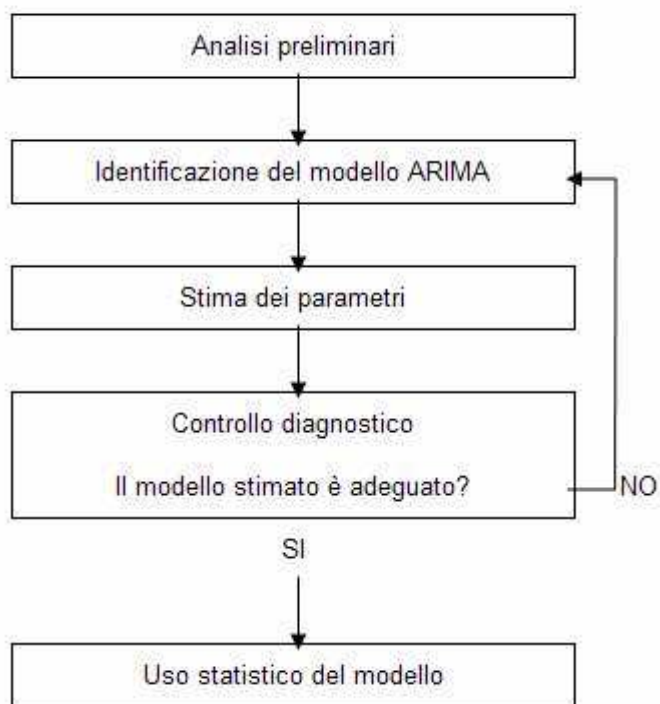


Figura 2.3 – Schema logico della procedura di Box e Jenkins

IDENTIFICAZIONE

Il primo passo consiste nella specificazione dell'ordine del modello, ovvero nell'individuare i parametri p , d , q , P , D , Q . I principali strumenti da utilizzare sono la funzione di

autocorrelazione e la funzione di autocorrelazione parziale, entrambe stimate sul campione.

L'idea di fondo è quella di riconoscere nella struttura della funzione di autocorrelazione empirica la struttura di una funzione teorica. Dunque, a tale scopo, le relazioni viste nei paragrafi precedenti possono essere riassunte come segue:

- la costruzione di un modello ARMA poggia sull'assunto che il processo sia stazionario e dunque le funzioni di autocorrelazione e di autocorrelazione parziale tendono ad annullarsi a velocità esponenziale al crescere di k . Se tendono ad annullarsi troppo lentamente, allora il processo generatore della serie è probabilmente non stazionario. È quindi opportuno operare una delle trasformazioni già descritte in modo tale da arrivare ad avere una serie stazionaria in media e in varianza; dopodiché si può procedere all'identificazione del modello.
- Per un processo MA(q) l'ACF è diversa da zero per $k \leq q$ e si annulla per $k \geq q+1$. La PACF invece tende a zero con una modalità che dipende dal valore del parametro θ_i .
- Per un processo AR(p) l'ACF tende a zero con un andamento dipendente dal valore dei parametri ϕ_i , mentre la PACF è diversa da zero per $k \leq p$ e si annulla per $k \geq p+1$.
- Quando non c'è un chiaro segno di trovarsi in una delle condizioni precedenti, è probabilmente più appropriato un modello misto ARMA.
- Quando invece tali componenti a ritardi stagionali, allora è opportuna ipotizzare la presenza di una qualche forma di stagionalità e quindi utilizzare un modello ARIMA stagionale (SARIMA).

STIMA DEI PARAMETRI

Una volta fissati i valori p e q si può passare alla fase di stima dei parametri che lo caratterizzano, ossia di $\delta = (\phi_0, \phi_1, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2)$. Ci sono diversi metodi per la stima, tra i più utilizzati il *metodo dei minimi quadrati non lineari* ed il *metodo della massima verosimiglianza*; quest'ultimo è quello più utilizzato dato che, sebbene richieda la conoscenza della distribuzione del termine d'errore, fornisce stimatori con migliori

proprietà statistiche. L'ipotesi di *gaussianità* dell'errore permette di effettuare un test sulla significatività o meno di ogni singolo coefficiente per il modello stimato.

CONTROLLO DIAGNOSTICO

L'idea di base per il controllo dell'adeguatezza del modello stimato consiste nella considerazione che, se il modello è correttamente specificato, allora sui residui $e_t = (y_t - \hat{y}_t)$ devono potersi riscontrare le ipotesi fatte a priori sui disturbi ε_t .

La principale verifica che si può fare consiste nell'analisi delle *autocorrelazioni dei residui*, infatti, se il modello ARIMA prescelto è un'adeguata rappresentazione del processo generatore dei dati, allora la funzione di autocorrelazione $\rho(k)$ della serie dei residui e_t , stimata tramite $\hat{\rho}(k)$, non dovrebbe essere significativamente diversa dalla autocorrelazione di un processo *white noise*, che com'è noto è nulla. Poiché per un *white noise* la varianza di $\hat{\rho}(k)$ è approssimativamente $1/n$ per ogni k , sotto ipotesi di normalità

la regione di accettazione al livello di significatività del 5% è data da $\left[\frac{-1.96}{\sqrt{n}}; \frac{1.96}{\sqrt{n}} \right]$. Un

valore di $\hat{\rho}(k)$ fuori da quest'intervallo porta a ritenere $\rho(k)$ significativamente diversa da zero. Questo risultato è utile perché permette di studiare l'ipotesi di in correlazione seriale tracciando le due parallele all'asse delle ascisse in $-1.96/\sqrt{n}$ e $1.96/\sqrt{n}$ e osservando il comportamento della funzione di autocorrelazione empirica al variare di k .

2.2.10 – Misura dell'accuratezza delle previsioni

La previsione è uno degli obiettivi prioritari per l'analisi delle serie storiche perché, oltre ad essere fine a se stessa, la determinazione dei valori futuri costituisce uno strumento per effettuare la scelta tra più modelli alternativi grazie alla bontà delle previsioni.

Se y_t è l'osservazione e quindi il dato reale, e \hat{y}_t è il valore previsto all'istante t , allora l'errore di previsione è definito come differenza tra i due valori: $e_t = y_t - \hat{y}_t$. Ovviamente

tale errore si può calcolare solo una volta che il dato reale sia stato osservato quindi, se si hanno a disposizione n osservazioni ed n previsioni per n istanti temporali, allora si potranno calcolare n errori e con essi una serie di indici per valutare la bontà delle previsioni:

$$EM = \frac{1}{n} \sum_{t=1}^n e_t$$

$$EMA = \frac{1}{n} \sum_{t=1}^n |e_t|$$

$$EQM = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}$$

L'EM (Errore Medio) dà chiaramente una semplice indicazione sulla presenza o meno di una sistematica sovrastima o sottostima, in esso infatti valori positivi e valori negativi si compensano tra loro e non danno una precisa valutazione della bontà della previsione.

L'EMA (Errore Medio Assoluto) rimedia al difetto dell'EM portando gli errori in valore assoluto che non permette la compensazione tra valori positivi e negativi.

L'indice su cui principalmente si baseranno le valutazioni del seguente elaborato è l'EQM (Errore Quadratico Medio) che indica la discrepanza quadratica media fra i valori osservati ed i valori stimati.

CAPITOLO III – APPLICAZIONE DEI MODELLI

3.1 – Il progetto di analisi

L'analisi che si è deciso di svolgere, in accordo con l'ufficio *Planning & Controlling Analysis*, è volta ad integrare alcune attività dell'ufficio stesso, in particolare il processo di *forecasting* che rappresenta una fase delicata ed importantissima dell'intero processo di previsione poiché, come già detto, fornisce supporto statistico e, quindi oggettivo, alle decisioni che dovranno poi essere comunicate alla Produzione.

La scelta di applicare le serie storiche ai dati di vendita di Lotto Sport Italia deriva quindi dalla grande importanza che ricoprono le previsioni in azienda, ma in gran parte dal desiderio di approfondire e applicare alla realtà concetti visti durante gli studi universitari. Sono stati utilizzati strumenti molto semplici, quali i modelli lineari per serie storiche, proprio per avvicinare quanto studiato in questi anni alla realtà aziendale in cui ci si potrebbe trovare a lavorare.

3.2 – I dati

Il processo di *forecasting* nell'azienda in esame ha, come già detto, l'obiettivo di fornire delle previsioni di vendita per tutti i circa 3000 articoli in cui consiste l'offerta Lotto. Infatti Lotto, avendo esternato tutta la sua produzione, ha necessità di organizzare per tempo i quantitativi di merce da richiedere ai fornitori esteri e questi dati verranno man mano aggiornati: da quando inizia la campagna vendite fino a poco prima della sua chiusura, viene comunicata all'ufficio Produzione un'indicazione della quantità da produrre a breve, cosicché possa aggiornare gli ordini precedentemente inviati ai produttori esterni in termini di tempistiche e volumi richiesti. Ovviamente ad inizio campagna non si potrà avere una precisa previsione del dato di chiusura poiché non si dispongono di sufficienti informazioni per quanto riguarda il passato recente; ad esempio, nel 2008, anno in cui si fa iniziare la crisi economica attuale, ma ancor più negli anni successivi, le vendite non avranno seguito fedelmente gli andamenti degli anni precedenti, ma ci potrebbero essere stati degli sconvolgimenti, questi ultimi intuibili solo da dati molto prossimi al momento della previsione.

Il dato di chiusura viene quindi aggiornato di mese in mese, per tener conto dell'andamento attuale delle vendite. È nell'ambito di questo processo che sono state inserite le previsioni mensili di questo elaborato, prima cercando di capire come organizzare i dati, poi modellando le varie serie e, dopo aver scelto quelle più adeguate, calcolando le previsioni vere e proprie.

Come ricordato più volte, Lotto Sport Italia produce in *outsourcing*; le fonti produttive si procurano le materie prime e svolgono l'intero processo di produzione mentre Lotto si occupa della progettazione. L'azienda rivende il prodotto finito ai negozi (clienti) che a loro volta poi lo rivenderanno ai consumatori finali. In Figura 3.1 le fasi del processo di vendita:

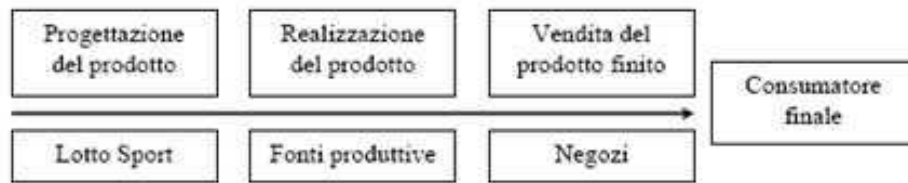


Figura 3.1 - Il processo di vendita

Una volta che il prodotto finito viene venduto da Lotto ai negozi nasce il dato di vendita denominato "sell-in" mentre il secondo processo di vendita, quello che termina con il consumatore, genera il dato di vendita detto "sell-out".

I dati di vendita di cui l'azienda dispone e che sono stati utilizzati in questo lavoro sono dati di vendita "sell-in".

Per chiarire brevemente le due tipologie di dato si può dire che il "sell-out" è un'operazione di vendita rivolta ai consumatori e quindi l'azienda deve attuare strategie commerciali atte ad attirare l'attenzione del consumatore sul prodotto e spingerlo all'acquisto mentre i dati "sell-in" focalizzano l'attenzione sui rivenditori: si tratta infatti di mettere in atto strategie che favoriscano l'ingresso del prodotto nell'espositore del punto vendita. Una volta che il produttore è riuscito ad ottenere questo, il problema si concentra principalmente sul distributore che deve far sì che la merce che espone venga comprata; la questione non è in realtà del tutto concentrata su quest'ultimo ma, soprattutto di recente, le aziende per essere competitive nel mercato devono proporsi come potenziali partner commerciali in grado di migliorare la performance dell'intero canale distributivo.

I dati utilizzati in questo lavoro sono quindi *sell-in* e strutturati in serie mensili, dal Gennaio 2005 ad Aprile 2011. Le serie sono ricavate dall'aggregazione di dati di vendita dell'abbigliamento dei singoli modelli per macrocategorie merceologiche come segue:

Desc Macrocategoria Merc	Desc Categoria Merc
HEAVY TOPS	GILET
	JACKET
	SWEAT LONG SLEEVE

LIGHT TOPS	JERSEY LONG SLEEVE
	JERSEY SHORT SLEEVE
	POLO LONG SLEEVE
	POLO SHORT SLEEVE
	TANK
	TOP
	T-SHIRT LONG SLEEVE
	T-SHIRT SHORT SLEEVE
	T-SHIRT SLEEVELESS

PANTS	PANT
	PANT MID

SHORTS/SKIRTS	BERMUDA
	SHORT

TRACKSUITS	TRACKSUIT FL
	TRACKSUIT MI
	TRACKSUIT PL

Si è scelta questa suddivisione perché rappresenta una sintesi e una linea guida per i singoli modelli appartenenti alla macrocategoria. Infatti le prime stime di vendita non vengono fatte sui singoli modelli, ma sulle macrocategorie, solo successivamente, per modelli particolarmente rilevanti o con caratteristiche singolari, verrà fatta una previsione per il singolo articolo. In questo modo si ha un notevole risparmio di tempo e una prima sintesi per l'ordinativo da comunicare alla Produzione; ovviamente all'interno di ogni raggruppamento ci sono delle differenze nei volumi di vendita, ma queste sono quasi sempre proporzionali ai dati della macrocategoria. Dove questo non avvenisse per

caratteristiche singolari del prodotto, si va ad analizzare, come già detto, la serie delle vendite del singolo articolo.

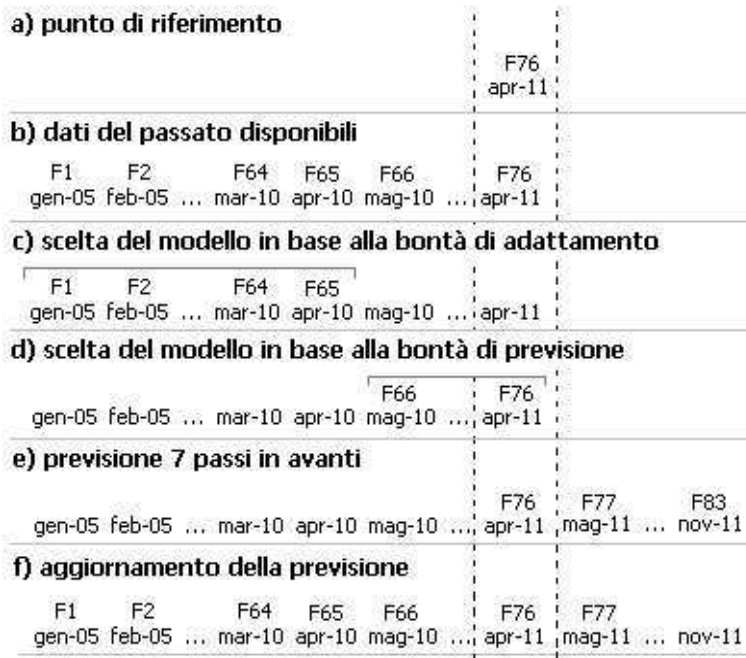
È importante sottolineare una caratteristica delle serie storiche analizzate: gli articoli venduti non sono ovviamente sempre gli stessi di anno in anno poiché le mode cambiano così come cambiano i gusti dei consumatori, quindi Lotto, come del resto tutte le aziende che operano nel settore dell'abbigliamento e della calzatura, effettua delle atte ad adattare la propria offerta alle esigenze di mercato, per far sì di essere sempre competitivi e al passo con le esigenze dei consumatori. Sorge dunque spontaneo chiedersi se può essere effettuata un'analisi considerando i dati di anni passati e quindi relativi a modelli per qualche aspetto diversi dagli attuali; il problema viene superato grazie ai cosiddetti modelli continuativi che permettono all'azienda di produrre lo stesso articolo applicando delle minime variazioni ma continuando a considerare il modello uguale a quello degli anni precedenti. Inoltre, osservando le serie di un modello continuativo, si può capire se queste piccole modifiche che sono state fatte al prodotto hanno portato conseguenze positive o se, invece, non sono da mantenere e si deve andare quindi cercando un altro cambiamento, semplicemente osservando il trend della serie in questione.

3.3 – L'applicazione ai dati dei modelli SARIMA

Nel secondo capitolo è stato illustrato lo schema e le procedure per la scelta del modello, la stima e le previsioni, ora applicheremo quelle tecniche ai dati a nostra disposizione e sarà illustrata nello specifico la procedura adottata. È importante ricordare che il valore d'interesse per l'azienda è la quantità di vendite previste a fine stagione dunque la previsione sarà sufficiente eseguirla per tanti passi in avanti quanti sono i mesi che mancano alla chiusura della stagione e utilizzare come dato finale la somma di queste previsioni.

Al momento in cui è stata fatta l'estrazione dei dati si è in possesso dei dati dei primi 4 mesi del 2011 e di un dato parziale per il mese di maggio. La stagione estiva delle vendite ha inizio con il mese di giugno e la previsione andrebbe eseguita per i successivi 6 passi in avanti. Non potendo però considerare il mese di maggio al completo, si effettuerà una previsione 7 passi in avanti, a partire dal dato di aprile 2011; il mese di maggio 2011 non verrà ovviamente compreso nella somma del dato finale per la stagione estiva 2011.

In figura 3.2 è schematizzato il procedimento di stima e scelta del modello per ognuna delle serie.



- g) Il punto di riferimento è l'ultimo dato disponibile al momento in cui sono stati estratti i dati per l'analisi, aprile 2011
- h) Si dispone una serie di 76 dati, a partire da gennaio 2005
- i) Si deve, come prima cosa, scegliere un modello della classe dei SARIMA basandosi sul suo adattamento ai dati. È stato scelto di troncare il campione togliendo le ultime 12 osservazioni, scegliendo il miglior modello sulla base delle osservazioni fino ad Aprile 2010, quindi con (76-12) dati e per ogni modello da testare si andranno a calcolare gli errori di stima (valore osservato – valore stimato). Allo stesso modo si andrà a prendere, sul campione di dati troncato, il miglior modello con il metodo del Lisciamento esponenziale. Verranno quindi scelti, per ogni macrocategoria, il miglior modello SARIMA e il migliore con il Lisciamento esponenziale e successivamente, all'interno di ogni coppia di modelli, quello da preferire in base all'Errore Quadratico Medio (EQM).
- j) Dopo aver scelto il miglior modello si passerà alla previsione, andando a prevedere, nel nostro caso, fino a 7 passi in avanti.
- k) A questo punto, visto che l'analisi si è svolta su dati reali, di mese in mese si potrà andare ad aggiornare la previsione, disponendo dunque di un dato in più e quindi di maggiori informazioni. Ad esempio a giugno 2011 si avrà a disposizione l'intero dato

per il mese di maggio e quindi 77 dati, da cui si potrà stimare il dato finale senza dover riformulare un modello ma semplicemente aggiornando quello scelto.

Senza ripetere lo studio dei modelli per tutte le cinque macrocategorie che risulterebbe ripetitivo, verrà riportata in dettaglio l'analisi per la sola categoria *Tracksuit* che risulta la più redditizia per l'azienda, per quanto riguarda il settore abbigliamento. Va sottolineato che le analisi e le relative previsioni sono state svolte con i medesimi metodi per tutte le rimanenti serie.

3.3.1 – Analisi della serie di vendita di Tracksuits

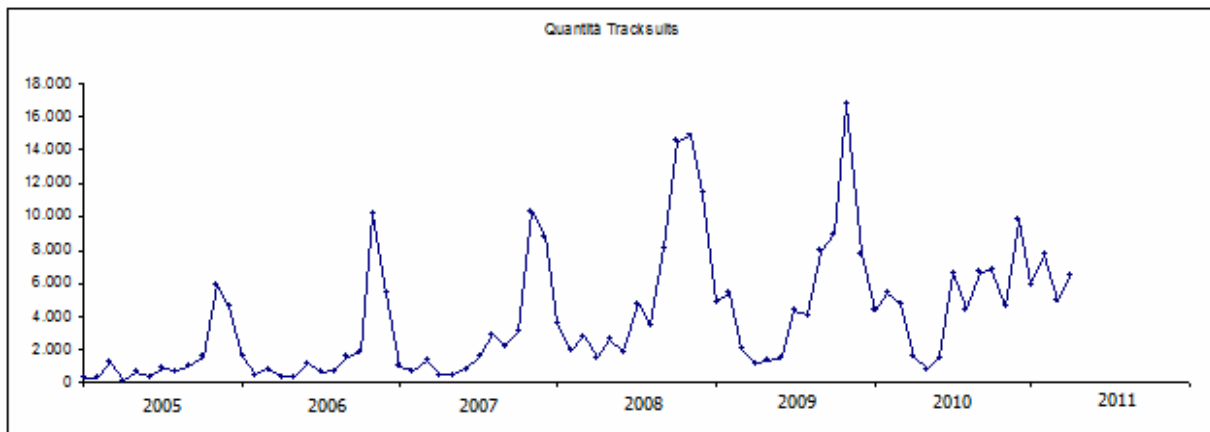


Figura 3.3 – Grafico serie storica vendite Tracksuits

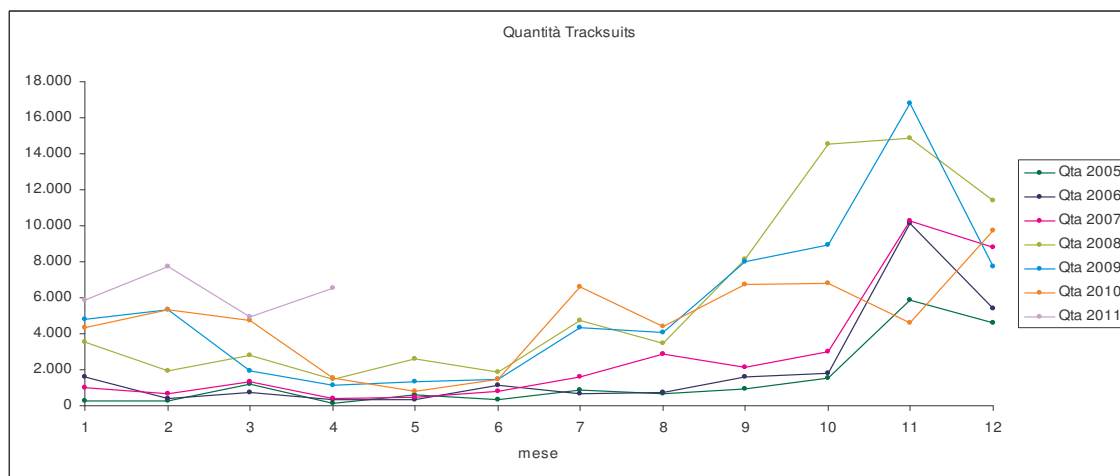


Figura 3.4 – Grafico serie storica vendite Tracksuits per anno

Come si può notare dal grafico in figura 3.3 la serie presenta una forte stagionalità con picchi di vendita negli ultimi mesi dell'anno. La media sembra non essere costante e denota dunque la presenza di un trend leggermente crescente, così come crescente risulta la varianza.

I modelli verranno calcolati sulla base di (76-12) dati per poter poi calcolarne l'EQM e scegliere il migliore.

IL MODELLO A LISCIAMENTO ESPONENZIALE

L'osservazione del grafico della serie permette dunque di ipotizzare le caratteristiche del modello: nel nostro caso abbiamo un trend leggermente crescente e dato l'aumentare della varianza con il passare del tempo il modello potrebbe essere moltiplicativo. La serie sembra essere non stazionaria nè in media nè in varianza.

Proviamo a stimare i modelli additivo e moltiplicativo sulle prime 76 osservazioni (per la stima utilizziamo il pacchetto statistico R). I risultati della stima e i vari indicatori di bontà del modello (l'errore medio (EM), l'errore medio assoluto (EMA) e l'errore quadratico medio (EQM)) sono riportati nella Tabella 3.1:

<u>Modello Additivo</u>		<u>Modello Moltiplicativo</u>	
alpha	0.2439099	alpha	0.996357601
beta	0.9	beta	0.9
gamma	0.2967339	gamma	0.49385
EM	-1163,08	EM	-1599,51
EMA	2483,14	EMA	3097,77
EQM	3918,48	EQM	5202,17

Tabella 3.1 – Stime dei parametri di liscio e degli errori di previsione

Confrontando dunque i valori ottenuti, sembra preferibile il modello additivo con formule di

$$\hat{L}_n = 0.244(\hat{L}_{n-1} + \hat{T}_{n-1}) + 0.756(y_n - \hat{S}_{n-s})$$

aggiornamento $\hat{T}_n = 0.9 \cdot \hat{T}_{n-1} + 0.1 \cdot (\hat{L}_n - \hat{L}_{n-1})$ dove i valori iniziali per \hat{L}_n , \hat{T}_n

$$\hat{S}_n = 0.297 \cdot \hat{S}_{n-s} + 0.703(y_n - \hat{L}_n)$$

e \hat{S}_n sono dati rispettivamente dai coefficienti $a=1351.9276$, $b=-408.8864$ e dai vari S .

IL MODELLO SARIMA

Poichè, come abbiamo già precedentemente osservato, dal grafico della serie storica in Figura 3.3 emerge una non stazionarietà in varianza, ci occupiamo quindi subito di risolvere quest'ultimo problema con una trasformazione logaritmica. La figura 3.5 riporta il logaritmo della serie trasformata.

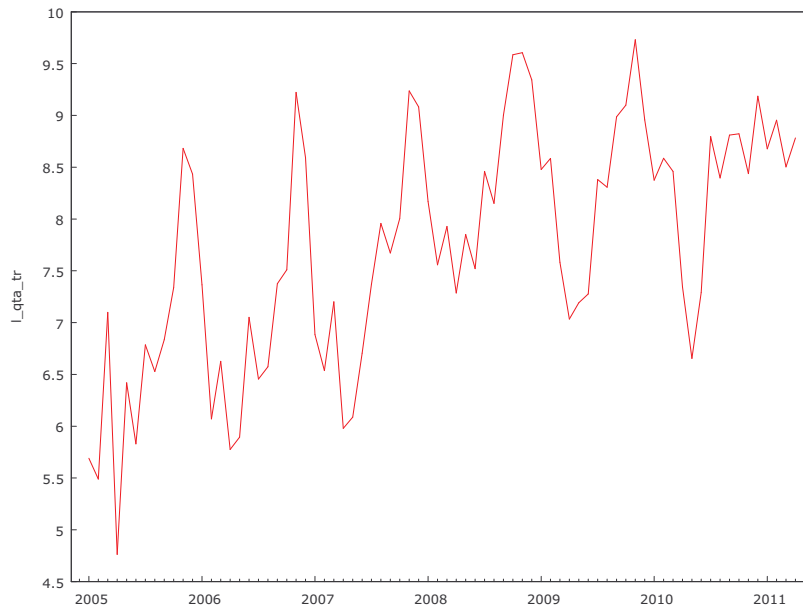


Figura 3.5 – Logaritmo della serie trasformata

Per procedere con l'identificazione del modello, bisogna analizzare le funzioni di autocorrelazione empiriche riportate in Figura 3.6:

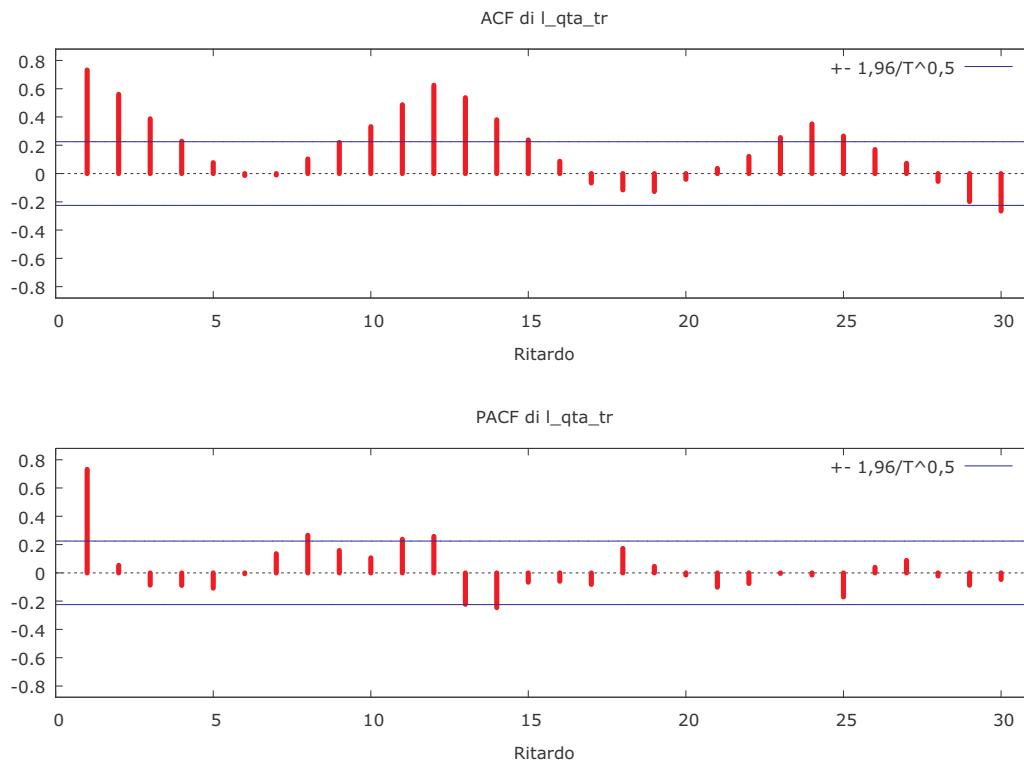


Figura 3.6 – Funzioni di autocorrelazione empiriche della serie trasformata

Da tali grafici emerge un'evidente componente stagionale e anche una componente autoregressiva di ordine 1.

Applicando la procedura Box-Jenkins si perviene, dopo diversi passi, all'identificazione di un modello SARIMA(1,0,1)(1,0,1)₁₂ con costante. Le stime, ottenute con il software Gretl, sono riportate di seguito in Tabella 3.2 mentre le funzioni di autocorrelazione empiriche dei residui si possono vedere in Figura 3.7.

gretl: modello 3

File Modifica Test Salva Grafici Analisi LaTeX

Valutazioni della funzione: 57
Valutazioni del gradiente: 20

Modello 3: ARMA, usando le osservazioni 2005:01-2010:04 (T = 64)
Stimato usando il filtro di Kalman (MV esatta)
Variabile dipendente: l_gta_tr
Errori standard basati sull'Hessiana

	coefficiente	errore std.	rapporto t	p-value	
const	7,37285	1,64279	4,488	7,19e-06	***
phi_1	0,950610	0,0484581	19,62	1,10e-085	***
Phi_1	0,926856	0,0545664	16,99	1,05e-064	***
theta_1	-0,613577	0,164042	-3,740	0,0002	***
Theta_1	-0,429275	0,167586	-2,562	0,0104	**

Media var. dipendente 7,592900 SQM var. dipendente 1,160936
Media innovazioni 0,123823 SQM innovazioni 0,525787
Log-verosimiglianza -57,52536 Criterio di Akaike 127,0507
Criterio di Schwarz 140,0040 Hannan-Quinn 132,1537
Note: SQM = scarto quadratico medio; E.S. = errore standard

	Reale	Immaginario	Modulo	Frequenza
AR				
Radice 1	1,0520	0,0000	1,0520	0,0000
AR (stagionale)				
Radice 1	1,0789	0,0000	1,0789	0,0000
MA				
Radice 1	1,6298	0,0000	1,6298	0,0000
MA (stagionale)				
Radice 1	2,3295	0,0000	2,3295	0,0000

Tabella 3.2 – Stime del modello SARIMA(1,0,1)(1,0,1)₁₂

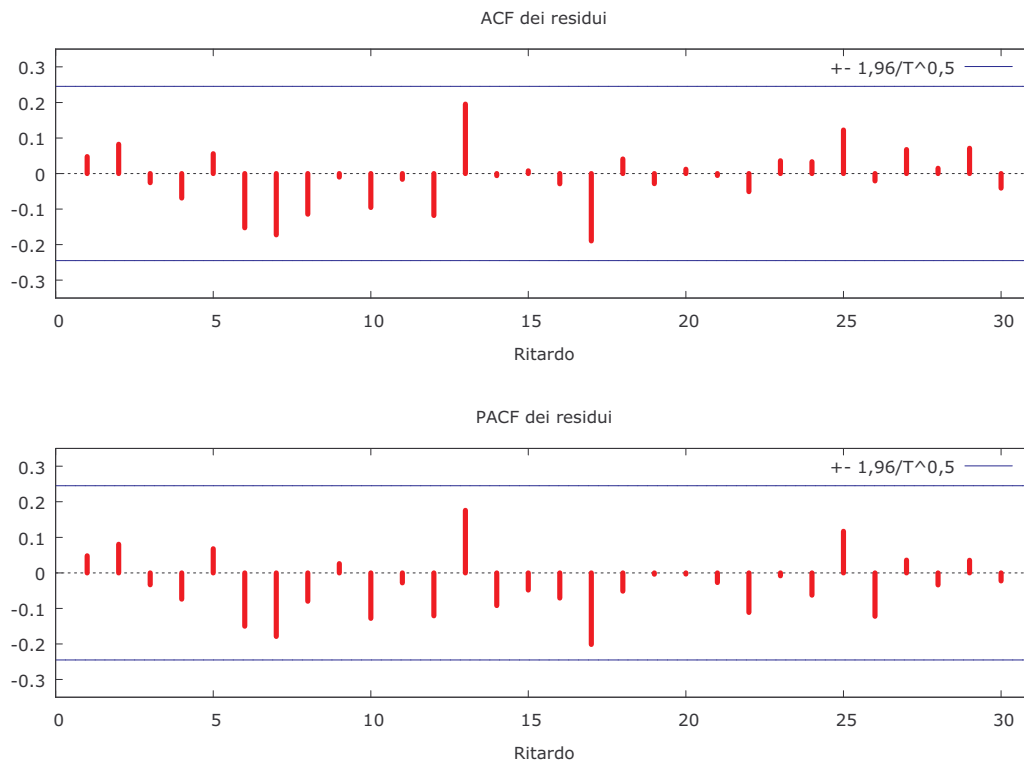


Figura 3.7 – Funzioni di autocorrelazione empiriche dei residui del modello $SARIMA(1,0,1)(1,0,1)_{12}$

Osserviamo che i coefficienti delle componenti autoregressive sono molto vicini a 1, il che potrebbero suggerire una differenziazione.

Applicando sempre la procedura Box-jenkins si ottiene che altri due modelli competitive con quello appena trovato sono i modelli $SARIMA(0,1,1)(1,0,1)_{12}$ e $SARIMA(5,0,0)(0,1,1)_{12}$ dove solo il quinto ritardo AR viene preso in considerazione in quanto gli altri risultano non significativi. I risultati relativi alle stime di tali modelli sono riportati di seguito nelle Tabelle 3.3 e 3.4.

gretl: modello 5

File Modifica Test Salva Grafici Analisi LaTeX

Valutazioni della funzione: 61
 Valutazioni del gradiente: 20

Modello 5: ARIMA, usando le osservazioni 2005:02-2010:04 (T = 63)
 Stimato usando il filtro di Kalman (MV esatta)
 Variabile dipendente: (1-L) l_qta_tr
 Errori standard basati sull'Hessiana

	coefficiente	errore std.	rapporto t	p-value
Phi_1	0,954892	0,0388103	24,60	1,14e-133 ***
theta_1	-0,713730	0,122798	-5,812	6,16e-09 ***
Theta_1	-0,521289	0,175083	-2,977	0,0029 ***

Media var. dipendente 0,026270 SQM var. dipendente 0,828209
 Media innovazioni 0,049554 SQM innovazioni 0,513909
 Log-verosimiglianza -56,00466 Criterio di Akaike 120,0093
 Criterio di Schwarz 128,5819 Hannan-Quinn 123,3809

Note: SQM = scarto quadratico medio; E.S. = errore standard

	Reale	Immaginario	Modulo	Frequenza
AR (stagionale)				
Radice 1	1,0472	0,0000	1,0472	0,0000
MA				
Radice 1	1,4011	0,0000	1,4011	0,0000
MA (stagionale)				
Radice 1	1,9183	0,0000	1,9183	0,0000

Tabella 3.3 – Stime del modello SARIMA(0,1,1)(1,0,1)₁₂

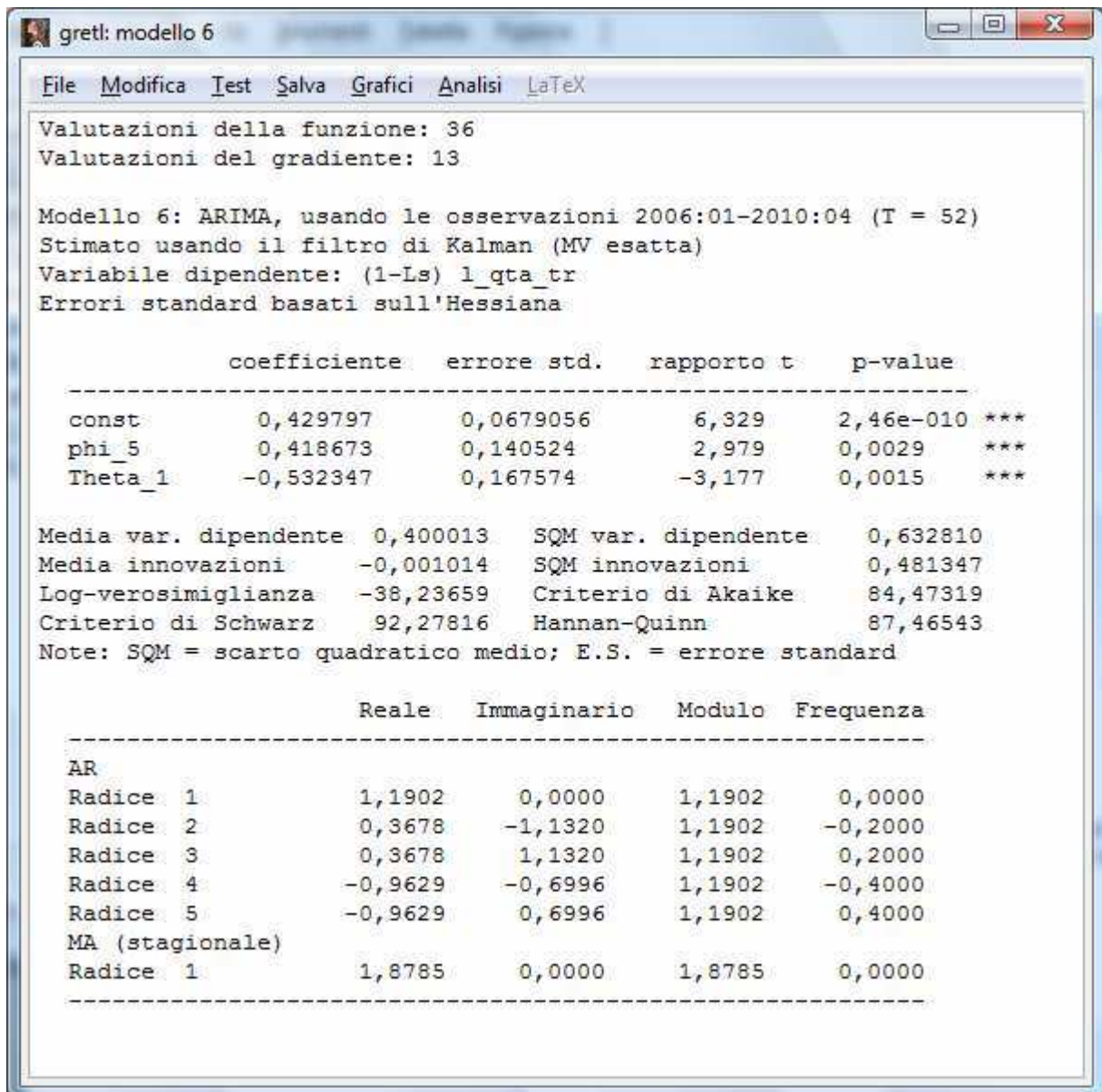


Tabella 3.4 – Stime del modello SARIMA(5,0,0)(0,1,1)₁₂

A questo punto non ci resta che confrontare tra loro questi modelli confrontando la bontà delle previsioni. Calcoliamo dunque, per i 3 modelli stimati, gli Errori Quadratici Medi (EQM), riportati nelle Tabelle 3.5, 3.6, 3.7.

gretl: previsioni

Per intervalli di confidenza al 95%, $z(0,025) = 1,96$

Oss	l_qta_tr	Previsione	Errore std	Intervallo al 95%
2009:05	7,191429	7,639384		
2009:06	7,275865	7,512038		
2009:07	8,380686	8,018984		
2009:08	8,305978	8,098387		
2009:09	8,986447	8,607085		
2009:10	9,099074	9,117014		
2009:11	9,730324	9,599904		
2009:12	8,956995	9,378655		
2010:01	8,374938	8,340399		
2010:02	8,586346	8,166950		
2010:03	8,459564	7,993526		
2010:04	7,345365	7,514406		
2010:05	6,654153	7,865671	0,525787	6,835148 - 8,896194
2010:06	7,290293	8,130656	0,554846	7,043178 - 9,218135
2010:07	8,795734	9,019736	0,579855	7,883241 - 10,156230
2010:08	8,395703	8,730905	0,601560	7,551869 - 9,909942
2010:09	8,810012	9,119714	0,620522	7,903514 - 10,335915
2010:10	8,822617	9,119472	0,637172	7,870639 - 10,368306
2010:11	8,438799	9,646046	0,651852	8,368440 - 10,923652
2010:12	9,185740	9,071584	0,664839	7,768524 - 10,374644
2011:01	8,677951	8,616954	0,676360	7,291312 - 9,942595
2011:02	8,954415	8,608191	0,686605	7,262470 - 9,953913
2011:03	8,502080	8,174695	0,695734	6,811083 - 9,538308
2011:04	8,782630	7,101951	0,703881	5,722370 - 8,481531

Statistiche della previsione

Errore medio	-0,15795
Errore quadratico medio	0,58721
Radice dell'errore quadratico medio	0,76629
Errore assoluto medio	0,57953
Errore percentuale medio	-2,3883
Errore percentuale assoluto medio	7,1882
U di Theil	1,0533
Proporzione del bias, UM	0,042488
Proporzione della regressione, UR	0,24893
Proporzione del disturbo, UD	0,70858

Tabella 3.5 – Previsioni e EQM del modello $SARIMA(1,0,1)(1,0,1)_{12}$



Figura 3.8 – Grafico previsioni modello SARIMA(1,0,1)(1,0,1)₁₂

gretl: previsioni

Per intervalli di confidenza al 95%, $z(0,025) = 1,96$

Oss	l_qta_tr	Previsione	Errore std	Intervallo al 95%
2009:05	7,191429	7,956290		
2009:06	7,275865	7,430325		
2009:07	8,380686	8,091373		
2009:08	8,305978	7,958858		
2009:09	8,986447	8,732587		
2009:10	9,099074	9,268812		
2009:11	9,730324	9,224506		
2009:12	8,956995	9,204612		
2010:01	8,374938	8,455496		
2010:02	8,586346	8,508768		
2010:03	8,459564	7,771595		
2010:04	7,345365	7,599437		
2010:05	6,654153	7,628325	0,574066	6,503177 - 8,753474
2010:06	7,290293	7,692754	0,609733	6,497699 - 8,887809
2010:07	8,795734	8,535792	0,643426	7,274700 - 9,796884
2010:08	8,395703	8,478786	0,675441	7,154946 - 9,802625
2010:09	8,810012	8,998020	0,706005	7,614275 - 10,381764
2010:10	8,822617	9,083960	0,735300	7,642799 - 10,525122
2010:11	8,438799	9,565638	0,763472	8,069262 - 11,062015
2010:12	9,185740	8,975547	0,790640	7,425921 - 10,525173
2011:01	8,677951	8,531406	0,816905	6,930301 - 10,132511
2011:02	8,954415	8,692722	0,842352	7,041742 - 10,343702
2011:03	8,502080	8,595980	0,867053	6,896588 - 10,295372
2011:04	8,782630	7,745786	0,891068	5,999324 - 9,492249

Statistiche della previsione

Errore medio	-0,10122
Errore quadratico medio	0,31474
Radice dell'errore quadratico medio	0,56102
Errore assoluto medio	0,42042
Errore percentuale medio	-1,587
Errore percentuale assoluto medio	5,197
U di Theil	0,73177
Proporzione del bias, UM	0,03255
Proporzione della regressione, UR	0,047757
Proporzione del disturbo, UD	0,91969

Tabella 3.6 – Previsioni e EQM del modello SARIMA(0,1,1)(1,0,1)₁₂

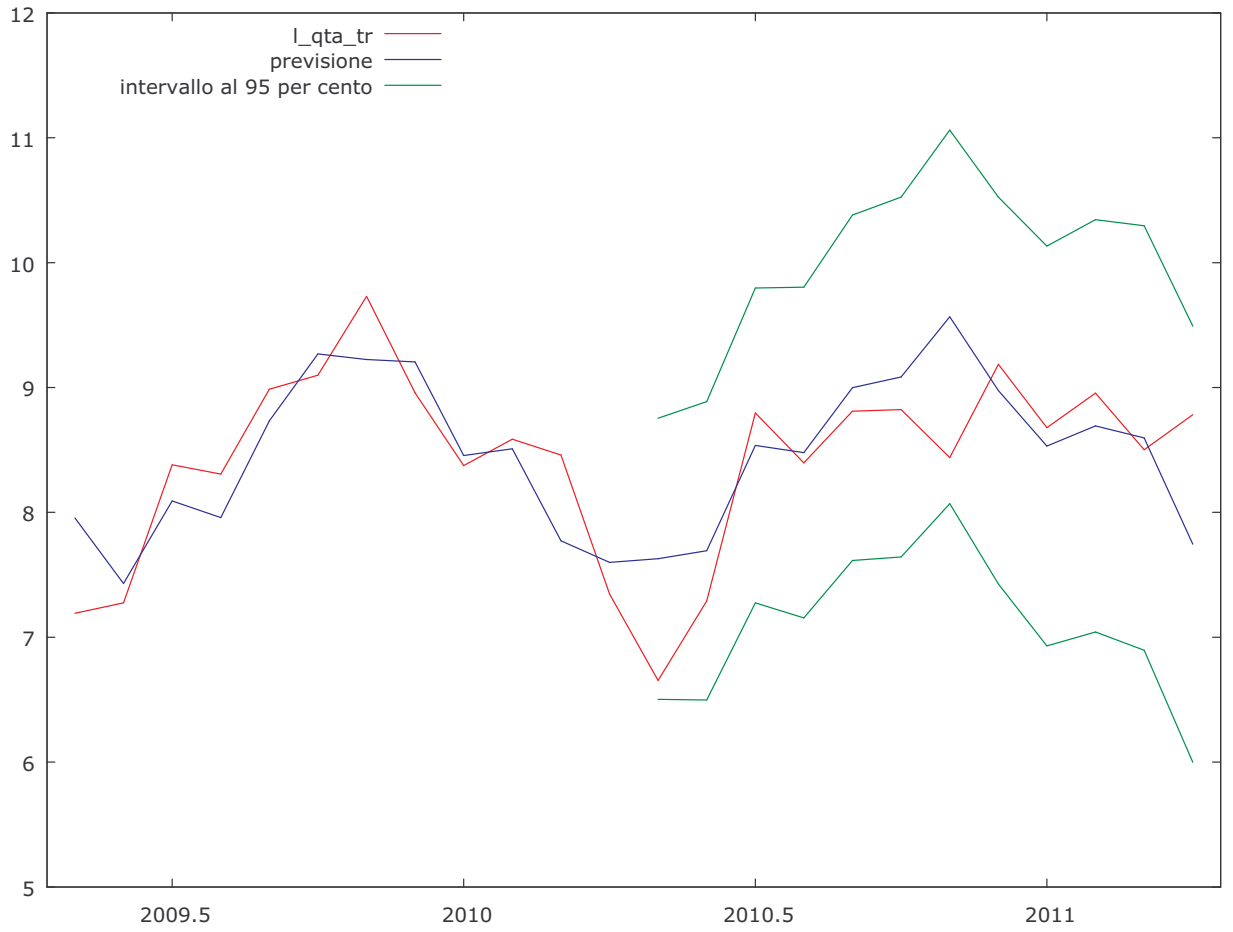


Figura 3.9 – Grafico previsioni modello $SARIMA(0,1,1)(1,0,1)_{12}$

greti: previsioni

Per intervalli di confidenza al 95%, $z(0,025) = 1,96$

Oss	l_qta_tr	Previsione	Errore std	Intervallo al 95%
2009:05	7,191429	7,658562		
2009:06	7,275865	7,954425		
2009:07	8,380686	8,897368		
2009:08	8,305978	8,191340		
2009:09	8,986447	8,884754		
2009:10	9,099074	9,205507		
2009:11	9,730324	9,904584		
2009:12	8,956995	9,843685		
2010:01	8,374938	8,718007		
2010:02	8,586346	8,527597		
2010:03	8,459564	8,224500		
2010:04	7,345365	7,403073		
2010:05	6,654153	7,528261	0,481347	6,584839 - 8,471684
2010:06	7,290293	7,843957	0,481347	6,900534 - 8,787379
2010:07	8,795734	8,906062	0,481347	7,962639 - 9,849484
2010:08	8,395703	8,861907	0,481347	7,918484 - 9,805329
2010:09	8,810012	9,312361	0,481347	8,368938 - 10,255783
2010:10	8,822617	9,546608	0,521831	8,523837 - 10,569378
2010:11	8,438799	10,310788	0,521831	9,288018 - 11,333559
2010:12	9,185740	9,898835	0,521831	8,876065 - 10,921606
2011:01	8,677951	9,040174	0,521831	8,017404 - 10,062945
2011:02	8,954415	8,941375	0,521831	7,918604 - 9,964145
2011:03	8,502080	8,771651	0,528608	7,735598 - 9,807705
2011:04	8,782630	7,868963	0,528608	6,832909 - 8,905016

Statistiche della previsione

Errore medio	-0,46007
Errore quadratico medio	0,59403
Radice dell'errore quadratico medio	0,77073
Errore assoluto medio	0,61452
Errore percentuale medio	-5,6824
Errore percentuale assoluto medio	7,4405
U di Theil	1,083
Proporzione del bias, UM	0,35632
Proporzione della regressione, UR	0,1838
Proporzione del disturbo, UD	0,45988

Tabella 3.7 – Previsioni e EQM del modello SARIMA(5,0,0)(0,1,1)₁₂

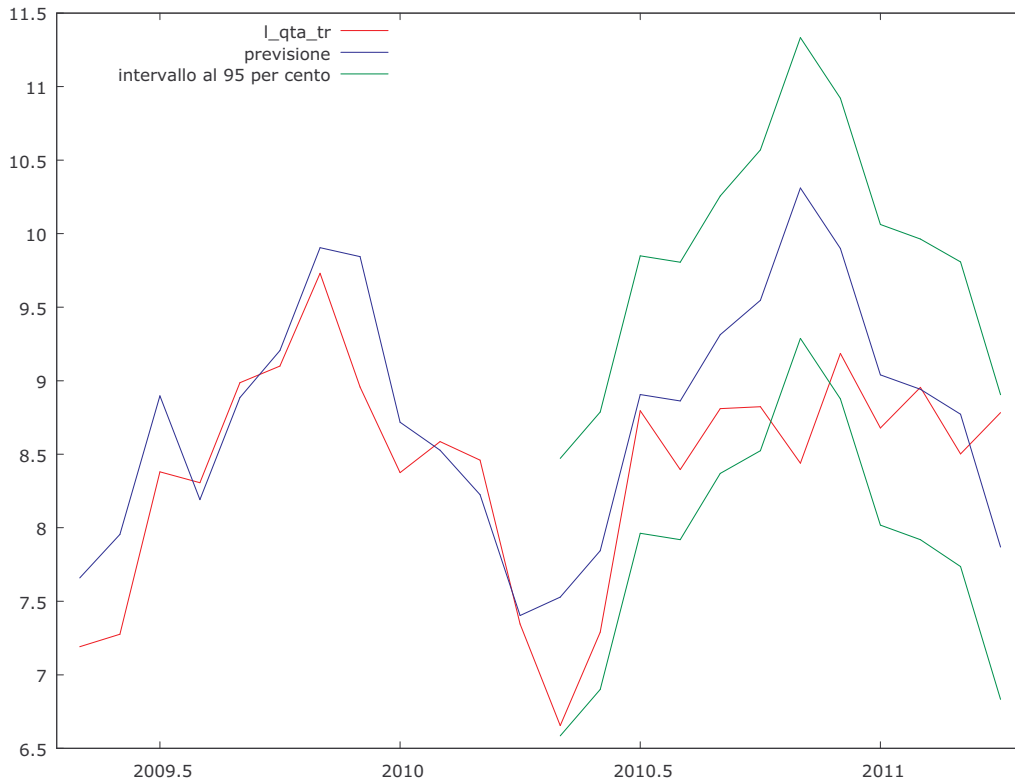


Figura 3.10 – Grafico previsioni modello SARIMA(5,0,0)(0,1,1)₁₂

Modello	EQM
SARIMA(1,0,1)(1,0,1) ₁₂	0.58721
SARIMA(0,1,1)(1,0,1) ₁₂	0.31474
SARIMA(5,0,0)(0,1,1) ₁₂	0.59403

Dall'analisi dell'EQM sembra preferibile il modello SARIMA(0,1,1)(1,0,0)₁₂ senza costante:

$$(1 - B) \cdot (1 - 0.763) \cdot \log(Y_t) = (1 + 0.642) \cdot \log(\varepsilon_t)$$

Verifichiamo, infine, normalità e casualità dei residui che come si può osservare dalla Figura 3.11 risultano buone:

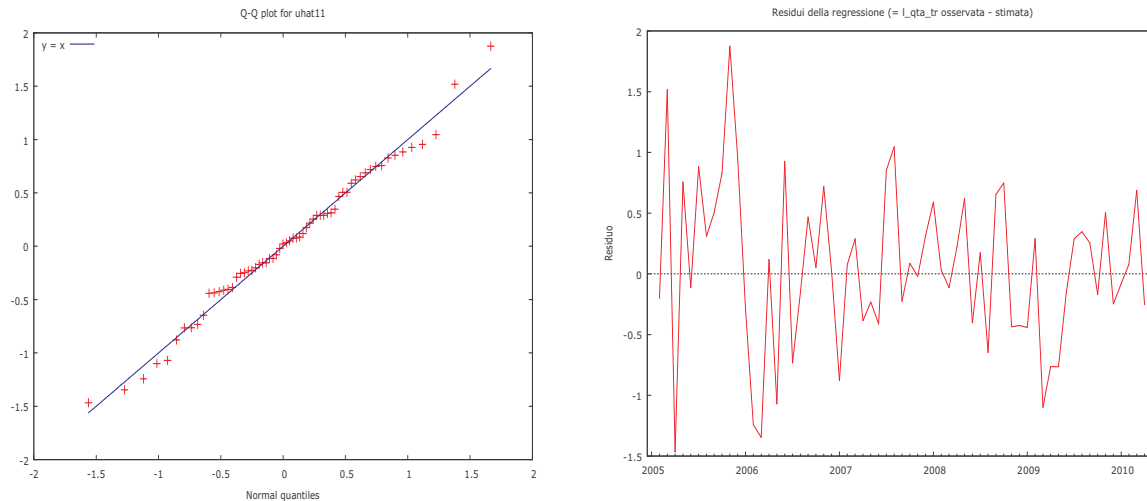


Figura 3.11 – Grafico normalità e casualità residui

Per poter confrontare le previsioni ottenute da questo modello con quelle ottenute dal modello di Holt-Winters, dobbiamo ritrasformare le previsioni tramite la funzione esponenziale. L'EQM risulta dunque 3276,175.

Confrontando dunque i due modelli, quello con metodo Holt Winters e quello con metodo SARIMA, il modello che presenta EQM minore, ed è quindi da preferire, è il modello SARIMA.

Effettuiamo ora la stima del modello SARIMA scelto con il campione al completo (76 osservazioni) e calcoliamo i valori delle vendite previsti fino a conclusione della stagione Spring-Summer 2011 (SS11), quindi 7 passi in avanti.



Figura 3.12 – Grafico previsioni 7 passi in avanti

I valori della previsione non sono pubblicati per motivi di privacy dell'azienda.

3.3.2 – Le serie rimanenti

Di seguito sono riportati i modelli scelti per le serie delle altre macrocategorie, con i relativi grafici di previsioni 7 passi in avanti.

Heavy Tops

Modello	Criterio di Akaike	EQM
SARIMA(0,1,2)(1,0,0) ₁₂ senza costante	203	0.49295
SARIMA(0,0,1)(0,0,1) ₁₂ con costante	206	1.2392
SARIMA(0,1,1)(0,0,1) ₁₂ senza costante	211	3.1003
SARIMA(1,0,0)(0,0,1) ₁₂ con costante	204	1.6716

Modello	EQM (non logaritmi)
SARIMA(0,1,2)(1,0,0) ₁₂ senza costante	466.05
Holt Winters stagionalità moltiplicativa	4976.81

Quindi la scelta ricade sul modello SARIMA(0,1,2)(1,0,0)₁₂ senza costante:

gretl: modello 5

File Modifica Test Salva Grafici Analisi LaTeX

Valutazioni della funzione: 56
 Valutazioni del gradiente: 20

Modello 5: ARIMA, usando le osservazioni 2005:02-2010:04 (T = 63)
 Stimato usando il filtro di Kalman (MV esatta)
 Variabile dipendente: (1-L) l_v1
 Errori standard basati sull'Hessiana

	coefficiente	errore std.	rapporto t	p-value	
Phi_1	0,387704	0,129053	3,004	0,0027	***
theta_1	-0,525099	0,121126	-4,335	1,46e-05	***
theta_2	-0,367446	0,114042	-3,222	0,0013	***

Media var. dipendente -0,002915 SQM var. dipendente 1,334296
 Media innovazioni 0,122964 SQM innovazioni 1,112049
 Log-verosimiglianza -97,62980 Criterio di Akaike 203,2596
 Criterio di Schwarz 211,8321 Hannan-Quinn 206,6312

Note: SQM = scarto quadratico medio; E.S. = errore standard

	Reale	Immaginario	Modulo	Frequenza
AR (stagionale)				
Radice 1	2,5793	0,0000	2,5793	0,0000
MA				
Radice 1	1,0833	0,0000	1,0833	0,0000
Radice 2	-2,5123	0,0000	2,5123	0,5000

Tabella 3.8 – Stime del modello SARIMA(0,1,2)(1,0,0)₁₂

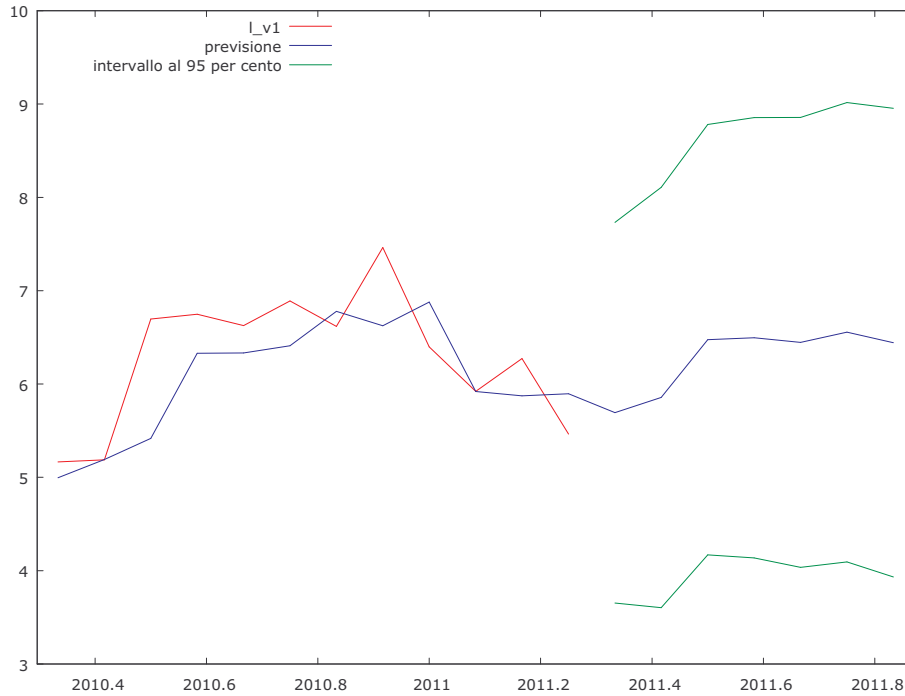


Figura 3.13 – Grafico previsioni 7 passi in avanti (Heavy Top)

Light Tops

Modello	Criterio di Akaike	EQM
SARIMA(1,0,1)(1,0,1) ₁₂ con costante	129.8638	0.59765
SARIMA(1,0,1)(1,1,0) ₁₂ con costante	99.63076	0.99129
SARIMA(0,1,1)(1,0,2) ₁₂ senza costante	120.9669	2.1369

Quindi la scelta ricade per ora sul modello SARIMA(1,0,1)(1,0,1)₁₂ con costante, nonostante il criterio di Akaike suggerisse un altro modello come migliore ma il nostro obiettivo è la previsione, dunque si basa la scelta del modello sull'EQM di previsione.

Modello	EQM (non logaritmi)
SARIMA(1,0,1)(1,0,1) ₁₂ con costante	4064.596
Holt Winters stagionalità additiva	3516

Quindi va scelto il modello di Holt Winters additivo:

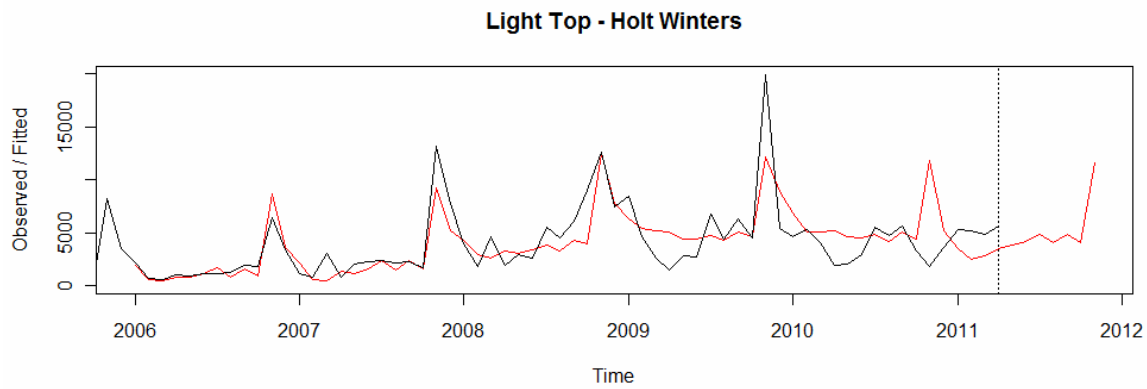


Figura 3.14 – Grafico previsioni 7 passi in avanti (Light Top)

Pants

<u>Modello</u>	<u>Criterio di Akaike</u>	<u>EQM</u>
SARIMA(1,0,0)(0,0,1) ₁₂ con costante	191.8060	1.44
SARIMA(1,1,7)(0,0,1) ₁₂ senza costante ($\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6=0$)	191.4478	1.0757
SARIMA(0,0,0)(0,1,0) ₁₂ con costante	153.4707	0.68001

<u>Modello</u>	<u>EQM (non logaritmi)</u>
SARIMA(0,0,0)(0,1,0) ₁₂ con costante	484.0152
Holt Winters stagionalità additiva	5512.32

Quindi la scelta ricade sul modello SARIMA(0,0,0)(0,1,1)₁₂ con costante:



Figura 3.15 – Grafico previsioni 7 passi in avanti (Pants)

Shorts/Skirts

Modello	Criterio di Akaike	EQM
SARIMA(1,0,0)(0,0,2) ₁₂ con costante	122.6977	0.30004
SARIMA(1,0,1)(0,0,2) ₁₂ con costante	120.0509	1.5266
SARIMA(1,1,0)(0,0,2) ₁₂ senza costante	120.13	0.31371
SARIMA(0,0,0)(0,1,1) ₁₂ con costante	82.07826	0.93508

In questo caso la scelta sta tra il modello SARIMA(1,0,0)(0,0,2)₁₂ con costante e il modello SARIMA(1,1,0)(0,0,2)₁₂ senza costante poiché presentano valori di EQM molto simili. Si sceglie infine quello tra i due con Criterio di Akaike minore:

Modello	EQM (non logaritmi)
SARIMA(1,1,0)(0,0,2) ₁₂ senza costante	1358.061
Holt Winters stagionalità moltiplicativa	4417.7

Anche questa volta si preferisce il modello SARIMA:



Figura 3.15 – Grafico previsioni 7 passi in avanti (Shorts/Skirts)

CONCLUSIONI

È stato introdotto il settore Montebellunese della calzatura sportiva in cui si colloca Lotto Sport Italia ed è stata descritta l'attuale struttura dell'azienda e come questa sia riuscita ad affermarsi nel mondo della calzatura sportiva e nell'abbigliamento tecnico, oltre che per il tempo libero. Ci si è poi soffermati ad illustrare le mansioni svolte dall'ufficio *Planning & Controlling Analysis*, sede dello stage che ha portato alla stesura di questo elaborato.

All'interno di tale ufficio, è stato possibile prendere parte ad alcuni processi aziendali come il *reporting*, il *pricing* e il *forecasting*.

Durante il processo di *reporting* è stato possibile imparare ad utilizzare una delle più diffuse applicazioni di *business intelligence* che permette di analizzare le grandi quantità di dati di cui l'azienda dispone, il software *Business Objects*. Con l'ausilio di questo software si sono creati numerosi report per manager, direttori e responsabili d'azienda. Il processo di *reporting* non può chiaramente funzionare solo tramite un software che permette l'estrazione di dati, ma ci deve essere, alla base, un analista che conosca le potenzialità di tale strumento e soprattutto conosca il contenuto delle variabili che utilizza, in modo da fornire dati coerenti e corretti. Infatti, è di fondamentale importanza, quando viene fatta una richiesta, esaminare quello che sarebbe stato l'oggetto dell'analisi in modo tale da capire quali potevano essere le informazioni più utili da inserire e quali invece da escludere perché superflue o addirittura fuorvianti per l'interpretazione del report.

Per quanto riguarda il processo di *pricing*, si è preso parte a tutto ciò che riguarda l'aggiornamento dei prezzi e delle marginalità e la gestione del complesso universo listini aziendale. L'azienda utilizza listini manuali, le cui modifiche vanno inserite da chi segue questo processo, e listini automatici che seguono un aggiornamento notturno automatico, basato su una programmazione ben definita che dipende dai costi a cui vengono poi applicati varie marginalità aggiunte.

Il processo di *forecasting* invece, a cui si è preso parte solo marginalmente fino a quando non vi sono state inserite le tecniche statistiche utilizzate in questo lavoro per fare previsione, ha permesso di comprendere a fondo l'importanza della previsione di vendita. Nell'ambito degli studi universitari si sono apprese tecniche e metodologie che permettono di arrivare ad un risultato con il minor errore possibile; lo stage svolto presso un'azienda così grande e complessa, mi ha permesso di capire quanto tali tecniche possano essere

d'aiuto, ma, soprattutto, mi ha permesso di toccare con mano quanto di teorico è stato studiato ed applicarlo alla realtà che un'impresa deve affrontare.

Sono quindi state introdotte le tecniche statistiche di modellazione delle serie storiche nell'ambito di questo processo, individuando assieme ai responsabili dell'ufficio Analisi Vendite come potessero essere strutturate le serie storiche degli oltre 3.000 articoli arrivando a produrre delle previsioni mensili dettagliate per macrocategoria produttiva.

Sono stati descritti i modelli *SARIMA* e i criteri di scelta per il modello migliore. Per svolgere le analisi è stato fatto uso dei software statistici *Gret* e *R*.

In conclusione, si può ribadire quanto già detto: in azienda servono strumenti statistici semplici per effettuare le previsioni di vendita, possibilmente implementati tramite strumenti software che rendano il più possibile automatica l'attività di stima.

BIBLIOGRAFIA

Di Fonzo T., Lisi F., *Serie storiche economiche. Analisi statistiche e applicazioni*, Carrocci Editore, 2005.

Durante A., *Rapporto OSEM 2009* – Montebelluna Sportsystem, Veneto Banca, Montebelluna, Treviso, 2006.

Favotto F., *Economia aziendale: Modelli, misure e casi*, Mc Graw Hill, Milano, 2001.

Grandinetti R., *"Marketing – Mercati, prodotti e relazioni"*, Carrocci Editore, 2008.

Rispoli M., *L'impresa industriale*. Il Mulino, 1984.

Rampin V., *"La previsione delle vendite di Lotto Sport Italia S.p.A"*, Facoltà di Scienze Statistiche, Università degli studi di Padova (a.a. 2007-2008).

Stock, Watson, *Introduzione all'econometria*, Pearson, 2005.

Box G. E. P., Jenkins S. G. M., *Times Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, 1976.

Makridakis, Hyndman, *Forecasting methods and applications*, Wiley & Sons, 1998

www.osservatoriodistretti.org

www.montebellunadistrict.com

www.lottosport.com

PROGRAMMI STATISTICI

Gretl 1.8.7 (2010-01-24) <http://gretl.sourceforge.net/>

R version 2.9.0 (2009-04-17) Copyright (C) 2009 The R Foundation for Statistical Computing ISBN 3-900051-07-0 <http://www.R-project.org>