



**UNIVERSITÀ DEGLI STUDI DI PADOVA**

**FACOLTÀ DI SCIENZE STATISTICHE**

**CORSO DI LAUREA IN STATISTICA E TECNOLOGIE INFORMATICHE**

**TESI DI LAUREA**

**APPLICAZIONE DI TECNICHE STATISTICHE  
DEL CONTROLLO DI PROCESSO:  
CASO STUDIO**

**RELATORE: Prof.ssa CAPIZZI GIOVANNA**

**LAUREANDO: PONCINA GIULIANO**

**ANNO ACCADEMICO 2005/06**



# Indice

Premesse e finalità.....	5
Capitolo 1	
Controllo statistico di processo .....	7
Capitolo 2	
Fibre acriliche.....	11
2.1 Introduzione .....	11
2.2 Storia e sviluppo delle fibre sintetiche.....	11
2.3 Classificazione delle fibre chimiche .....	13
2.4 Descrizione generale del processo .....	14
2.4.1 Polimerizzazione .....	14
2.4.2 Dissoluzione del polimero in solvente.....	14
2.4.3 Filatura.....	15
2.4.4 Tintura della fibra .....	16
2.4.5 Taglio ed imballo.....	16
2.5 Le principali caratteristiche fisico-meccaniche .....	17
2.5.1 Lunghezza della fibra .....	18
2.5.2 Massa lineare o titolo.....	18
2.5.3 Proprietà dinamometriche.....	19
2.5.4 Ensimaggio .....	20
2.6 Misure delle proprietà chimico-fisiche delle fibre.....	22
2.6.1 Analisi del laboratorio fisico (su bava) .....	22
2.6.2 Analisi del laboratorio chimico .....	24
2.7 LIMS .....	25
Capitolo 3	
Validazione di un metodo per analisi di laboratorio e quantificazione dell'incertezza nelle misure analitiche.....	27
3.1 Introduzione .....	27
3.2 Scopo e campo di applicazione.....	28
3.3 Incertezza .....	29
3.3.1 Definizione d'incertezza.....	29
3.3.2 Fonti d'incertezza .....	29
3.3.3 Componenti dell'incertezza.....	30
3.3.4 Errore e incertezza .....	30
3.4 Misure analitiche e incertezza.....	31
3.4.1 Validazione del metodo .....	31
3.5 Riferibilità.....	33
3.5.1 Taratura degli strumenti di misurazione.....	34
3.5.2 Misurazioni utilizzando metodi primari .....	34
3.5.3 Misurazioni utilizzando una sostanza pura come Materiale di Riferimento (MR) .....	34
3.5.4 Misurazioni su un Materiale di Riferimento Certificato (MRC).....	35
3.5.5 Misurazioni utilizzando un procedimento accettato.....	35

3.6 Il procedimento per la stima dell'incertezza di misura .....	35
3.7 Incertezza nei procedimenti analitici.....	39

## Capitolo 4

<b>Analisi statistica dati .....</b>	<b>41</b>
4.1 Premessa.....	41
4.2 Analisi preliminare dei dati .....	42
4.3 Analisi della distribuzione dei dati .....	43
4.4 Analisi della correlazione.....	50
4.5 Analisi della media e della varianza.....	51
4.5.1 Suddivisione tra le due macchine di produzione .....	51
4.5.2 Analisi dei dati raggruppati per colore.....	53
4.5.2.1 Primo raggruppamento, versione dei dati non aggiornata. ....	54
4.5.2.2 Secondo raggruppamento, versione dati aggiornati.....	58
<b>4.6 Considerazioni finali .....</b>	<b>60</b>
4.7 Studio delle componenti di variabilità del dato.....	60

## Capitolo 5

<b>Carte di controllo .....</b>	<b>63</b>
5.1 Nozioni di base.....	63
5.2 Caratteristiche di un disegno di controllo.....	65
5.3 Fasi operative per il calcolo del periodo di controllo.....	66
5.3.1 Fase 1 e fase 2 .....	66
5.3.2 Caso studio.....	66
5.4 Carte di controllo per misure singole .....	67
5.4.1 Carte per misure singole .....	67
5.4.2 Costruzione carte singole .....	68
5.4.3 Carte EWMA .....	70
5.4.4 Caso studio.....	71
5.5 Controllo multivariato .....	74
5.5.1 Introduzione .....	74
5.5.2 La distribuzione $T^2$ di Hotelling.....	74
5.5.3 Costruzione della carta $T^2$ di Hotelling.....	75
5.5.4 Caso studio.....	76
5.5.5 La carta MEWMA .....	78
5.5.6 ARL e determinazione dei parametri per la carta .....	78
5.5.7 Costruzione delle carte MEWMA nel caso studio.....	81
5.5.8 Interpretazione dei segnali fuori controllo .....	81
5.6 Caratteristiche e confronti .....	82

<b>Conclusioni.....</b>	<b>83</b>
-------------------------	-----------

<b>Appendice.....</b>	<b>85</b>
-----------------------	-----------

<b>Bibliografia.....</b>	<b>95</b>
--------------------------	-----------

<b>Ringraziamenti.....</b>	<b>97</b>
----------------------------	-----------

# Premesse e finalità

Questo elaborato nasce dall'esigenza da parte dell'autore di coniugare nozioni e concetti teorici riguardanti particolari campi della statistica applicata tra i quali, certificazione e miglioramento della qualità, ad una realtà produttiva industriale manifatturiera. Negli ultimi dieci anni il numero di realtà industriali che ha investito in qualità è esponenzialmente aumentato. Molte infatti sono le aziende che hanno scelto di creare un metodo organizzativo-gestionale capace di garantire nel rispetto della normativa di settore il proprio prodotto o servizio per la soddisfazione delle esigenze del cliente e tale metodo viene riconosciuto come **Sistema di Gestione della Qualità**. Il Sistema di Gestione della Qualità consente di mantenere tali garanzie impegnando risorse, metodi di lavoro, procedure ed utilizzando la tecnologia. Tale sistema porta spesso alla certificazione aziendale e con essa a tutti i vantaggi per poter concorrere in un mercato ormai sempre più globale e competitivo. Molti aspetti e diversi significati sono legati al concetto di qualità o meglio ad un Sistema Qualità, non solo per un fatto di "certificazione" ma anche per risolvere problemi di pianificazione del lavoro, sfruttare al meglio la tecnologia disponibile, controllare le non conformità del prodotto o del servizio, integrare ed interagire tra le varie fasi di sviluppo. Quindi tra le varie definizioni quella che rispecchia nel modo migliore il termine di qualità è: *"per qualità si intende il grado in cui un insieme di caratteristiche soddisfa requisiti espressi ed impliciti"*. Le tecniche per il miglioramento della qualità sono un'insieme di attività operative, gestionali e tecnologiche, che un'azienda attua per assicurarsi che le caratteristiche di un prodotto o servizio siano quelle di livello nominale o vincolate a specifiche richieste. Affinché un prodotto possa soddisfare le esigenze dei consumatori deve essere il risultato di un processo produttivo stabile e ripetibile. Per raggiungere questo scopo il processo deve essere in grado di produrre pezzi tali che la variabilità del valore nominale specifico del prodotto sia la più bassa possibile. Il controllo statistico di un processo produttivo è un'insieme di potenti strumenti, utili per raggiungere la stabilità del processo e per migliorare la produttività attraverso la riduzione della variabilità. Il primo capitolo di questa tesi tratta proprio il Controllo Statistico di Processo (SPC, *Statistical Process Control*), quali metodi e quali tecniche vengono maggiormente impiegati nelle realtà produttive. Nel capitolo due verrà esposta una presentazione generale della società Montefibre, del prodotto e del processo produttivo nel campo manifatturiero della fibra acrilica, specificando le caratteristiche di interesse considerate nell'analisi statistica di questo elaborato.

Vista la natura e la provenienza dei dati (laboratorio d'analisi chimico-fisiche) il capitolo tre tratterà l'argomento delle analisi di laboratorio definendo concetti e metodi (linee guida) per la valutazione dell'**incertezza**, termine di primaria importanza nelle misure d'analisi nei laboratori. Questo capitolo, tratto da una guida EURICHEM (associazione che raggruppa numerose organizzazioni in Europa tra cui istituti di metrologia, università e centri di ricerca), ha lo scopo di presentare i principi della validazione dei metodi cercando di mettere ordine tra definizioni e concetti. Ogniqualvolta si debbano prendere decisioni in base a risultati analitici, è importante avere qualche indicazione circa la loro qualità, cioè sul loro livello di affidabilità rispetto allo scopo specifico. In alcuni settori della chimica analitica viene richiesto formalmente (spesso per legge) che i laboratori introducano provvedimenti per la garanzia della qualità dei dati prodotti. Tra i provvedimenti di questo tipo vi sono l'uso di metodi analitici validati e l'uso di procedure specificate di controllo interno della qualità. L'analisi statistica effettuata nei dati viene trattata all'interno del capitolo quattro. Questa descrive e sottolinea i problemi insorti suggerendo alcune tra le possibili soluzioni. Tale capitolo riassume inoltre alcune delle tecniche statistiche di controllo utilizzate per il monitoraggio delle caratteristiche di interesse qualitativo. Partendo da considerazioni grafiche tramite istogrammi, grafici di dispersione, distribuzione della frequenza e boxplot (tecniche EDA), che consentono di avere un primo quadro riassuntivo, vengono successivamente utilizzati test statistici per specifiche ed opportune verifiche d'ipotesi. Infine sulla base dei risultati ottenuti, viene proposto e formulato un modello che tiene conto delle fonti di variabilità del dato. Lo strumento statistico che consente di seguire il processo produttivo e di stabilire se un processo è o meno sotto controllo è costituito dalle **carte di controllo**. La carta di controllo è un utile tecnica per il monitoraggio del processo e consente di correggere ed eliminare eventuali errori intervenuti nel processo causati da fonti di variabilità anomale. Il quinto capitolo quindi è relativo alla costruzione delle carte di controllo. Suddiviso nei due paragrafi, in cui sono considerate carte di controllo univariate e multivariate, vengono sviluppate le specifiche basi teoriche e l'applicazione al caso studio. La tesi si conclude con alcune considerazioni che, sulla base delle difficoltà incontrate, descrivono le scelte effettuate e gli opportuni suggerimenti. Nell'appendice sono riportati i comandi utilizzati con il software "R", l'output con i relativi risultati dei test statistici applicati e le funzioni create per tutto lo sviluppo dell'analisi statistica (grafici e carte di controllo).

---

# Capitolo 1

## Controllo statistico di processo

Tra le varie definizioni, quella più classica del concetto di qualità si basa sul presupposto che beni e servizi devono soddisfare le richieste di coloro che li utilizzano. Ciò significa che la qualità deve essere appropriata all'uso nei due aspetti generali:

- qualità di progetto
- conformità alle normative.

La normativa UNI EN ISO 9001 nel capitolo 8<sup>1</sup> descrive, nell'ambito della valutazione della qualità, strumenti e tecniche di natura statistica, finalizzate al miglioramento della qualità. Le caratteristiche della qualità sono spesso valutate in relazione alle specifiche e, come nel nostro caso per un prodotto manifatturiero, le specifiche sono le misure stabilite per alcune caratteristiche tangibili dei componenti e dei sottocomponenti che costituiscono il prodotto. Il valore di misura che corrisponde al valore desiderato per una caratteristica di qualità è definito **valore nominale**. Il maggiore valore ammissibile o accettabile per una caratteristica viene definito limite di **specifica superiore** (*Upper specification limit* USL) mentre il valore minore accettabile per una caratteristica di qualità è il limite di **specifica inferiore** (*Lower specification limit* LSL).

Generalmente i prodotti che non rientrano nei limiti per alcune caratteristiche si definiscono non conformi alle specifiche e vengono scartati a seconda del grado di rilevanza della non conformità<sup>2</sup>.

Negli ultimi cinquant'anni l'idea di qualità è stata intesa soprattutto come capacità di un bene prodotto di soddisfare determinate prestazioni o requisiti (da definire a priori e da misurare in una forma in qualche modo oggettiva, sulla base di difettosità del singolo bene o più propriamente, in riferimento alla omogeneità della produzione in un determinato periodo di tempo).

A questa visione che privilegia gli aspetti tecnici e manifatturieri, si possono collegare le tecniche organizzative riguardanti la qualità e l'impiego degli strumenti statistici che si riferiscono ai piani di campionamento in fase di accettazione e collaudo da un lato, e alla verifica dello stato di regolarità del processo di produzione dall'altro, dando origine a

---

<sup>1</sup> Misurazioni analisi e miglioramento.

<sup>2</sup> Riferimento alla normativa per il trattamento della non conformità.

quell'apparato di metodi e di tecniche noto con il nome di Controllo Statistico della Qualità (*Statistical Quality Control, SQC*)<sup>3</sup>.

Fanno parte dei metodi statistici del controllo qualità:

- il controllo statistico di un processo produttivo (SPC);
- la programmazione degli esperimenti;
- il campionamento in accettazione (anche se in minor misura).

La carta di controllo é una delle principali tecniche adottate per il controllo statistico di un processo produttivo SPC ed è un utile tecnica per il monitoraggio del processo stesso. Infatti, quando sono presenti fonti di variabilità anomale, può capitare di osservare alcuni punti al di fuori dei limiti di controllo<sup>4</sup>. Tale indicazione segnala la necessità di effettuare indagini per rimuovere e correggere gli eventuali errori intervenuti.

Un esperimento programmato è invece estremamente utile per scoprire quali sono state le variabili che hanno maggiormente influenzato la realizzazione del livello qualitativo osservato e per sapere in che misura esso è influenzato dalle variazioni di livello dei fattori studiati. Questo strumento è quindi essenziale per ridurre la variabilità della qualità del prodotto e nel determinare i livelli che devono assumere le variabili controllabili per avere una resa ottimale. Il campionamento in accettazione è strettamente legato all'analisi ed alla verifica del prodotto finito o delle materie prime impiegate. Il campionamento in accettazione è definibile come l'ispezione e la classificazione di un campione di unità, selezionate a caso da un lotto, per la valutazione della qualità dello stesso lotto. Si possono distinguere tre fasi di ispezione:

- ispezione realizzata immediatamente dopo la produzione e prima della consegna al cliente;
- ispezione in entrata, ovvero quando il controllo dei pezzi è effettuato prima della loro acquisizione;
- ispezione di rettifica, quando cioè i campioni esaminati possono portare all'accettazione o al rifiuto del lotto e in quest'ultimo caso i pezzi esaminati possono essere sottoposti a un'ulteriore lavorazione o scartati.

Tra i vari strumenti che offre l'SPC oltre alle carte di controllo troviamo di rilevante importanza l'EDA (*Exploratory Data Analysis*) che con l'uso di grafici (istogrammi, grafici di dispersione, boxplot, grafici di Pareto) definisce una prima linea guida nell'analisi dei fattori di variabilità dei dati. La loro efficacia nel miglioramento della qualità dipende tuttavia da come il management aziendale assimila ed applica tali strumenti. Devono essere addestrate

---

<sup>3</sup> Iniziale formulazione del testo di Juran, "Quality Control Handbook".

<sup>4</sup> Vedi capitolo quattro: carte di controllo.



non solo le persone direttamente dedicate al miglioramento della qualità della produzione ma anche e soprattutto quelle che dirigono l'azienda. Solo un'applicazione continua e sistematica degli strumenti soprascritti, sostenuta dal management, consente di fare di questi strumenti una parte integrante del modo di pensare alla programmazione della produzione ed al miglioramento della qualità.

Ogni processo produttivo è sempre soggetto a fonti di variabilità intrinseca o naturale. Questa variabilità naturale, o rumore di fondo, è il risultato dell'effetto cumulato di molti piccoli ma non eliminabili fattori costanti o casuali. Un processo la cui variabilità sia provocata solo da fattori casuali verrà detto sotto controllo.

Tra le varie fonti di variabilità ne esistono tuttavia alcune che influiscono sulla qualità risultante dei prodotti e possono essere solo occasionalmente presenti nel processo produttivo.

Le cause "identificabili" possono essere raggruppate in cinque categorie:

1. uomo
2. macchina
3. metodo
4. materiale
5. ambiente

La variabilità prodotta da questi fattori è molto più evidente di quella prodotta da fattori casuali e dà luogo ad una prestazione del processo non ottimale.

Le fonti di variabilità che non sono riconducibili a fattori casuali vengono chiamate "fattori specifici". Un processo che stia funzionando in presenza di fattori specifici verrà detto fuori controllo.

- cause di tipo accidentali → Processo in controllo
- cause identificabili o determinabili → Processo fuori controllo

L'obiettivo primario del controllo statistico di un processo produttivo è quello di individuare il più velocemente possibile il verificarsi di fattori specifici. Quanto più veloce è l'individuazione delle cause, tanto prima potranno essere avviate le azioni correttive, così da evitare la produzione di una elevata quantità di pezzi di qualità non accettabile.

Spesso le carte di controllo vengono usate per controllare i parametri di un processo e per determinare la capacità del processo stesso. Si tenga presente comunque che lo scopo del SPC è di eliminare o ridurre il più possibile la variabilità all'interno del processo stesso.

In questo contesto viene sviluppato, nel seguente elaborato, l'analisi ed il controllo statistico di un processo produttivo per la produzione di fibre tessili (realtà industriale manifatturiera).



---

# Capitolo 2

## Fibre acriliche

### 2.1 Introduzione

Nel seguente capitolo vengono descritte alcune nozioni di base riguardanti le **fibre acriliche** ed il processo produttivo attuato nello stabilimento MONTEFIBRE di P.to Marghera riportando tra le varie definizioni e fasi produttive quelle più interessanti ed esplicitamente coinvolte per i dati d'analisi proposti nel capitolo successivo.

Tipologie diverse di fibra hanno proprie caratteristiche che le definiscono. Le caratteristiche devono essere determinate e costanti all'interno di ristretti limiti qualitativi che devono permettere di soddisfare le esigenze pratiche di impiego. Montefibre, come altri produttori di fibra, è affiliata alla CIRFS (*Comitè International del le Rayon et des Fibres Syntetiques*) fondata nel 1928, che nel BISFA (*Bureau International pour la Standardisation des Fibres Artificielles*), armonizza i metodi d'analisi ed i metodi di campionamento tra produttori di fibra, ne definisce le varie terminologie, procedure, e tests.

Infine verranno riportate le fasi di estrapolazione dei dati nelle misure di laboratorio esplicitandone le metodiche d'analisi.

### 2.2 Storia e sviluppo delle fibre sintetiche

Lo sviluppo e la produzione delle fibre sintetiche (fibre di sintesi a partire da composti chimici semplici) sono piuttosto recenti; il ritardo è da connettere ad una carenza della struttura dei **polimeri**<sup>5</sup> naturali (quali cellulosa, gomma, fibre naturali), il cui studio risultava difficoltoso in quanto chimicamente essi si presentavano infusibili, non reattivi e insolubili, del tutto dissimili dalle usuali sostanze chimiche.

Solo negli anni '20 grazie ad un ricercatore tedesco (Staudinger) si comprese che i polimeri naturali sono formati da lunghe catene filiformi (macromolecole lineari), riproducibili mediante reazione di appropriate molecole relativamente semplici.

Anche se la data di nascita delle fibre sintetiche si fa risalire alla produzione di una fibra a base clorovinilica<sup>6</sup>, di fatto la prima vera fibra sintetica prodotta industrialmente e destinata ad

---

<sup>5</sup> Macromolecole che compongono la struttura delle materie plastiche; l'unità base della struttura è il monomero.

<sup>6</sup> Cloruro di vinile.

avere grande impatto sul mercato, deve essere considerata la fibra poliammidica<sup>7</sup> lanciata da DuPont con il nome commerciale Nylon (produzione sperimentale anno 1938).

Verso la fine degli anni '30 e all'inizio degli anni '40 si sviluppano e vengono prodotti vari tipi di fibre tra cui la fibra acrilica brevettata nel 1942; tuttavia a causa del periodo bellico gli impianti industriali verranno attivati solo agli inizi degli anni '50.

E' importante notare come nel giro di pochi anni vengano messe a punto tutte le fibre sintetiche di maggiore importanza nel settore tessile (poliestere, poliammidica, acrilica).

Nel contesto della produzione delle fibre sintetiche, l'Italia fa la comparsa nel 1939 con piccole produzioni di nylon (Montecatini).

I produttori di fibre sintetiche ripetono negli anni '60 il grande sforzo che avevano fatto negli anni '30 i produttori di fibre artificiali; negli anni 1960-1970 la produzione italiana rappresenta circa il 5% della produzione mondiale. A partire degli anni '70-'80 si determinerà però un lento declino a causa di una mancata razionalizzazione degli impianti, carenza di ricerca e sviluppo, sovrapproduzioni, crisi petrolifere e spostamento delle produzioni da Paesi di antica industrializzazione (Europa, USA, Giappone) a Paesi di più recente industrializzazione dell'Estremo Oriente (Cina, Taiwan, Corea del Sud).

Come per le altre fibre chimiche, il primo obiettivo da raggiungere postosi nel passato fu quello di una parziale sostituzione delle fibre naturali (lana e cotone). Si adattarono quindi le caratteristiche della fibra ai cicli di lavorazione tessile (filatura) già esistenti. In seguito però si sono individuate anche delle nuove applicazioni tanto che, con alcuni prodotti, si è arrivati a staccarsi dal settore tessile stesso.

A partire dallo stesso polimero si possono ottenere delle fibre per applicazioni diverse. Quindi, nei limiti consentiti dalla materia prima, si possono progettare e produrre fibre adatte a particolari usi e/o lavorazioni.

Le fibre acriliche di Montefibre sono usate sia nel campo tessile che in quello tecnico. L'utilizzo della fibra acrilica è normalmente destinato al 90% nell'abbigliamento, all'8% per l'arredamento ed il rimanente 2% per usi tecnico industriali.

Le tipologie di prodotto finale ottenibili adoperando la fibra acrilica sono parecchie e di utilizzi diversi. Le caratteristiche che la rendono appetibile sul mercato riguardano sia la qualità e la facilità di manutenzione del manufatto, che la lavorabilità della fibra durante tutto il procedimento produttivo.

---

<sup>7</sup> Prodotto di condensazione tra molecole caratterizzate da due gruppi reattivi amminici e da due gruppi reattivi carbossilici.

## 2.3 Classificazione delle fibre chimiche

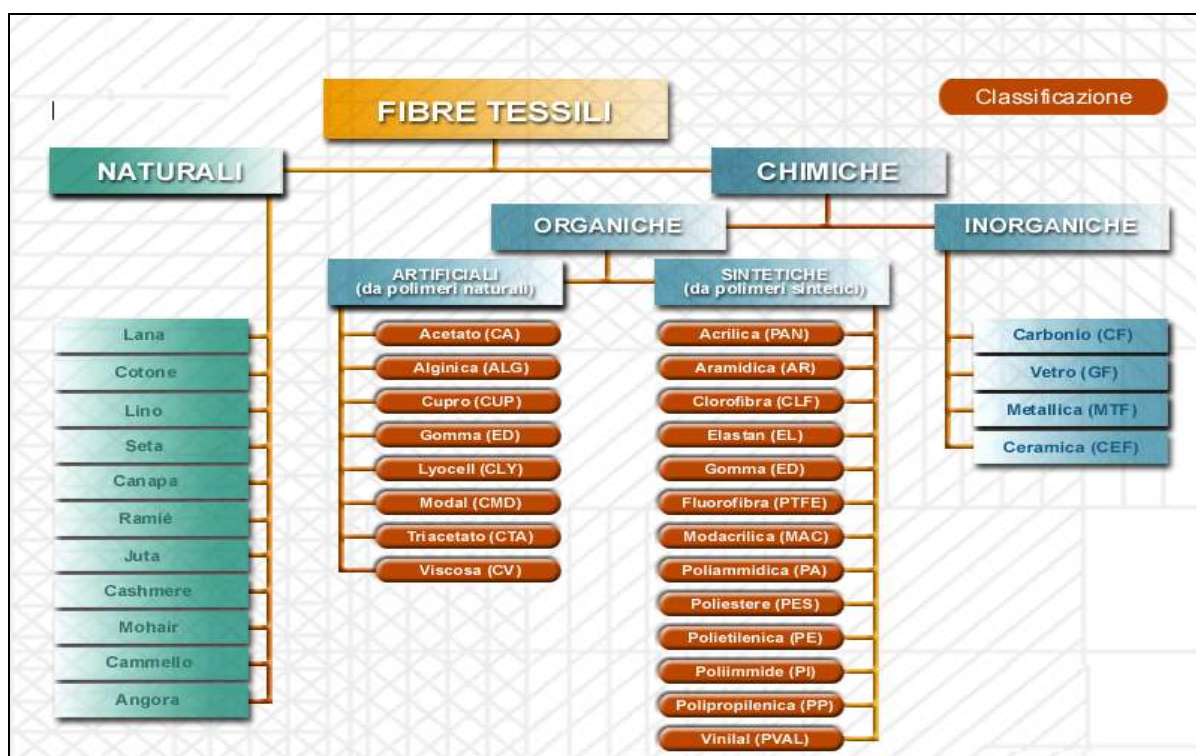
La norma ISO 2076-1999 (E) fornisce un elenco delle denominazioni comuni usate per designare le differenti categorie di fibre chimiche correntemente prodotte su scala industriale per usi tessili ed altri impieghi.

Ciascuna denominazione comune viene definita mediante attributi, normalmente basati su differenze chimiche, esplicitate mediante formule chimiche, a cui spesso corrispondono differenti proprietà distintive (vedi figura 1).

Come ausilio alla comunicazione tecnica e commerciale, alle fibre vengono assegnati codici o abbreviazioni di identificazione, basati sulle denominazioni comuni e costituite da lettere alfabetiche (da 2 a 4).

Si sottolinea che le denominazioni comuni, su cui si basa la legislazione di etichettatura della composizione fibrosa, devono essere distinte e non confuse con i marchi commerciali che ciascun produttore può utilizzare per il riconoscimento del proprio prodotto.

FIGURA 1



## 2.4 Descrizione generale del processo

### Produzione delle fibre acriliche

Nel presente paragrafo si descrive sinteticamente il processo di produzione della fibra acrilica nelle sue principali fasi. Lo stabilimento MONTEFIBRE di P.to Marghera produce fibre acriliche, la fibra (greggia e/o tinta) nasce come nastro continuo di più fibre (tow) che può essere commercializzato tale e quale o venduto in forma di fiocco ottenuto tagliando il tow. Il processo di produzione della fibra acrilica si articola nella seguente serie di operazioni unitarie, che partendo dalla produzione del polimero permettono di arrivare alla fibra finita:

- Polimerizzazione
- Dissoluzione del polimero in solvente
- Filatura ad "umido", produzione della fibra
- Tintura della fibra
- Taglio, imballo e spedizione

### 2.4.1 Polimerizzazione

In questo impianto viene realizzata la reazione di polimerizzazione tra **Acrilonitrile** e **Acetato di Vinile** in sospensione acquosa in presenza di un catalizzatore radicalico. La reazione si effettua in reattori muniti di agitatori e "incamiciati", asportando il calore di reazione mediante circolazione di acqua nella camicia esterna del reattore. L'aggiunta del comonomero<sup>8</sup> è necessaria per conferire buone caratteristiche tessili e di lavorabilità alla fibra una volta prodotta.

Dal polimero di solo Acrilonitrile (omopolimero<sup>9</sup>) si ottiene una fibra che ha caratteristiche meccaniche e chimico-fisiche particolari e che la rendono adatta per impieghi tecnici. Il polimero si separa dalla sospensione attraverso un processo di filtrazione ed i monomeri non reagiti vengono recuperati e rimessi in ciclo. Il polimero infine essiccato e macinato viene mandato al reparto in cui viene sciolto mediante opportuno solvente.

### 2.4.2 Dissoluzione del polimero in solvente

Viene preparata una soluzione viscosa "**dope**", di concentrazione opportuna sciogliendo il polimero in un solvente organico, **dimetilacetammide**, facendo molta attenzione alle condizioni di temperatura e tempo per avere la viscosità desiderata evitandone la degradazione e la formazione di "gel".

---

<sup>8</sup> Monomeri vinilici solfonati.

<sup>9</sup> Formato da monomeri uguali.

La soluzione ottenuta subisce le operazioni di miscelazione, disaerazione, filtrazione, eventualmente addizionata con additivi opacizzanti e stoccata in serbatoi per essere successivamente inviata al reparto di filatura.

### 2.4.3 Filatura

È l'operazione che permette di produrre la fibra tessile sotto forma di nastro continuo multibava (**tow**) partendo dalla soluzione di polimero (**dope**).

Questa operazione viene condotta in particolari macchine di filatura che comprendono anche successive sezioni di finitura della fibra ottenuta.

Le operazioni che si realizzano nelle macchine di filatura sono le seguenti:

- Coagulo: il “dope” (soluzione di polimero nel solvente) alimentato dalle pompe passa attraverso delle filiere<sup>10</sup> (estrusione<sup>11</sup>) immerse in una soluzione di acqua/solvente, in questo ambiente il dope impoverendosi di solvente permette la precipitazione del polimero formando così il fascio di filamenti (tow).
- Lavaggio, Stiro, Tintura: il lavaggio si fa con acqua per eliminare il solvente, la fibra viene quindi stirata tra rulli che ruotano a velocità diverse per migliorare le caratteristiche meccaniche, nella zona di tintura la fibra può venire colorata.
- Ensimaggio: è l'operazione con la quale in una vasca vengono aggiunte alla fibra sostanze lubrificanti, antistatiche, ammorbidenti che sono necessarie per la manipolazione e le lavorazioni tessili.
- Essiccamento: passando su rulli caldi l'acqua residua viene eliminata e vengono chiuse eventuali fessurazioni eventualmente presenti sulla fibra.
- Crettatura: questa operazione serve per "arricciare" la fibra per permettere le lavorazioni tessili.
- Stabilizzazione/Annilazione: la fibra subisce un trattamento in autoclave (vari cicli vuoto/vapore) per farle acquistare stabilità dimensionale e per migliorare le caratteristiche meccaniche e tessili.

---

<sup>10</sup> Le filiere sono costituite da un disco metallico di circa 15 cm (tantalo) su cui sono stati realizzati migliaia di fori del diametro di alcune decine di micron.

<sup>11</sup> Il dope incontra prima la tela filtrante e quindi il distributore che lo suddivide su tutta la superficie interna della filiera e per effetto della pressione è costretto a passare attraverso i fori della filiera ed entrare nel bagno di filatura sotto forma di getti liquidi cilindrici.

### 2.4.4 Tintura della fibra

#### Fibra tinta in Linea (TL)

La fibra presenta, grazie alla presenza di siti “tingibili” impartiti dal sistema catalitico al polimero acrilico, la possibilità di essere tinta mediante particolari coloranti.

La reazione di tintura della fibra si effettua “in linea” tramite appositi bagni di colorante collocati in opportune posizioni della macchina di filatura (dopo l'estrusione), le miscele vengono preparate dosando i singoli coloranti solubili in funzione della ricetta desiderata con un sistema completamente automatizzato e successivamente inviate al bagno di tintura.

#### Fibra tinta in "Pasta" (TP)

Il metodo TP non prevede l'utilizzo di coloranti in grado di reagire con i siti tingibili della fibra ma l'aggiunta di dispersioni di pigmenti che, opportunamente macinati ed omogeneizzati in appositi mulini, vengono miscelati al dope prima della fase di filatura (prima dell'estrusione), in tale modo dal bagno di coagulo esce una fibra già tinta (pigmentata). L'utilizzo di pigmenti ed il particolare processo produttivo, rendono questa fibra idonea negli *end uses* dove è necessaria un'elevata solidità del colore alla luce ed alle intemperie (tendaggi ed ombrelloni da sole).

### 2.4.5 Taglio ed imballo

La fibra proveniente dalla filatura:

in parte rimane come **tow** (nastro continuo multibava)

in parte viene tagliata a “**fiocco**” con lunghezze variabili in funzione dell'uso tessile finale

- **taglio a fiocco laniero**

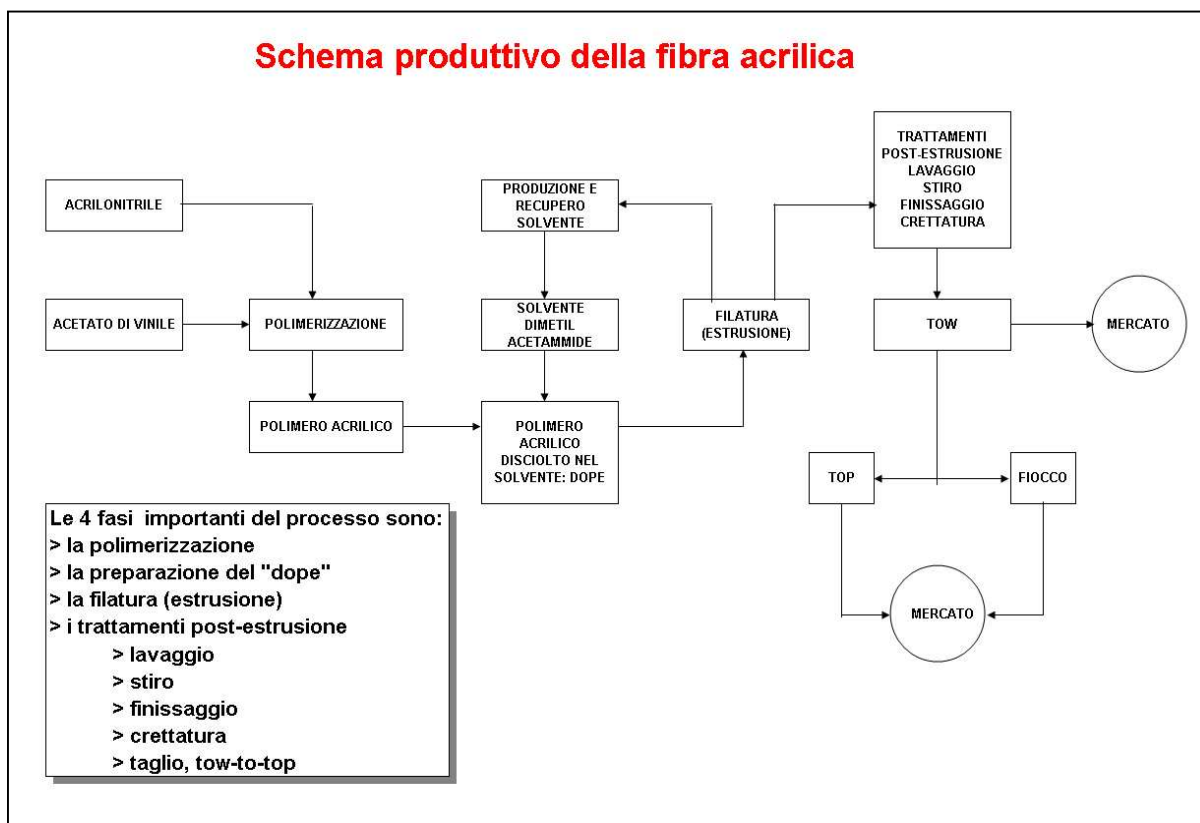
- **taglio a fiocco cotoniero**

Successivamente sia il fiocco che il tow vengono inviati alle presse per il confezionamento in balle per la spedizione ai clienti.

Le operazioni di **taglio** (trasformazione tow → fiocco), di **pressatura** e di **imballo** sia del fiocco che del tow sono completamente automatizzate.

Di seguito viene raffigurato uno schema riassuntivo della produzione di fibra acrilica.





## 2.5 Le principali caratteristiche fisico-meccaniche

L'utilizzo delle fibre finalizzato al settore tessile passa attraverso cicli di lavorazione in grado di trasformare il prodotto nei vari manufatti: tessuti, maglieria, velluti ecc.

Storicamente i cicli principali di filatura derivano dalle due fibre naturali più utilizzate: il ciclo "cotoniero" (dedicato a fibre corte e generalmente di uguale lunghezza come il cotone) ed il ciclo "laniero" (fibre mediamente lunghe e con lunghezza variabile, tipico del vello di pecora).

Sono possibili comunque anche altri cicli tessili che generalmente risultano dedicati alle fibre chimiche (ad esempio nell'uso del filo continuo); in ogni caso è di basilare importanza che i prodotti presentino delle caratteristiche tali da permettere la loro agevole trasformazione attraverso i cicli tessili.

**Proprietà fisiche:** tra i parametri principali che definiscono le caratteristiche e la processabilità di una fibra chimica vi sono:

- Lunghezza (o taglio)
- Massa lineare (o titolo)
- Proprietà dinamometriche (o di trazione)
- Ensimaggio (od oli di prima finitura)

### 2.5.1 Lunghezza della fibra

La lunghezza della fibra rappresenta una delle grandezze dimensionali fondamentali sulle quali si basa la definizione morfologica di fibra intesa come elemento di materia caratterizzato da elevato rapporto lunghezza-finezza. Naturalmente ha senso parlare di valore di lunghezza solo nel caso in cui una fibra è caratterizzata da una lunghezza definita (fibra discontinua) con valori normalmente compresi tra qualche centimetro e qualche decimetro.

La lunghezza della fibra, oltre a condizionare strettamente i sistemi e le regolazioni delle macchine tessili, influenza molte delle caratteristiche dei filati.

### 2.5.2 Massa lineare o titolo

L'altro parametro che definisce le dimensioni di una fibra è la misura della sua finezza, o più precisamente della sua sezione trasversale. Normalmente per i prodotti usati nei campi tecnici, tale proprietà è espressa dalla misura del diametro; qualora si tratti di un prodotto a sezione circolare è costante per tutta la sua lunghezza.

Le fibre sono caratterizzate invece da una grande varietà di sezioni, che a volte possono pure variare nel senso longitudinale. Anche per questo aspetto comunque, le fibre chimiche presentano una notevole semplificazione rispetto alla variabilità riscontrabile per le fibre naturali e il modo universalmente usato per definire la dimensione trasversale, questo almeno per le fibre d'impiego tessile, è la massa per unità di lunghezza.

Tale grandezza, nota con il termine "titolo" (o più precisamente "densità lineare") ha come unità di misura il tex definito come la massa espressa in grammi di un chilometro di fibra. Si tratta di un'unità derivata dal Sistema Internazionale di Misura, che prevede multipli e sottomultipli, di cui i più usati sono il decitex (dtex) per le fibre e i filamenti ed il Kilotex (Ktex) per tow e top.

E' opportuno accennare anche ad una preesistente unità di misura, che tende ad essere sostituita, è cioè il denaro, definito come la massa espressa in grammi di 9000 metri di fibra (1 tex = 9 den).

Assumendo che una fibra sia assimilabile ad un corpo cilindrico di sezione circolare uniforme e di densità costante, si ottiene una notevole relazione tra massa lineare e diametro:

$$T = \frac{\pi}{4} d D^2 10^{-2} \quad \text{oppure} \quad D = 11,3 \sqrt{\frac{T}{d}}$$

dove :

T = Titolo in dtex

d = densità della fibra in g/cm<sup>3</sup>

D = diametro della fibra in μm

L'importanza della finezza di una fibra ai fini della definizione delle caratteristiche qualitative di un filato o di un manufatto tessile è intuitivamente comprensibile (si pensi alla “mano” di un manufatto inteso come morbidezza e soffici ta).

La gamma teorica delle masse lineari ottenibili in fase di produzione   notevole ma di fatto esistono intervalli definiti da vincoli di carattere processuale e di trasformazione tessile che limitano la gamma a valori normalmente compresi tra 1 dtex e 17 dtex.

E' importante comunque sottolineare l'attuale tendenza a produrre fibre (o filamenti) pi  fini di 1 dtex (microfibre) sia per evoluzione tecnologica dei macchinari sia per esigenze qualitative di settori specifici (tipo abbigliamento sportivo) sia per l'introduzione di nuovi prodotti (come materiali microfibrosi).

Risulta interessante notare inoltre come la densit  del materiale di cui   costituita la fibra influenzi il titolo e che pertanto a parit  di titolo, fibre di diversa natura (poliestere, nylon, acrilico, ecc.) abbiano in realt  diverso diametro.

### 2.5.3 Propriet  dinamometriche

Per propriet  dinamometriche (o a trazione) si intende il comportamento di una fibra sottoposta a forze e deformazioni nel senso longitudinale.

E' opportuno definire alcuni termini e concetti di carattere generale:

Forza di rottura: forza massima sopportata dalla fibra in una prova di trazione condotta fino a rottura e normalmente espressa in centi-newton (cN).

Allungamento a rottura: allungamento della fibra rilevato in corrispondenza della forza di rottura. L'allungamento pu  essere espresso come aumento in lunghezza assoluta della provetta o, pi  comunemente, come aumento relativo percentuale.

Tenacit  di rottura: rapporto tra la forza di rottura e il titolo, normalmente espressa in cN/tex. Nei settori tecnici   invece comunemente usata la forza rapportata alla sezione trasversale della fibra (sforzo di trazione) ed   normalmente espressa in Mpa (N/mm<sup>2</sup>).

Si tratta di due grandezze specifiche (unitarie) la cui relazione   la seguente:

$$F_s = 10 d T$$

dove:

T = tenacit  in cN/tex

F<sub>s</sub> = forza per unit  di sezione in N/mm<sup>2</sup> (Mpa)

d = densit  in g/cm<sup>3</sup>

Giacch  le due grandezze sono legate tramite la densit , anche in questo caso deriva che il confronto tra i materiali di diversa densit  assumer  un valore diverso a seconda della grandezza impiegata (le fibre hanno densit  pi  bassa rispetto a materiali tipo fili metallici).

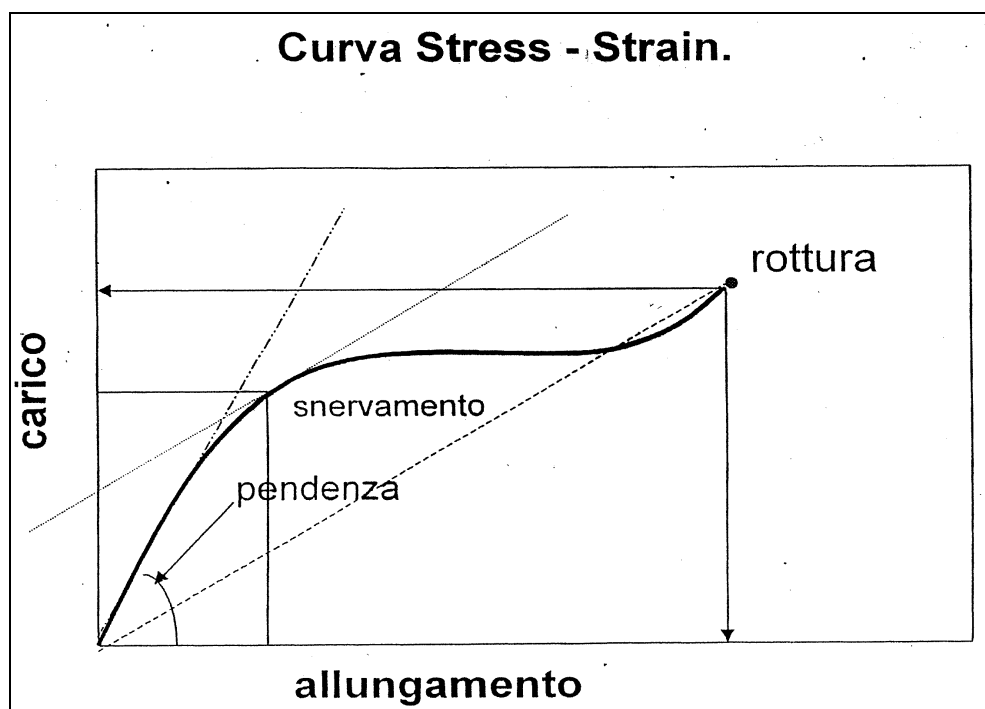
Curva forza-allungamento: è la curva che definisce il comportamento di una fibra sottoposta a trazione. Normalmente la curva è definita in termini di tenacità-allungamento percentuale, vedi figura 2 (stress-strain).

Lavoro di rottura: lavoro necessario a portare la fibra alla rottura.

Lavoro specifico di rottura: rapporto tra il lavoro di rottura ed il titolo. Esso è dato dall'area sottesa alla curva tenacità-allungamento.

Modulo elastico: rappresenta la proprietà che definisce la resistenza alla deformazione. Nelle fibre il modulo è espresso come rapporto della tenacità rispetto alla deformazione, calcolato tra due tratti della curva o a un punto tangente della curva. Si fa riferimento normalmente al modulo iniziale di elasticità inteso come misura dell'inclinazione del tratto iniziale della curva tenacità-allungamento percentuale ed espresso in cN/tex.

FIGURA 2



### 2.5.4 Ensimaggio

Un parametro che non è strettamente legato alla struttura della fibra, ma che rappresenta un fattore fondamentale ai fini della lavorabilità, è relativo ad una miscela di additivi applicati sulla fibra in fase di processo di produzione, che va sotto il nome di ensimaggio (olio di prima finitura o finish). Si tratta di una sostanza, o meglio di una miscela di sostanze, che conferisce

alla fibra chimica un insieme di proprietà antistatiche, lubrificanti ed ammorbidenti che permettono un buon rendimento sia in produzione che in trasformazione tessile.

La finalità dell'applicazione dell'ensimaggio è quella di ottimizzare e prevenire la formazione di cariche elettrostatiche, di rendere idonei per la lavorazione gli attriti fibra-fibra e tra fibra e gli organi di lavorazione. Finish opportuni impartiscono anche caratteristiche di "mano" ai manufatti finali.

E' pertanto costante preoccupazione del produttore assicurarsi, mediante adeguati controlli, il corretto dosaggio dell'ensimaggio sia in termini qualitativi (di composizione chimica) che quantitativi (determinato come contenuto percentuale base peso della fibra).

Quelle elencate sono solo alcune delle caratteristiche della fibra che vengono coinvolte nell'elaborato, tuttavia si possono notare nella figura sottostante un elenco complessivo delle caratteristiche.

FIGURA 3



## **2.6 Misure delle proprietà chimico-fisiche delle fibre**

L'industria utilizza per il proprio fabbisogno materie prime che trasforma, durante il ciclo di produzione, in prodotto finito. È pertanto necessario controllare le caratteristiche delle materie prime, delle varie fasi di trasformazione e del prodotto finito, per ottenere alla fine del processo un materiale che soddisfi criteri prestabiliti. Ogni caratteristica del materiale in esame deve avere un determinato valore, con una tolleranza compresa tra un minimo ed un massimo: questi limiti vengono chiamati “specifiche” di quella caratteristica. Il valore rilevato in laboratorio sarà ritenuto accettabile se compreso nell'intervallo delle specifiche<sup>12</sup>, e non altrimenti.

Lo scopo del laboratorio fisico e chimico è il controllo delle proprietà fisiche e chimiche delle fibre prodotte nello stabilimento, con le modalità e gli strumenti specifici.

### **2.6.1 Analisi del laboratorio fisico (su bava)**

#### **Titolo**

L'analisi del titolo della bava viene generalmente effettuata su campioni significativi di tow o di fiocco. Il titolo della bava è un dato assai importante perchè da esso dipendono in modo notevole le possibilità ed i limiti di applicazione della fibra stessa. Infatti, ad ogni titolo corrisponde un limite di filabilità ben preciso che non si può superare senza rischio di alterare le caratteristiche di regolarità del filato che si vuole ottenere. Le fibre di diverso titolo non presentano uguali proprietà di “mano”, cioè non danno uguali sensazioni di morbidezza, nervosità, ecc.

Per questo bisogna che il titolo medio della fibra sia come stabilito ed entro limiti prefissati. Anche la variabilità del titolo delle bave deve essere contenuta entro un intervallo il più stretto possibile perchè essa influisce negativamente sulla regolarità dei filati allestiti.

Il controllo del titolo non va eseguito misurando il diametro delle fibre in quanto sarebbe un'operazione complessa e dai risultati poco precisi, considerando inoltre che le sezioni delle fibre in genere non sono rotonde.

Nel caso particolare del prodotto Montefibre la sezione della fibra risulta reniforme, in altri processi tecnologici di produzione di fibra acrilica invece, il filamento può avere sezioni ovoidali, ad “osso di cane”, tondeggianti, ecc.

Si preferisce pertanto seguire altri procedimenti indiretti, rivelatisi più semplici e pratici.

Il primo metodo che viene illustrato è quello per “pesata”.

---

<sup>12</sup> Valori di riferimento BISFA

Si tratta semplicemente di pesare un certo numero di fibre (generalmente un centinaio) di lunghezza esattamente definita ed applicare la formula:

$\text{Tex} = 1000 \times P/L$     dove P = peso della fibra, L = lunghezza della fibra

Ovviamente le bilance utilizzate risultano essere strumenti di particolare sensibilità ed accuratezza in quanto una pesata media per la determinazione del titolo di una fibra tessile si aggira nell' intorno di un milligrammo.

Altro strumento per la determinazione del titolo risulta essere il **vibroscopio**. Il funzionamento dell'apparecchio si basa sul seguente principio: una fibra si comporta come una qualsiasi corda quand'è tesa tra due punti fissi e sottoposta a vibrazione trasversale. È noto che la frequenza di vibrazione di una corda (o numero di oscillazioni al secondo) dipende dalla sua lunghezza, dal suo diametro e dalla tensione cui è sottoposta.

Nel nostro caso le lunghezza e la tensione della fibra sono fisse; la frequenza di vibrazione dipende solo dal diametro della fibra, tanto più la fibra ha diametro grande, tanto più basso è il numero delle sue oscillazioni, quando entra in vibrazione.

L'analisi del titolo bava al vibroscopio viene svolta nel seguente modo: vengono fatti alcuni prelievi di fibre dal campione da controllare e le si dispone su una tavoletta coperta di velluto. Servendosi della pinza si afferra una fibra, la si inserisce nella molletta tendifibre, di peso calibrato, e la si sistema nel morsetto dello strumento facendola appoggiare al vibratore.

Si agisce sulla manopola dell'oscillatore per variare la frequenza di vibrazione, allo scopo di ricercare quella particolare frequenza in cui la fibra entra in risonanza; per facilitare questa osservazione l'immagine della fibra è proiettata e ingrandita su di uno schermo. Quando sullo schermo si osserva il punto di massima oscillazione si può leggere sul disco graduato del vibroscopio il valore del titolo della fibra (espresso in tex). Si ripete questa operazione con altre fibre fino a controllarne il numero prescritto. Si ricava quindi il titolo medio che è quello preso in considerazione.

### **Caratteristiche fisico-meccaniche**

Oltre al titolo si controllano pure le caratteristiche fisico-meccaniche su bava. Si tratta di determinare i valori del carico e dell'allungamento a rottura delle fibre sottoposte a trazione al dinamometro.

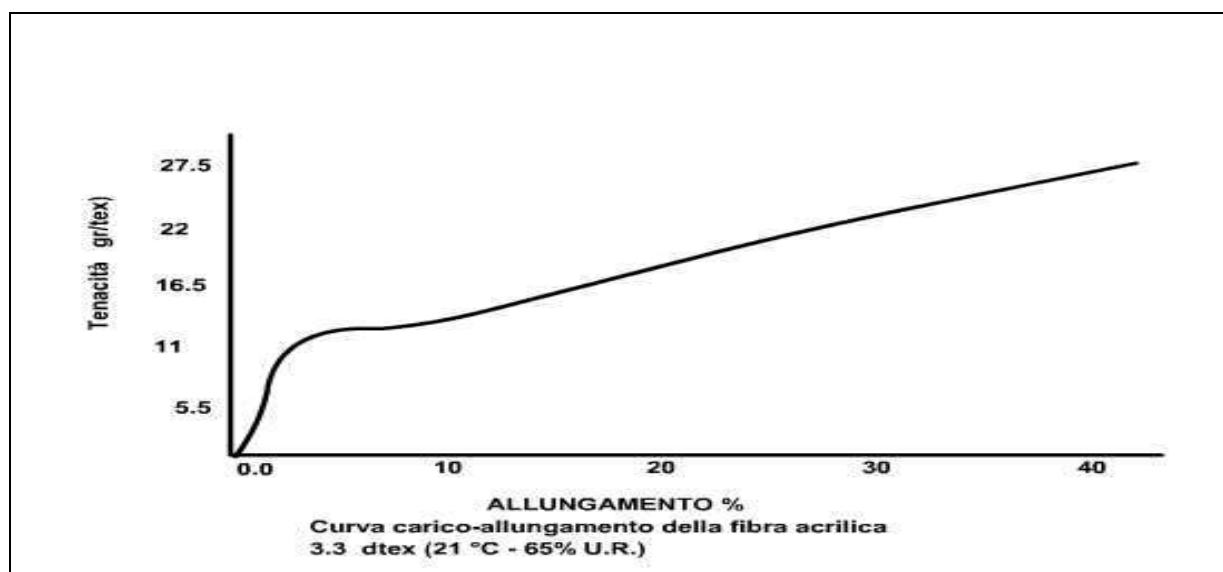
Il controllo di queste grandezze è di notevole importanza poichè esse influiscono in modo determinante sul carico ed allungamento a rottura dei manufatti (filati, tessuti) confezionati con il materiale in questione. Ad esempio, più una fibra è tenace, più resistente risulterà il filato prodotto con essa, con notevoli vantaggi per la lavorazione del filato stesso e per le applicazioni successive.

Il dinamometro è uno strumento molto preciso e versatile. Ha un morsetto fisso ed uno mobile tra i quali viene fissata la fibra in esame. Il morsetto mobile è montato su di una traversa che scende a velocità costante, a comando meccanico. Lo sforzo sopportato dal morsetto fisso, che è dovuto alla resistenza che la fibra oppone alla trazione cui è sottoposta, viene rilevato da un congegno elettronico e registrato su una carta diagrammabile che si muove a velocità nota e costante. Si ottiene una curva carico-allungamento (vedi figura 4) dalla quale si possono ricavare i valori di carico e di allungamento a rottura della bava.

Un altro valore interessante è dato dalla resistenza a rottura divisa per il titolo della fibra, che si denomina **tenacità**.

Infatti è ragionevole supporre che una fibra con basso titolo sia meno resistente di una avente titolo alto; ma se si rapportano i valori del carico all'unità di tex si può osservare che la resistenza specifica della fibra più grossa non si discosta molto da quella della fibra più sottile.

FIGURA 4



## 2.6.2 Analisi del laboratorio chimico

### Ensimaggio su fibra estraibile in alcool

Oltre alle misure di alcune caratteristiche fisico-meccaniche della fibra viene preso in considerazione il rilevamento, in laboratorio chimico, dell'ensimaggio estraibile.

Il metodo descrive la determinazione ponderale dell'ensimaggio su fibra estraibile a caldo in alcool. Il principio su cui si basa tale metodo è il seguente: il campione di fibra esattamente pesato viene posto in un estrattore secondo Soxlet (ove il solvente che imbeve la fibra viene continuamente rinnovato per distillazione) e mantenuto in estrazione per un tempo



determinato. La soluzione di ensimaggio in alcool è fatta evaporare permettendo così di determinare per pesata la quantità di finish estratto dalla fibra.

## **2.7 LIMS**

I valori determinati con i metodi descritti sopra, vengono inseriti in un sistema informatico in grado di gestire l'acquisizione, l'elaborazione e l'archiviazione di tutti i dati generati in uno o più laboratori o processi, acquisiti con strumentazione diversa. Tale sistema è il LIMS (*Laboratory Information Management System*). Il LIMS è un sistema computerizzato disegnato per catturare, analizzare, riportare e organizzare le informazioni e i dati via banca dati, non solo per le funzioni principali all'interno del laboratorio, ma anche nell'organizzazione globale del lavoro. I dati secondo opportuni metodi di ricerca possono essere estrapolati dalla rete intranet aziendale ed in tal modo essere aggiornati.



---

# Capitolo 3

## Validazione di un metodo per analisi di laboratorio e quantificazione dell'incertezza nelle misure analitiche

### 3.1 Introduzione

Uno dei termini ricorrenti nelle misure di laboratorio è l'**incertezza**. La stima dell'incertezza di misura da parte dei laboratori di prova, di taratura e clinici, fornisce un'indicazione del grado di fiducia che l'utente finale può riporre nei risultati di misurazione, cosa particolarmente importante quando su di essi si basano decisioni rilevanti come ad esempio la conformità a leggi, specifiche tecniche o norme. Tale esigenza è stata tradotta, già da alcuni anni, nella norma UNI CEI EN ISO/IEC 17025:2000 (requisiti generali per la competenza dei laboratori di prova e di taratura).

I principi generali per la stima dell'incertezza di misura sono descritti in dettaglio nella ISO *"Guide to the expression of uncertainty in measurement"* (ristampa del 1995), tradotta in italiano nel 1997 e successivamente adottata come norma UNI CEI ENV 13005:2000.

Tuttavia, la sua applicazione diretta nei laboratori di analisi chimiche e biologiche risulta di difficile attuazione. In particolare, mentre per la maggior parte delle misure di grandezze fisiche esistono campioni primari a cui fare riferimento, ed è possibile costruire delle catene di riferibilità fino ad essi, nella maggior parte delle misurazioni chimiche e biologiche l'applicazione di un modello matematico rigoroso presenta notevoli difficoltà, legate in gran parte alla variabilità delle matrici da analizzare, alla complessità dei metodi analitici ed alla carenza di materiali di riferimento certificati che garantiscano la riferibilità alle unità SI (Sistema Internazionale).

L'accreditamento in base alla norma ISO/IEC 17025 è la dimostrazione della riferibilità dei risultati delle misurazioni.

Per definire formalmente un livello di fiducia dei risultati, è essenziale che un risultato di misura sia riferibile ad un riferimento dato, quale una unità SI, un materiale di riferimento o, dove applicabile, un metodo specificato o empirico.

Sebbene il concetto d'incertezza di misura sia noto ai chimici da molti anni, è stata la pubblicazione nel 1993 della *"Guide to the Expression of Uncertainty in Measurement"* (GUM) da parte dell'ISO, in collaborazione con altri enti ed organizzazioni, a definire

formalmente le regole generali per valutare ed esprimere l'incertezza di misura. Il presente capitolo basato sulla guida EURACHEM illustra come i concetti della GUM possano essere applicati nelle misurazioni analitiche di laboratori. Per prima cosa viene introdotto il concetto d'incertezza e la distinzione tra incertezza ed errore, quindi segue una descrizione dei passaggi richiesti per la valutazione dell'incertezza con i relativi procedimenti.

Per la valutazione dell'incertezza l'analista deve esaminare attentamente tutte le possibili fonti d'incertezza. In pratica, uno studio preliminare permetterà di identificare rapidamente le fonti più significative d'incertezza ed il valore ottenuto per l'incertezza composta dipende quasi interamente dai contributi maggiori. Si può quindi ottenere una buona stima dell'incertezza concentrando l'impegno sui contributi preponderanti. Una volta valutata l'incertezza per un dato metodo applicato in un dato laboratorio (cioè per un dato procedimento di misurazione), la stima dell'incertezza ottenuta può essere applicata ai risultati ottenuti successivamente con lo stesso metodo, purché i dati pertinenti del controllo di qualità giustifichino questa estrapolazione.

### 3.2 Scopo e campo di applicazione

Alcune aree tipiche nelle quali sono necessarie misure di laboratorio, e nelle quali possono essere applicati i principi di questa Guida, sono:

- controllo della qualità e garanzia della qualità nell'industria manifatturiera;
- prove richieste per la verifica della conformità a norme, leggi o regolamenti;
- prove effettuate utilizzando un metodo concordato;
- taratura di campioni e apparecchiature;
- misure associate con lo sviluppo e la certificazione di materiali di riferimento;
- ricerca e sviluppo.

Dal momento che in numerosi settori sono stati adottati dai laboratori provvedimenti formali per la garanzia della qualità, questo sunto della Guida EURACHEM è in grado di illustrare come utilizzare per la stima dell'incertezza di misura i dati ottenuti con le seguenti procedure:

- valutazione dell'effetto delle fonti d'incertezza identificate sul risultato analitico per un particolare metodo adottato come **procedimento di misurazione** in un dato laboratorio;
- risultati ottenuti da procedure definite di controllo interno della qualità in un dato laboratorio;
- risultati ottenuti da studi collaborativi usati per validare metodi di analisi in un certo numero di laboratori competenti;

- risultati ottenuti in programmi di prove valutative usati per valutare la competenza analitica di laboratori.

L'ipotesi di partenza è che, sia quando si eseguono misurazioni, sia quando si valutano le prestazioni di un procedimento di misurazione, vengano attuati provvedimenti efficaci di garanzia e di controllo della qualità per assicurare che il processo di misurazione sia stabile e sotto controllo. Tali provvedimenti di solito includono, per esempio, personale qualificato, adeguata manutenzione, conservazione e taratura della strumentazione e dei reagenti, uso di idonei campioni di riferimento, procedimenti di misurazione documentati, uso di appropriati campioni di controllo e di carte di controllo.

### 3.3 Incertezza

#### 3.3.1 Definizione d'incertezza

La definizione del termine incertezza (di misura) può essere riassunta come: “parametro associato al risultato di una misurazione, che caratterizza la dispersione dei valori ragionevolmente attribuibili al misurando<sup>13</sup>”, dove il parametro può essere, per esempio, uno **scarto tipo**<sup>14</sup> (o un suo multiplo dato), o la semiampiezza di un intervallo avente un livello di fiducia stabilito.

L'incertezza di misura in generale, comprende più componenti. Talune di queste possono essere valutate dalla distribuzione statistica dei risultati di serie di misurazioni e possono dunque essere caratterizzate mediante scarti tipo sperimentali. Le altre componenti, anch'esse caratterizzabili mediante scarti tipo, sono valutate da distribuzioni di probabilità ipotizzate sulla base dell'esperienza o di informazioni di altro tipo.

La definizione d'incertezza citata mette l'accento sull'intervallo di valori che l'analista ritiene possano essere ragionevolmente attribuiti al misurando.

#### 3.3.2 Fonti d'incertezza

In pratica l'incertezza sul risultato può derivare da molte possibili fonti, come ad esempio, definizione incompleta del misurando, campionamento, effetti di matrice e interferenze, condizioni ambientali, incertezze di masse e di dispositivi volumetrici, valori di riferimento, approssimazioni e assunti incorporati nel metodo e nel procedimento di misurazione, nonché variazioni.

---

<sup>13</sup> Grandezza sottoposta a misurazione.

<sup>14</sup> Radice quadrata positiva della varianza.

### 3.3.3 Componenti dell'incertezza

Per stimare l'incertezza globale, può essere necessario considerare ciascuna fonte d'incertezza e trattarla separatamente per valutarne il contributo.

Ciascuno dei singoli contributi all'incertezza è denominato "componente dell'incertezza". Tra le varie componenti si possono trovare **incertezza tipo**, quando una componente dell'incertezza è espressa come scarto tipo. Per un risultato di misurazione  $y$ , l'incertezza totale, detta **incertezza tipo composta**, è uno scarto tipo stimato come la radice quadrata positiva della varianza totale ottenuta combinando tutte le componenti dell'incertezza, in qualsiasi modo stimate.

Per gli scopi della chimica analitica, nella maggior parte dei casi si dovrebbe utilizzare un'**incertezza estesa** ( $U$ ). L'incertezza estesa fornisce un intervallo all'interno del quale si può ritenere si trovi il valore del misurando con un più elevato livello di fiducia.  $U$  si ottiene moltiplicando l'incertezza tipo composta per un **fattore di copertura**  $k$  (tipicamente compreso tra 2 e 3). La scelta del fattore  $k$  dipende dal livello di fiducia desiderato. Per un livello di fiducia di circa il 95%,  $k$  è 2.

### 3.3.4 Errore e incertezza

È importante distinguere tra errore e incertezza. L'errore<sup>15</sup> è definito come "la differenza tra un singolo risultato di prova e il valore vero del misurando". Come tale, l'errore è un valore singolo.

In linea di principio, il valore di un errore noto può essere applicato come una correzione al risultato. L'incertezza, d'altra parte, ha la forma di un intervallo e, se stimata per un procedimento analitico e per un determinato tipo di materiale, può essere applicata a tutte le determinazioni di quel tipo. In generale, il valore dell'incertezza non può essere utilizzato per correggere un risultato di misurazione. L'incertezza del risultato di una misurazione non dovrebbe mai essere interpretata come se rappresentasse l'errore stesso, né come l'errore residuo dopo la correzione.

Un errore si considera costituito da una componente casuale e una sistematica.

**L'errore casuale** deriva da variazioni non prevedibili delle grandezze d'influenza. Questi effetti casuali danno origine a variazioni nelle osservazioni ripetute del misurando. L'errore casuale su un risultato analitico non può essere corretto, ma di solito può essere ridotto aumentando il numero di osservazioni.

---

<sup>15</sup> L'errore è un concetto ideale e gli errori non possono essere conosciuti con esattezza.

**L'errore sistematico** è definito come una componente di errore che, nel corso di un certo numero di analisi dello stesso misurando, rimane costante o varia in modo prevedibile. Esso è indipendente dal numero di misurazioni fatte e non può quindi essere ridotto aumentando il numero di analisi effettuate mantenendo costanti le condizioni di misurazione.

Gli errori sistematici a loro volta si distinguono in costanti e non costanti. Per esempio le inaccurately nella taratura a più punti di uno strumento, sono costanti per un dato livello del valore misurato, ma possono variare con esso. Gli effetti la cui entità varia in modo sistematico durante una serie di analisi, causati per esempio, da un controllo inadeguato delle condizioni sperimentali, danno origine a errori sistematici che non sono costanti. Il risultato di una misurazione dovrebbe essere corretto per tutti gli effetti sistematici significativi riconosciuti.

Un altro tipo di errore è **l'errore spurio**, o **errore grossolano**. Errori di questo tipo invalidano la misurazione e di solito derivano da errore umano o da un malfunzionamento strumentale. Le misurazioni per le quali sono stati identificati errori di questo tipo devono essere rifiutate e questi errori non devono mai essere inclusi nell'analisi statistica dei dati.

## 3.4 Misure analitiche e incertezza

### 3.4.1 Validazione del metodo

In pratica, l'idoneità allo scopo dei metodi analitici impiegati nelle analisi di routine si valuta di solito mediante studi di validazione del metodo. Tali studi forniscono risultati sulle prestazioni globali e su singoli fattori d'influenza, che possono essere utilizzati per la stima dell'incertezza associata con i risultati del metodo durante il suo uso corrente. Gli studi di validazione si basano sulla determinazione di parametri indicativi delle prestazioni globali del metodo. Questi si ottengono nella fase di sviluppo del metodo mediante uno studio interlaboratorio o seguendo protocolli di validazione interna. Lo scopo principale è l'identificazione e la rimozione (piuttosto che la correzione) degli effetti significativi. In questo caso si giunge a una situazione nella quale sono stati identificati la maggior parte dei fattori d'influenza potenzialmente significativi e ne è stata verificata, e dimostrata trascurabile, l'importanza rispetto alla precisione globale.

In queste circostanze, i dati a disposizione degli analisti sono principalmente risultati relativi alle prestazioni globali, documentazione della non-significatività della maggior parte degli effetti e misure degli effetti significativi residui.

Negli studi di validazione di metodi analitici quantitativi di solito vengono determinati uno o più dei seguenti parametri:

- **Precisione:** le principali misure della precisione includono lo scarto tipo di ripetibilità ( $S_r$ ), lo scarto tipo di riproducibilità ( $SR$ )<sup>16</sup> e la precisione intermedia. La ripetibilità indica la variabilità osservata all'interno del laboratorio, in un breve periodo di tempo, con lo stesso operatore, lo stesso strumento, ecc. e può essere stimata all'interno del laboratorio o mediante studi interlaboratorio. Lo scarto tipo della riproducibilità per un particolare metodo può essere stimato direttamente solo mediante uno studio interlaboratorio;  $SR$  indica la variabilità ottenuta quando laboratori diversi analizzano lo stesso campione. La precisione intermedia è in relazione alla variabilità dei risultati, viene osservata quando uno o più fattori, quali tempo, apparecchiature e operatori, cambiano all'interno dello stesso laboratorio; a seconda di quali fattori sono tenuti costanti si ottengono valori diversi. La precisione osservata di un procedimento analitico è una componente essenziale dell'incertezza globale, sia che venga determinata dalla combinazione di singole varianze o dallo studio relativo all'applicazione dell'intero metodo.
- **Scostamento sistematico:** lo scostamento sistematico di un metodo analitico si determina di solito mediante studi su materiali di riferimento idonei o con prove di recupero di aggiunte note. La determinazione dello scostamento sistematico globale, rispetto ad appropriati valori di riferimento, è importante per garantire la **riferibilità** a campioni riconosciuti. Lo scostamento sistematico può essere espresso come recupero analitico (valore trovato diviso il valore atteso).
- **Linearità:** la linearità è una proprietà importante dei metodi usati per eseguire misurazioni in un intervallo di concentrazioni. Può essere determinata la linearità della risposta sia per sostanze pure che per campioni realistici. La linearità generalmente non si quantifica, ma si verifica mediante controllo visivo o usando test di significatività per la non-linearità.
- **Limite di rivelabilità:** di solito durante la validazione del metodo si determina il limite di rivelabilità solo per stabilirne il limite inferiore di applicabilità. Sebbene le incertezze vicine al limite di rivelabilità richiedano un'attenta considerazione e un trattamento specifico, il limite di rivelabilità, comunque determinato, non è direttamente rilevante per la stima dell'incertezza.
- **Robustezza:** molti protocolli per lo sviluppo o la validazione di un metodo richiedono che sia studiata in modo diretto la sensibilità del metodo a variazioni di determinati parametri. Questo di solito si esegue mediante un "test di robustezza" preliminare, nel

---

<sup>16</sup> ISO 3534-1



quale si osserva l'effetto delle variazioni di uno o più parametri. Se l'effetto è significativo (paragonato alla precisione del test di robustezza) si esegue uno studio più approfondito per determinarne l'entità e, di conseguenza, scegliere l'intervallo operativo permesso. I risultati del test di robustezza possono quindi fornire informazioni sull'effetto di parametri importanti.

- **Selettività/Specificità:** sebbene definiti in modo vago, entrambi i termini si riferiscono al grado di univocità con cui un metodo risponde all'analita d'interesse. In un tipico studio di selettività si valutano gli effetti di possibili interferenti, spesso mediante l'aggiunta del potenziale interferente sia al bianco sia a campioni fortificati, e osservando l'effetto sulla risposta. I risultati sono di solito utilizzati per dimostrare che in pratica gli effetti non sono significativi. Tuttavia, poiché questi studi misurano direttamente le variazioni della risposta, è possibile utilizzarne i dati per stimare l'incertezza associata con interferenze potenziali, se è noto l'intervallo di concentrazioni dell'interferente.

### 3.5 Riferibilità

La riferibilità è definita formalmente come: “proprietà del risultato di una misurazione o del valore di un campione tale che esso possa essere collegato a riferimenti determinati, generalmente campioni nazionali o internazionali, attraverso una catena ininterrotta di confronti tutti con incertezza dichiarata”.

L'incertezza entra in gioco perché le incertezze nella catena della riferibilità di ciascun laboratorio limitano in parte la concordanza tra i laboratori. La riferibilità è dunque strettamente legata all'incertezza. La riferibilità fornisce il mezzo per collocare tutte le misurazioni dello stesso tipo su una scala di misura coerente, mentre l'incertezza caratterizza la “forza” dei legami della catena e il grado di concordanza prevedibile tra laboratori che eseguono misurazioni simili. In generale, l'incertezza di un risultato riferibile ad un dato riferimento sarà l'incertezza del riferimento insieme all'incertezza sull'esecuzione della misurazione rispetto a quel riferimento.

La riferibilità del risultato del procedimento analitico completo dovrebbe essere stabilita attraverso una combinazione dei seguenti procedimenti:

1. uso di campioni riferibili per la taratura di strumenti di misurazione;
2. uso di un metodo primario o confronto con i risultati di un metodo primario;
3. uso di una sostanza pura come materiale di riferimento (MR);
4. uso di un appropriato materiale di riferimento certificato (MRC) in matrice;

5. uso di un procedimento di misurazione accettato, strettamente definito.

### **3.5.1 Taratura degli strumenti di misurazione**

In tutti i casi, la taratura degli strumenti di misurazione usati deve essere riferibile a campioni idonei. Lo stadio di quantificazione del procedimento analitico implica spesso una taratura con una sostanza pura come MR, il cui valore è riferibile al SI. Questa prassi assicura la riferibilità dei risultati al SI per questa parte del procedimento. Tuttavia è necessario anche assicurare, con ulteriori procedure, la riferibilità dei risultati di operazioni che precedono lo stadio di quantificazione, come l'estrazione e la purificazione del campione.

### **3.5.2 Misurazioni utilizzando metodi primari**

La definizione in vigore di metodo primario è la seguente: “un metodo di misurazione primario è un metodo che possiede le più elevate qualità metrologiche, il cui svolgimento può essere completamente descritto e compreso e per il quale le incertezze sono espresse in unità SI. Un metodo primario diretto misura il valore di un'incognita senza far riferimento ad un campione della stessa grandezza. Un metodo primario di rapporto misura il valore di un rapporto tra un'incognita e un campione della stessa grandezza; lo svolgimento di questo metodo deve essere completamente descritto da un'equazione di misura”.

Il risultato di un metodo primario è di solito direttamente riferibile al SI e possiede la minima incertezza raggiungibile rispetto ad esso. I metodi primari sono di solito utilizzati solo dagli Istituti Metrologici Nazionali e sono raramente applicati a tarature e misurazioni di routine. Quando possibile la riferibilità ai risultati di un metodo primario si ottiene per confronto diretto dei risultati di misura del metodo primario e del metodo di prova o di taratura.

### **3.5.3 Misurazioni utilizzando una sostanza pura come Materiale di Riferimento (MR)**

La riferibilità può essere dimostrata mediante la misurazione di un campione costituito da, o contenente, una quantità nota di una sostanza pura con caratteristiche di MR (ad esempio mediante aggiunte note). In ogni caso è sempre necessario valutare la differenza nella risposta del sistema di misurazione ai materiali utilizzati per la taratura e ai campioni da analizzare. Purtroppo, in molte analisi chimiche e nel caso particolare delle aggiunte, sia la correzione delle differenze nella risposta, sia la sua incertezza, potrebbero essere grandi. Di conseguenza, sebbene la riferibilità del risultato alle unità del SI possa essere stabilita in linea di principio, in pratica, in tutti i casi tranne i più semplici, l'incertezza sul risultato può essere troppo

grande o addirittura non quantificabile. Se l'incertezza non è quantificabile, la riferibilità non è stata dimostrata.

#### **3.5.4 Misurazioni su un Materiale di Riferimento Certificato (MRC)**

La riferibilità può essere dimostrata mediante il confronto dei risultati di misurazione su un MRC in matrice con il valore certificato. Questo modo di operare può ridurre l'incertezza rispetto all'uso di una sostanza pura come MR, se è disponibile un MRC con una matrice idonea. Se il valore dell'MRC è riferibile al SI, allora queste misurazioni assicurano la riferibilità alle unità SI.

Anche in questo caso l'incertezza sul risultato potrebbe essere troppo grande o perfino non quantificabile, soprattutto se la composizione del campione e quella dell'MR non sono sufficientemente simili.

#### **3.5.5 Misurazioni utilizzando un procedimento accettato**

Spesso un adeguato confronto può essere eseguito solo mediante l'utilizzo di un procedimento definito nel dettaglio e generalmente accettato. Il procedimento sarà di solito definito rispetto ai parametri d'ingresso (per esempio specifici tempi di estrazione, dimensioni di particelle, ecc.). I risultati ottenuti applicando un procedimento di questo tipo sono considerati riferibili quando i valori dei parametri d'ingresso sono riferibili nel modo usuale a riferimenti dichiarati. L'incertezza sui risultati deriva sia dalle incertezze dei parametri d'ingresso specificati sia dagli effetti dovuti a una specificazione incompleta e alla variabilità. Se i risultati di un metodo o di un procedimento alternativo devono essere paragonabili ai risultati di un procedimento accettato, la riferibilità ai valori accettati si ottiene confrontando i risultati ottenuti con il procedimento accettato e con quello alternativo.

### **3.6 Il procedimento per la stima dell'incertezza di misura**

In linea di principio la stima dell'incertezza è semplice. I paragrafi seguenti riassumono le operazioni da eseguire in sequenza per ottenere una stima dell'incertezza associata con un risultato di misurazione.

Le fasi del procedimento sono:

#### ***Fase 1: Specificazione del misurando***

Definire chiaramente e in modo univoco cosa si sta misurando, comprese le relazioni tra il misurando e le grandezze d'ingresso dalle quali il misurando dipende (ad esempio grandezze misurate, costanti, valori di campioni di taratura, ecc.). Ove possibile, vanno incluse le correzioni per gli effetti sistematici noti. Queste informazioni dovrebbero essere riportate

nella relativa Procedura Operativa (PO) o altro documento che descriva il procedimento di misurazione. La distinzione tra metodi empirici e non empirici (qualche volta chiamati *razionali*) è importante perché ha delle conseguenze sulla stima dell'incertezza. Nei casi in cui vengono utilizzate delle convenzioni, le incertezze associate con alcuni effetti piuttosto grandi non sono rilevanti per l'uso corrente del dato. Di conseguenza, si dovrebbe dare dovuta considerazione alla dipendenza o meno dei risultati di misurazione dal metodo adottato e solo gli effetti rilevanti per il risultato dovrebbero essere inclusi nella stima dell'incertezza.

### ***Fase 2: Identificazione delle fonti d'incertezza***

Elencare le possibili fonti d'incertezza. L'elenco deve includere le fonti che contribuiscono all'incertezza sui parametri nella relazione specificata nella Fase 1, ma potrebbe comprenderne altre. Spesso viene utilizzato il diagramma di causa-effetto per elencare le fonti di incertezza in modo di evidenziare le reciproche relazioni e la loro influenza sull'incertezza del risultato.

Quando l'elenco delle fonti d'incertezza è stato compilato, gli effetti sul risultato possono in teoria essere rappresentati da un modello formale della misura, nel quale ogni effetto è associato con un parametro o una variabile in un'equazione. Tra le fonti comuni d'incertezza vi sono:

- **Campionamento:** se il campionamento in laboratorio o sul campo è parte della procedura specificata, effetti quali le variazioni casuali tra diverse aliquote e la possibilità di introdurre scostamenti sistematici nella procedura di campionamento, sono componenti dell'incertezza sul risultato finale.
- **Condizioni di conservazione:** se i campioni vengono conservati anche per un breve periodo di tempo prima dell'analisi, le condizioni di conservazione potrebbero influenzare i risultati. Sia la durata della conservazione sia le condizioni di conservazione dovrebbero quindi essere considerate come fonti d'incertezza.
- **Effetti strumentali:** gli effetti strumentali possono includere ad esempio i limiti di accuratezza per la taratura di una bilancia analitica, un sistema di controllo della temperatura che mantiene una temperatura media diversa (all'interno delle specifiche dichiarate) da quella impostata, un analizzatore automatico che potrebbe essere soggetto a effetti di memoria.
- **Condizioni di misurazione:** gli effetti di temperatura ed umidità potrebbero influenzare la misura.
- **Effetti del tipo di campione:** il recupero di un analita da una matrice complessa, o la risposta strumentale, potrebbero essere condizionati dalla composizione della matrice. La

stabilità di matrice/analita potrebbe cambiare durante l'analisi a causa di variazioni nel regime termico o di effetti fotolitici. Quando si usa un'aggiunta per valutare il recupero, il recupero dell'analita dal campione potrebbe differire dal recupero dell'aggiunta introducendo un'incertezza che deve essere valutata.

- **Effetti dei calcoli:** la scelta del modello di taratura, per esempio usare una retta di taratura su una risposta curvilinea, può portare a un'aderenza peggiore e a un'incertezza più elevata. Troncamenti e arrotondamenti possono portare a inaccurately sul risultato finale. Dal momento che queste sono raramente prevedibili, potrebbe essere necessario tenerne conto nella stima dell'incertezza.
- **Effetti legati all'operatore:** possibilità di leggere per esempio un indicatore o una scala sotto o sovrastimando sistematicamente. Possibilità di interpretare il metodo in modo leggermente differente.
- **Effetti casuali:** gli effetti casuali contribuiscono all'incertezza in tutte le determinazioni. Questo contributo dovrebbe essere incluso nell'elenco in ogni caso.

### *Fase 3: Quantificazione delle componenti dell'incertezza*

Misurare o stimare l'entità delle componenti dell'incertezza associate con ciascuna potenziale fonte d'incertezza identificata. Spesso è possibile stimare o determinare in un singolo contributo l'incertezza associata con un certo numero di fonti diverse. È importante anche considerare se i dati disponibili tengano adeguatamente conto di tutte le fonti d'incertezza, e pianificare attentamente studi ed esperimenti aggiuntivi per garantire che tutte le fonti d'incertezza siano state adeguatamente prese in considerazione. Il procedimento usato per stimare l'incertezza globale dipende dai dati disponibili sulle prestazioni del metodo.

Le fasi del procedimento sono:

- **Mettere a confronto i dati disponibili con quelli richiesti:** esaminando l'elenco delle fonti d'incertezza per vedere quali possano essere valutate in base ai dati disponibili, provenienti sia da uno studio esplicito di un singolo contributo, sia dalla sua variazione implicita durante lo svolgimento di esperimenti con l'intero metodo. Queste fonti dovrebbero essere spuntate dall'elenco preparato nella Fase 2 mantenendo tutte le rimanenti allo scopo di documentare quali contributi all'incertezza siano stati presi in considerazione.
- **Pianificare come ottenere gli altri dati necessari:** per le fonti d'incertezza non adeguatamente documentate da dati esistenti, si possono ottenere altre informazioni dalla letteratura scientifica o da dati altrimenti dichiarati (certificati, specifiche tecniche di strumenti, ecc.), oppure si possono pianificare esperimenti volti a ottenere gli ulteriori dati

necessari. Gli esperimenti supplementari potrebbero consistere in studi specifici di un singolo contributo all'incertezza o nei consueti studi di prestazione del metodo svolti in modo da assicurare variazioni rappresentative dei fattori importanti.

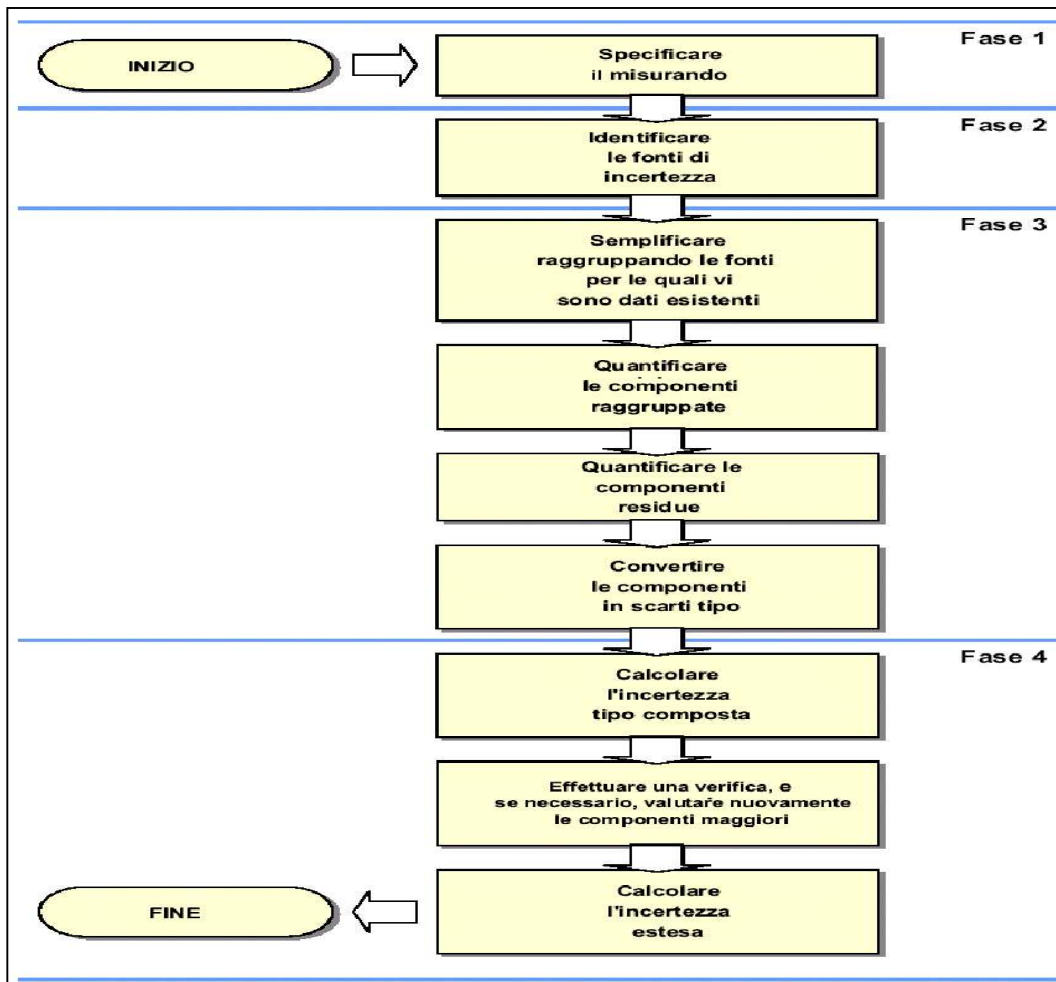
Deve essere chiaro che non tutte le componenti, anzi, in pratica, solo un piccolo numero di esse, daranno un contributo significativo. Non è necessario valutare in dettaglio le componenti che sono inferiori a un terzo di quella più grande, a meno che non siano molto numerose. Si dovrebbe fare una stima preliminare del contributo all'incertezza di ciascuna componente, o combinazione di componenti, ed eliminare quelle non significative.

**Fase 4: Calcolo dell'incertezza composta**

Le informazioni ottenute nella Fase 3 consistono in un certo numero di contributi quantificati all'incertezza globale, ciascuno associato o con una singola fonte, o con gli effetti combinati di numerose fonti. I contributi devono essere espressi come scarti tipo e combinati secondo appropriate regole matematiche per ottenere un'incertezza tipo composta.

La figura sottostante descrive schematicamente questo procedimento.

FIGURA 5: Il procedimento per la stima dell'incertezza



### 3.7 Incertezza nei procedimenti analitici

Allo scopo di identificare le possibili fonti d'incertezza in un procedimento analitico è utile suddividere l'analisi in una serie di fasi principali:

- Campionamento
- Preparazione del campione
- Scelta dei materiali di riferimento certificati in relazione al sistema di misura
- Taratura degli strumenti
- Analisi (acquisizione dei dati)
- Elaborazione dei dati
- Presentazione dei risultati
- Interpretazione dei risultati

Queste fasi possono essere ulteriormente suddivise in base ai contributi all'incertezza identificabili in ciascuna di esse. La lista seguente, sebbene non necessariamente esauriente, fornisce una guida ai fattori da considerare.

#### 1. Campionamento

- Omogeneità.
- Effetti di una strategia di campionamento specifica (casuale, casuale stratificata, proporzionale, ecc.).
- Effetti del movimento all'interno del campione (in particolar modo la distribuzione di densità).
- Stato di aggregazione del campione (solido, liquido, gas).
- Effetti di temperatura e pressione.
- Il procedimento di campionamento modifica la composizione? Ad esempio adsorbimento differenziale nel sistema di campionamento.

#### 2. Preparazione del campione

- Effetti di omogeneizzazione e/o di subcampionamento.
- Essiccamento.
- Macinazione.
- Dissoluzione.
- Estrazione.
- Derivatizzazione (effetti chimici).
- Errori di diluizione.
- Contaminazione.

- (Pre) Concentrazione.
- Controllo degli effetti di speciazione.

### **3. Scelta dei materiali di riferimento certificati in relazione al sistema di misura**

- Incertezza per l'MRC.
- Somiglianza di MRC al campione.

### **4. Taratura della strumentazione**

- Errori nella taratura degli strumenti con MRC.
- Materiale di riferimento e sua incertezza.
- Somiglianza del campione al materiale per la taratura.
- Precisione dello strumento.

### **5. Analisi**

- Effetti di memoria negli analizzatori automatici.
- Effetti dovuti all'operatore (daltonismo, parallasse, altri errori sistematici).
- Interferenze di matrice, reagenti ed altri analiti.
- Purezza dei reagenti.
- Impostazione dei parametri dello strumento, ad esempio parametri d'integrazione.
- Precisione tra prove analitiche ripetute.

### **6. Elaborazione dei dati**

- Calcolo della media.
- Controllo dell'arrotondamento e del troncamento.
- Elaborazioni statistiche.
- Algoritmi per l'elaborazione (adeguatezza dei modelli di calcolo, per esempio metodo dei minimi quadrati).

### **7. Presentazione dei risultati**

- Risultato finale.
- Stima dell'incertezza.
- Livello di fiducia.

### **8. Interpretazione dei risultati**

- In confronto a limiti/fasce.
- Per la conformità a requisiti normativi.
- Per l'idoneità allo scopo.



---

# Capitolo 4

## Analisi statistica dati

### 4.1 Premessa

Di seguito viene riportata la descrizione dell'analisi statistica effettuata su dei dati che rappresentano delle caratteristiche chimico/fisiche rilevate in laboratorio d'analisi presso lo stabilimento Montefibre di P.to Marghera.

Le caratteristiche interessate nell'analisi sono: Titolo, Tenacità, Allungamento ed Estraiabili<sup>17</sup>. I valori riportati riguardano solo un particolare tipo di produzione: fibra acrilica tinta in "pasta" (TP).

Il dataset in studio è stato aggiornato pari passo l'evolversi della tesi in modo da considerare due versioni finali: la prima effettuata nel mese di Gennaio e la seconda svolta nel mese di Aprile con una numerosità più elevata dei record (campioni). L'aggiornamento del file ha evidenziato delle variazioni nell'analisi conclusiva rispetto alla prima versione. Nell'elaborato si trovano tabelle riportanti i valori dei test applicati per verifiche d'ipotesi sia con i dati iniziali che con i dati aggiornati.

Tutti i test statistici applicati nell'analisi seguente, sono stati considerati con un livello di significatività (regione di accettazione o rifiuto) del test d'ipotesi basato su un valore del p-value<sup>18</sup> pari al 5%.

Nell'analisi preliminare e nelle considerazioni iniziali viene data particolare enfasi alla non normalità dei dati. Tuttavia il risultato dei grafici (istogrammi) ed i test applicati ai dati aggiornati, convalidano la scelta dell'ipotesi di normalità in distribuzione.

Tale ipotesi trova riscontro anche per la natura stessa dei dati d'origine, nonché per i riferimenti al BISFA ed alla normativa che regola la validazione dei metodi d'analisi di laboratorio ISO 5725 (ISO, 1994)<sup>19</sup>. I valori vengono rilevati secondo un piano di campionamento prefissato e si riferiscono a produzioni di articoli diversi (per colore), riguardando un periodo pari a circa dieci mesi di produzione delle due macchine di filatura (da Luglio 2005 ad Aprile 2006).

---

<sup>17</sup> Riferimenti approfonditi nel capitolo 2 paragrafo 2.5.

<sup>18</sup> Vedi capitolo successivo.

<sup>19</sup> Vedi capitolo 3.

L'analisi concentra l'attenzione sui seguenti punti (tutti corredati da grafici e test statistici):

1. distribuzione dei dati
2. correlazioni evidenti
3. raggruppamenti e considerazioni

ed è finalizzata sulla base dei dati disponibili, allo studio di possibili fonti di variabilità nel processo di produzione che concorrono in parte alla variabilità complessiva del dato finale. Tale variabilità viene considerata nella proposta del modello conclusivo.

Le analisi statistiche sono state effettuate con l'uso del software "R"<sup>20</sup>.

## 4.2 Analisi preliminare dei dati

Il file d'origine in studio contiene 165 osservazioni di quattro caratteristiche d'interesse qualitativo su due macchine di produzione di fibra acrilica numerate come Mc 9 e 19.

Nel file viene inoltre riportato (indicato come "PART") l'identificativo numerico del contenitore su cui viene messa la fibra che ne consente la rintracciabilità temporanea, il codice della tinta (colore) ed i valori rilevati in laboratorio delle quattro caratteristiche.

L'aggiornamento ai dati soprascritti aggiunge altri valori (record) ottenendo un dataset finale di 209 dati per le quattro caratteristiche d'interesse.

La prima sintesi dei dati riporta la dimensione numerica, i valori di minimo, massimo, primo e terzo quartile, mediana e media del vettore riferito ad ognuna delle caratteristiche rilevate.

TABELLE RIASSUNTIVE : TAB.1 (DATI DI PARTENZA), TAB. 2 (DATI AGGIORNATI)

**TAB.1**

	<b>dim</b>	<b>min</b>	<b>1°quartile</b>	<b>mediana</b>	<b>media</b>	<b>3°quartile</b>	<b>max</b>	<b>varianza</b>
<b>Titolo</b>	165	2,03	2,2	2,24	2,24	2,28	2,5	0,00634
<b>Tenacità</b>	165	32,8	38,2	39,4	39,72	41,2	46,7	5,683
<b>Allungamento</b>	165	25,6	28,6	29,5	29,39	30,3	34,6	2,416
<b>Estraibili</b>	165	0,41	0,59	0,63	0,641	0,7	0,95	0,00834

**TAB.2**

	<b>dim</b>	<b>min</b>	<b>1°quartile</b>	<b>mediana</b>	<b>media</b>	<b>3°quartile</b>	<b>max</b>	<b>varianza</b>
<b>Titolo</b>	209	2,03	2,21	2,24	2,25	2,3	3,39	0,013
<b>Tenacità</b>	209	31,4	38,1	39,4	39,51	41,1	46,7	5,988
<b>Allungamento</b>	209	20,2	27,8	28,9	28,92	30	34,6	3,178
<b>Estraibili</b>	209	0,41	0,6	0,64	0,65	0,71	0,95	0,008

<sup>20</sup> Programma per analisi statistica dei dati, open source. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: <http://www.r-project.org>.

### 4.3 Analisi della distribuzione dei dati

Di seguito viene riportata la rappresentazione grafica<sup>21</sup> dei dati attraverso istogrammi con sovrapposizione della curva normale, boxplot e grafico di dispersione per ognuna delle quattro caratteristiche, inoltre la tabella sottostante riporta i valori di simmetria e curtosi<sup>22</sup> dei dati aggiornati e non.

Dati di partenza	simmetria = 0; < 0= asim.neg.; >0= asim.positiva	curtosi(code +/-pesanti)
titolo	0,39	0,52
tenacità	-0,1	0,625
allungamento	-0,01	0,34
estraibili	0,47	1,12

Dati aggiornati	simmetria = 0; < 0= asim.neg.; >0= asim.positiva	curtosi(code +/-pesanti)
titolo	4,78 / 0,35*	44,93 / 0,35*
tenacità	-0,148	0,688
allungamento	-0,268 / 0,275*	2,3 / 0,234*
estraibili	0,173	0,296

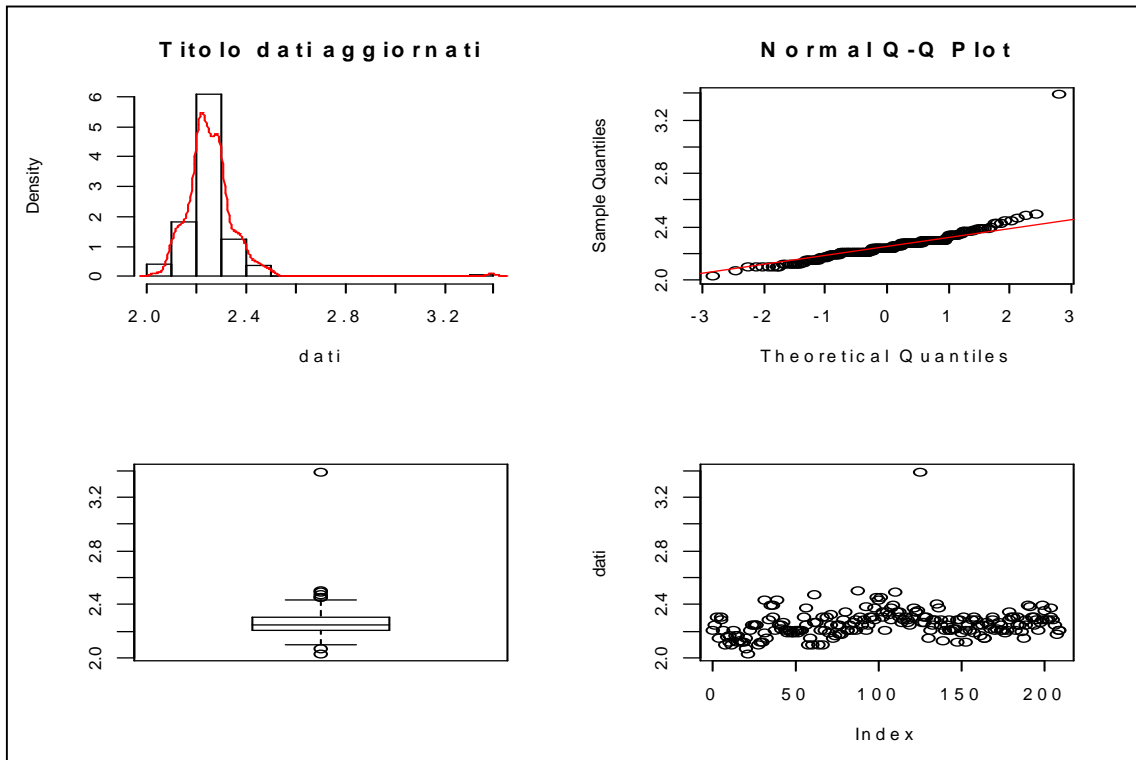
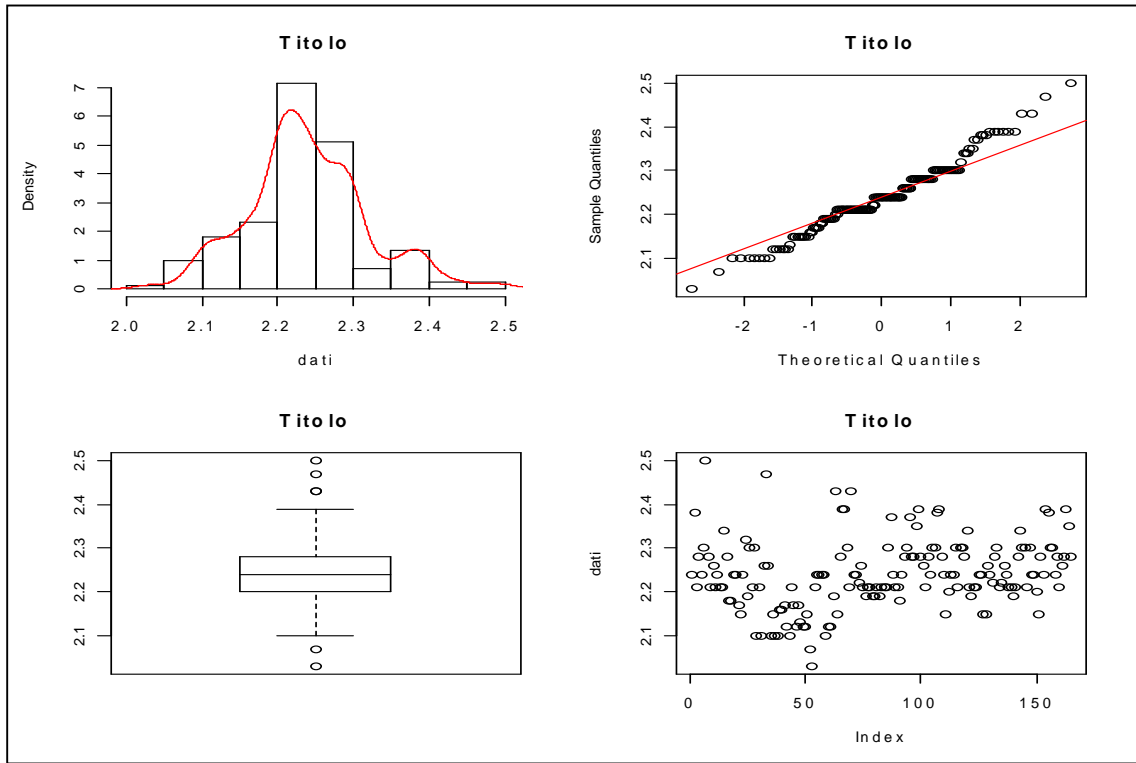
\* valore riportato con esclusione del dato anomalo (vedi plot).

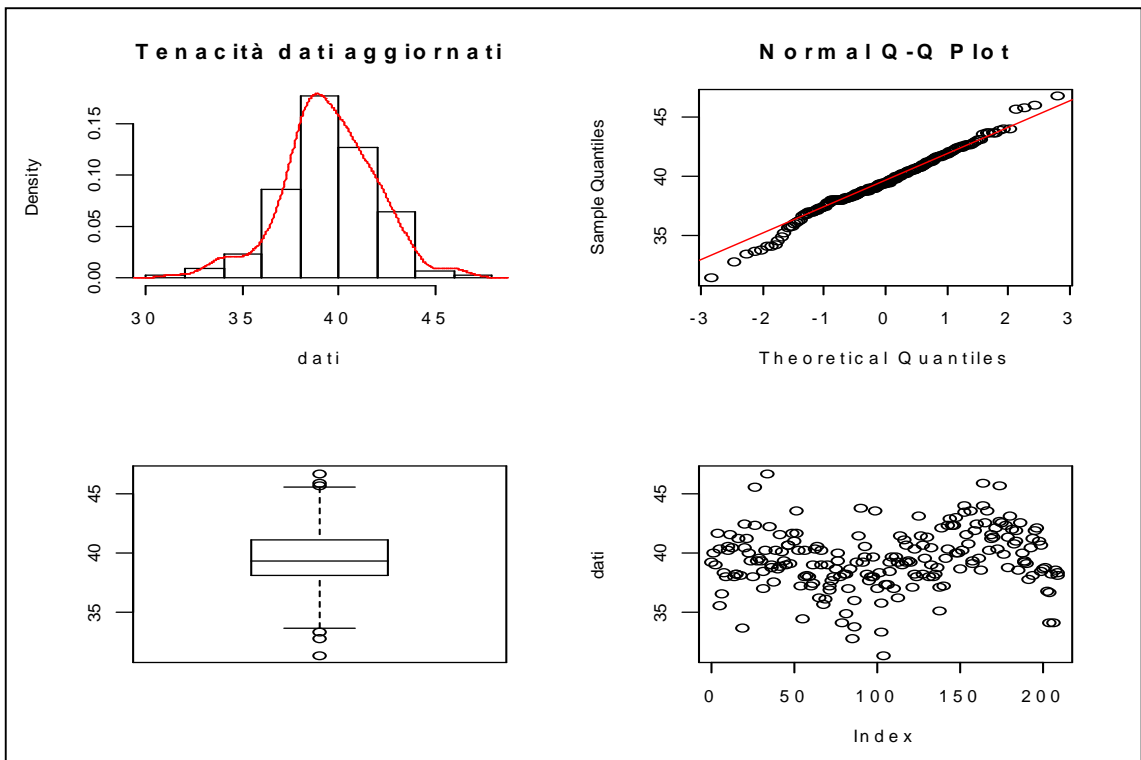
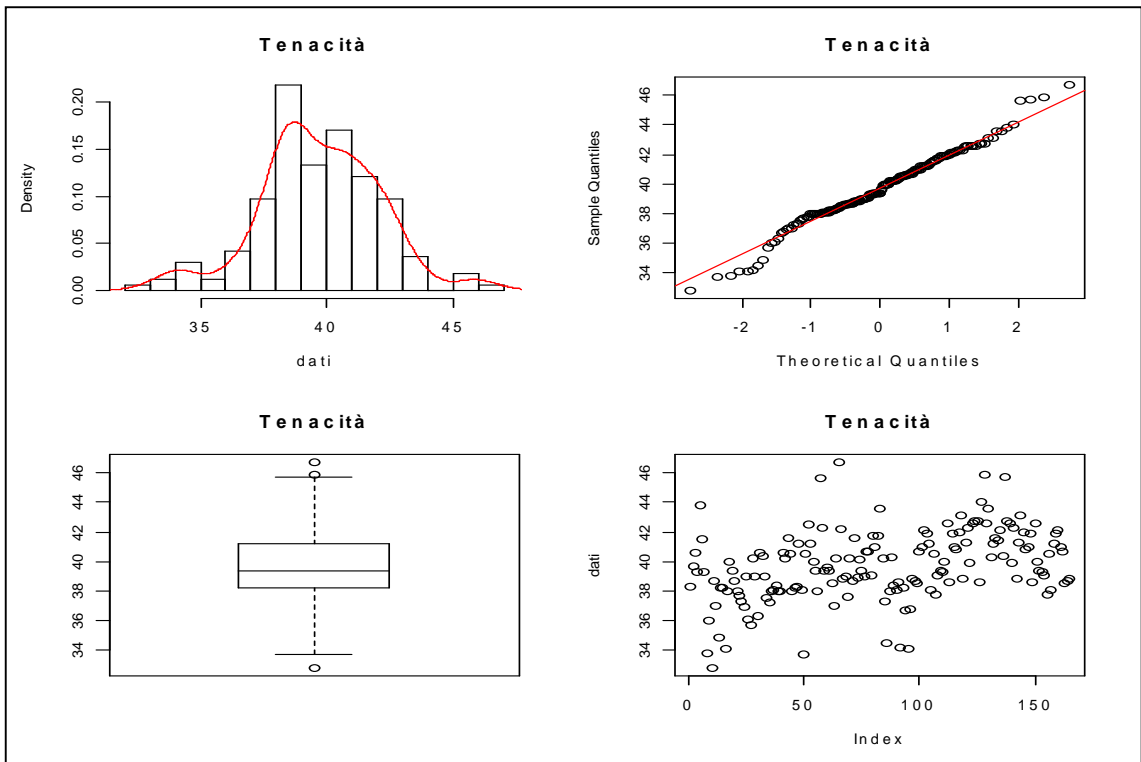
I primi grafici evidenziano la singola caratteristica per i dati di origine mentre nel successivo riquadro vi sono le caratteristiche aggiornate.

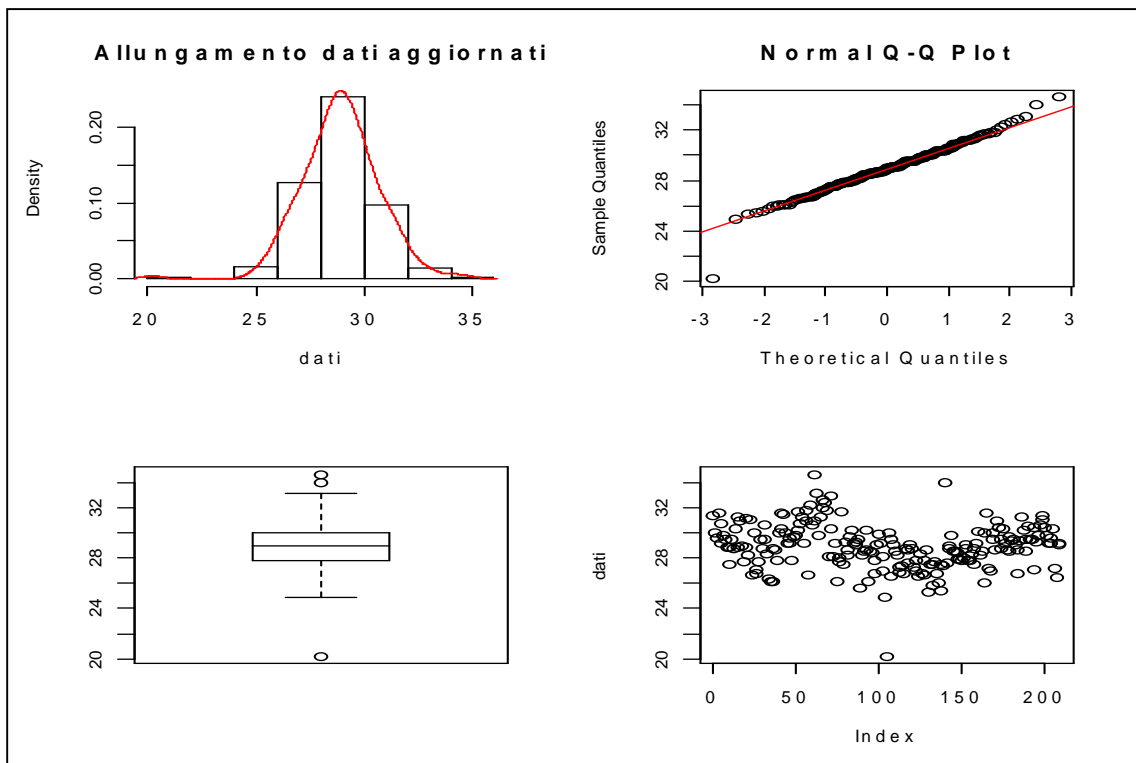
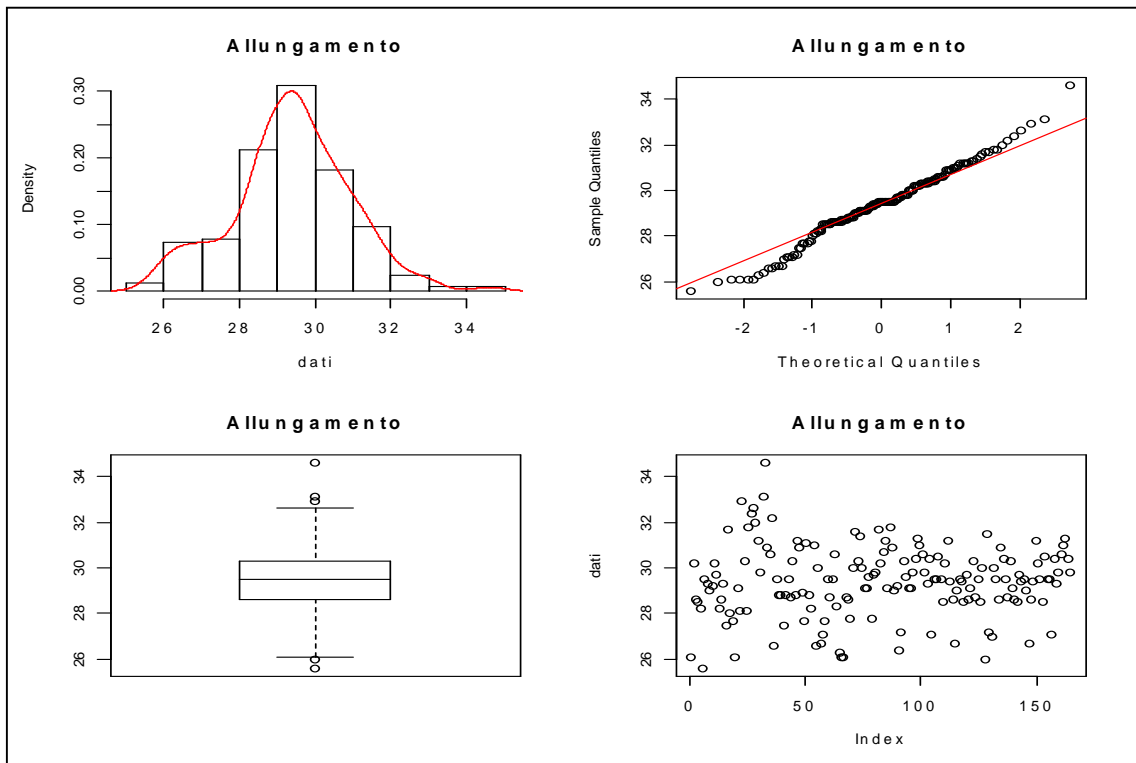
Dall'analisi visiva dei grafici si nota la presenza di valori anomali per la caratteristica titolo ed allungamento nel dataset aggiornato.

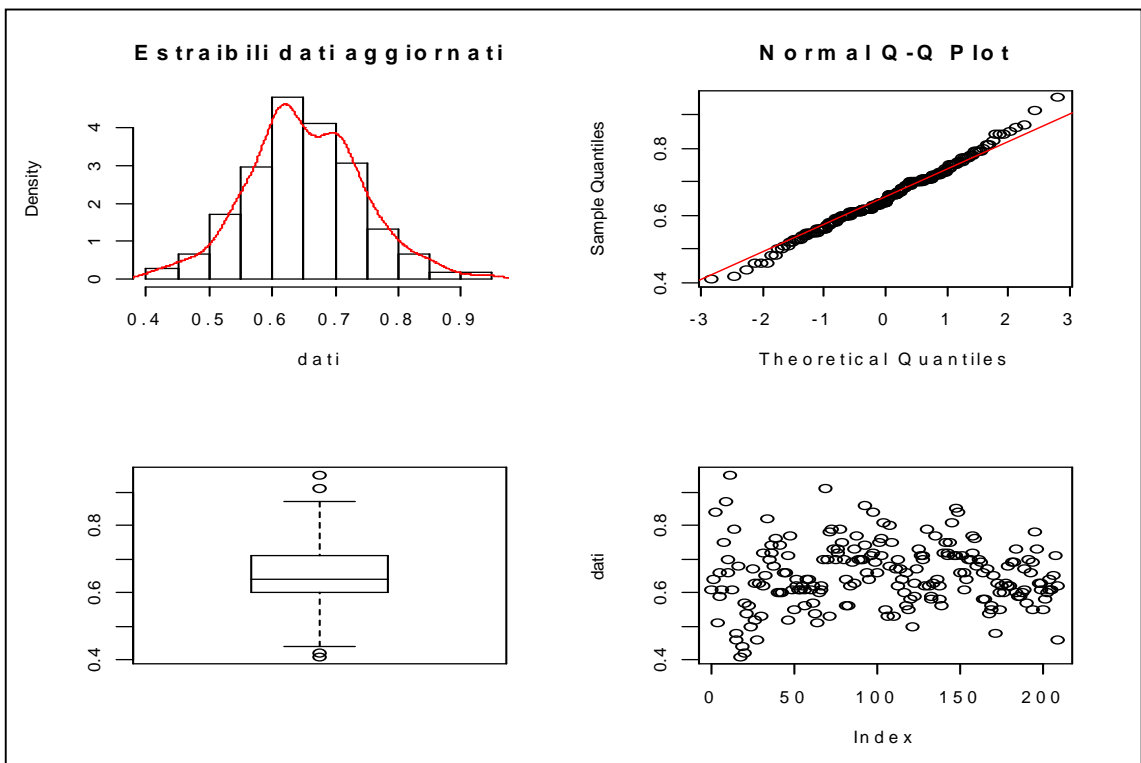
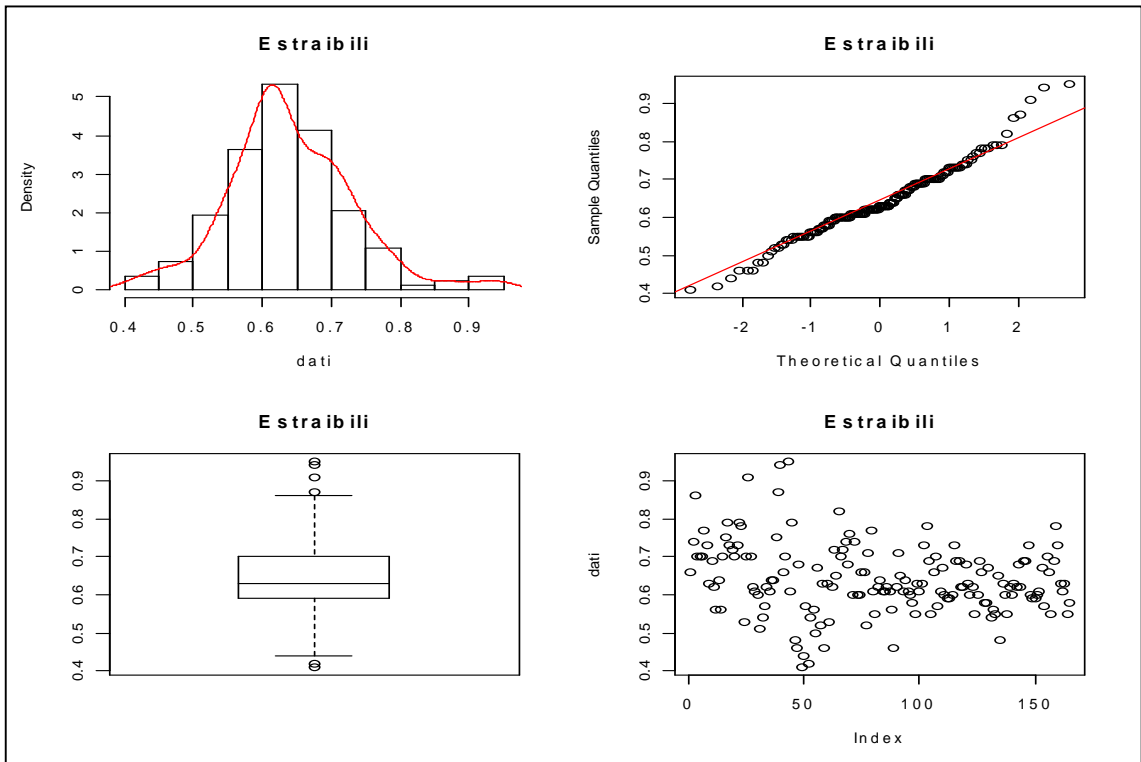
<sup>21</sup> Funzione in R riportata in Appendice E 1.

<sup>22</sup> Per calcolare le misure di asimmetria e curtosi di una distribuzione si è ricorso alle funzioni *skewness()* e *kurtosis()* contenute nel package *fBasics* di "R".









Dall'analisi effettuata con il Shapiro test<sup>23</sup> per i dati di origine “dati 1” si denota che contrariamente a quanto visto nell'analisi grafica iniziale, la distribuzione delle caratteristiche osservate (titolo, tenacità, estraibili) non è normale; solo per la caratteristica allungamento il test dà esito positivo.

Per i dati aggiornati invece, il test ci porta ad accettare la normalità in distribuzione per tenacità (anche se al limite) ed estraibili, mentre si abbassa il p-value per la caratteristica titolo ed allungamento. Da notare come cambia il valore del test eliminando i record anomali; il p-value è decisamente elevato rispetto ai livelli di significatività a cui si fa riferimento e ciò ci fa propendere per l'ipotesi nulla (normalità distribuzione) in particolare per la caratteristica allungamento.

<b>Shapiro.test per verifica normalità p-value</b>	<b>dati 1</b>	<b>dati 2</b>
<b>Titolo</b>	0,0031	0 / 0,003*
<b>Tenacità</b>	0,0336	0,07
<b>Allungamento</b>	0,1	0,0013 / 0,37*
<b>Estraibili</b>	0,0023	0,412

\* valori con esclusione del dato anomalo.

Vista l'importanza che ha l'assunzione di normalità dei dati nell'elaborazione statistica e per le finalità richieste nel presente elaborato, è stata svolta l'analisi dei dati con trasformazione degli stessi. Sono stati considerati inoltre altri test statistici per l'analisi della normalità. Il risultato ottenuto con le opportune trasformate tuttavia, non porta ad una miglior situazione<sup>24</sup>. A seguito della natura dei dati e per analogia con i riferimenti in materia (BISFA e normativa ISO 5725) si propende all'accettazione delle condizioni di normalità per la distribuzione dei dati. L'assunzione di normalità dei dati viene seguita anche dall'analisi svolta con la suddivisione delle due macchine di produzione.

Di seguito vengono riportati i valori dei p-value ottenuti con il shapiro test per le quattro caratteristiche osservate singolarmente nelle due macchine.

<b>Norm (shapiro test p-value)</b>	<b>dati di partenza</b>		<b>dati aggiornati</b>	
	<b>9</b>	<b>19</b>	<b>9</b>	<b>19</b>
<b>Titolo</b>	0,0012	0,005	0,0017	< 2.2e-16/0,021*
<b>Tenacità</b>	0,046	0,06655	0,07	0,152
<b>Allungamento</b>	0,731	0,0015	0,79	5.213e-05
<b>Estraibili</b>	0,3944	0,1461	0,66	0,55

\* valore ottenuto nel test eliminando l'osservazione anomala.

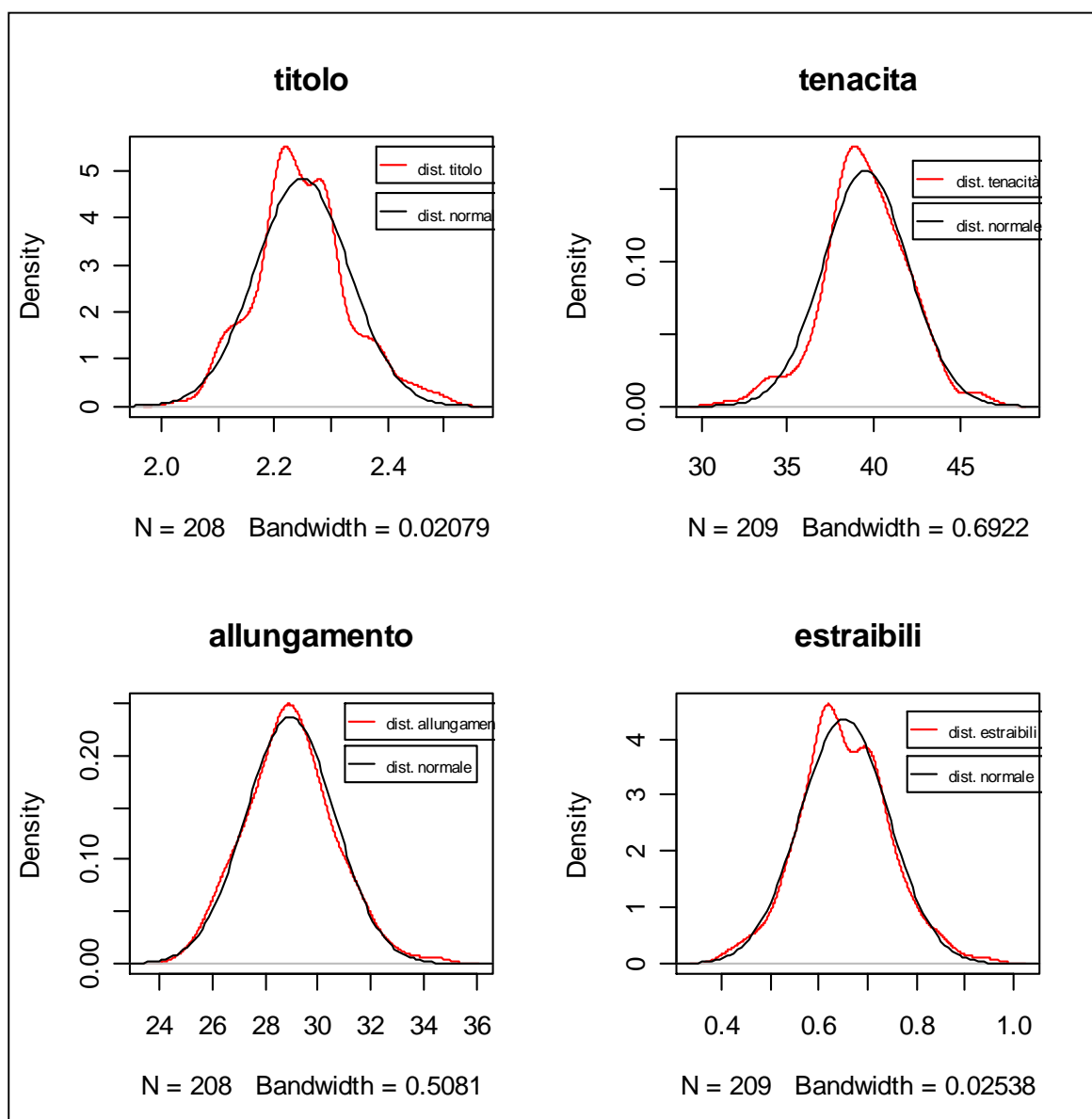
<sup>23</sup> Appendice B.

<sup>24</sup> Valori riportati in Appendice C.



Un aspetto importante per la scelta del modello normale nella distribuzione dei dati, viene affiancato osservando la distribuzione grafica delle caratteristiche (vedi sotto).

Tolta l'osservazione anomala per la caratteristica titolo ed allungamento, non si può trascurare il fatto che tendenzialmente vi sia presenza di una distribuzione normale nell'adattamento del modello. Nei grafici vengono riportate le distribuzioni delle caratteristiche (di colore rosso) con sovrapposizione della curva normale a parità di parametri.



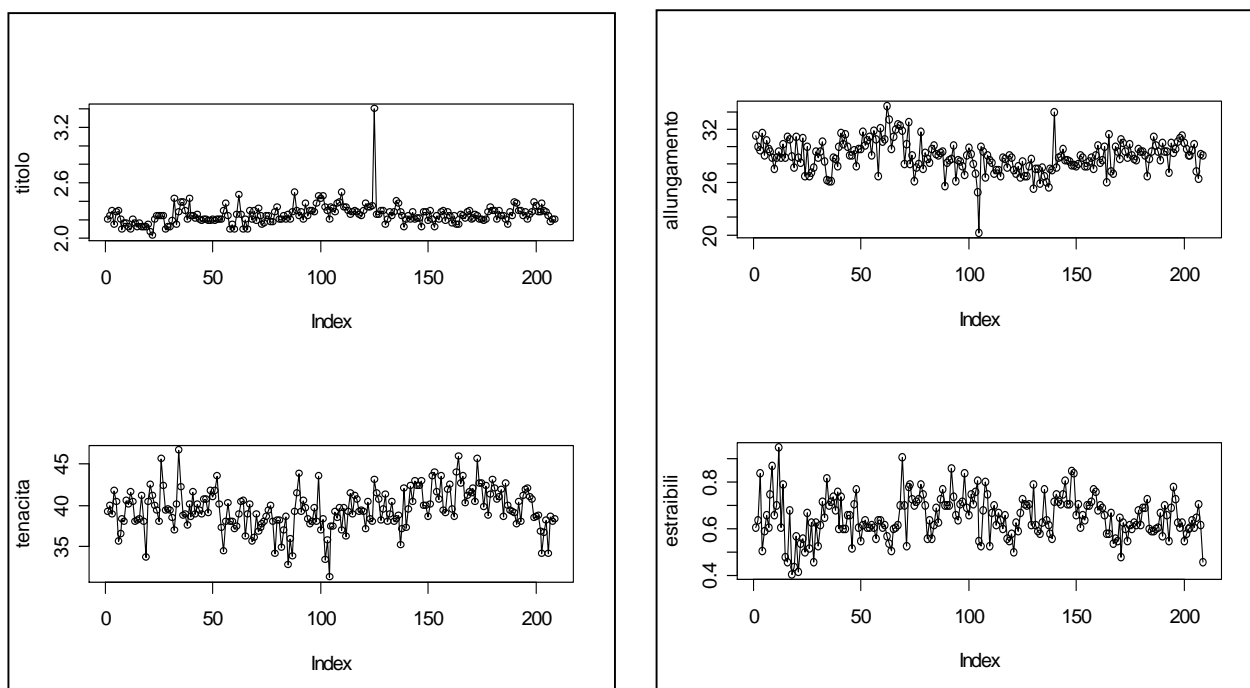
### 4.4 Analisi della correlazione

Di seguito vengono riportati i valori del coefficiente di correlazione per le quattro caratteristiche .

Dati di partenza				
	titolo	tenacita	allungamento	estraibili
titolo	1	-0,09077	0,09874	0,14159
tenacita	-0,09077	1	-0,09019964	-0,04354
allungamento	0,09874	-0,0902	1	-0,09461
estraibili	0,14159	-0,04354	-0,09461061	1

Dati aggiornati				
	titolo	tenacita	allungamento	estraibili
titolo	1	-0,05708	-0,06966984	0,162941
tenacita	-0,05708	1	0,005428953	-0,01782
allungamento	-0,06967	0,005429	1	-0,1238
estraibili	0,162941	-0,01782	-0,123797	1

Dall’analisi dei coefficienti di correlazione non si evince nessun tipo di dipendenza lineare in media tra le caratteristiche di interesse. Lo stesso si può notare anche dai grafici storici riportati sotto. L’andamento non evidenzia nessuna forma di legame tra i valori dei vettori del dataset.

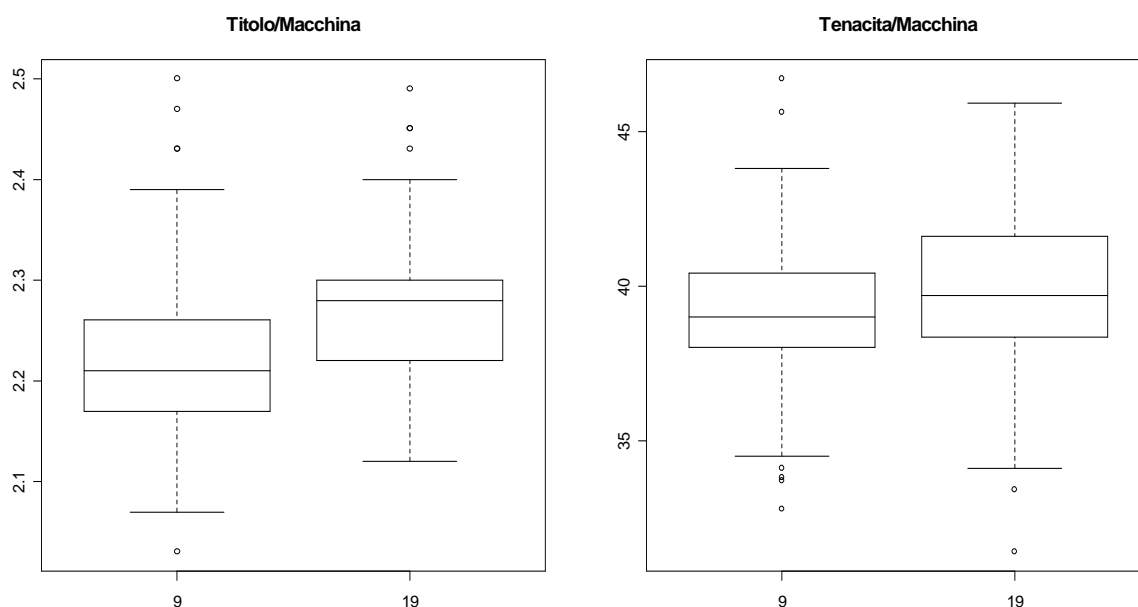


## 4.5 Analisi della media e della varianza

### 4.5.1 Suddivisione tra le due macchine di produzione

La prima analisi svolta, concentra l'attenzione alla possibile fonte di variabilità dei dati dovuta alla diversa macchina di produzione. Il dataset viene analizzato per vedere se tra le due macchine di produzione vi è una differenza in media ed in varianza delle caratteristiche. In sottolineato vengono riportati i valori dei p-value ottenuti nel test.

La verifica di ipotesi sulla media della popolazione da cui proviene un campione è effettuato tramite il "t.test" mentre per la varianza viene usato il "var.test"<sup>25</sup> ed è riferita per i soli dati aggiornati.

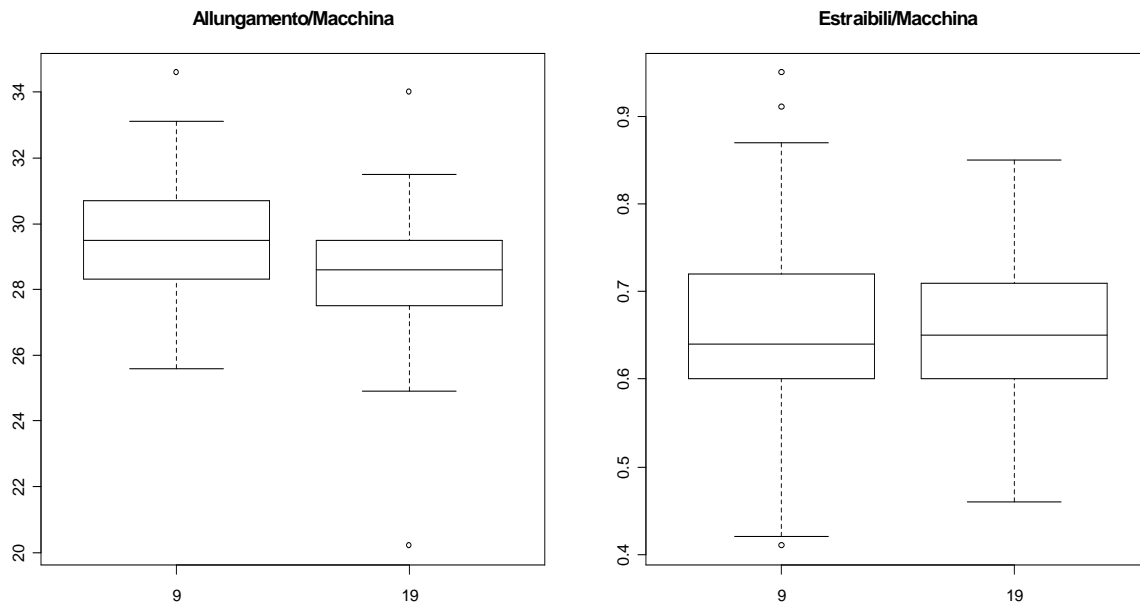


#### Analisi Titolo/Mc e Tenacità/Mc

→ t.test: p-value = 1.372e-05, var.test: p-value = 0.035. Il test ci porta ad accettare l'ipotesi che la media e la varianza non siano uguali ovvero i due gruppi hanno media e varianza diverse. NB: nel vettore considerato per la caratteristica titolo è stato tolto il valore anomalo.

→ t.test: p-value = 0.01870, var.test: p-value = 0.6722. Anche per la caratteristica Tenacità il test ci porta a rifiutare l'ipotesi di uguaglianza della media, mentre il test di omogeneità della varianza per i dati considerati ci porta ad accettare l'ipotesi di uguaglianza nei due gruppi della stessa.

<sup>25</sup> Le corrispondenti statistiche sono riportate in Appendice D.



### Analisi Allungamento/Mc e Estraiibili/Mc

→ t.test: p-value = 2.646e-05, var.test: p-value = 0.4485. Il test per quanto riguarda la caratteristica allungamento ci induce a considerare che la media dei due campioni non sia uguale. Per l'analisi della varianza il test (var.test) ci porta ad accettare l'ipotesi di omoschedasticità.

→ t.test: p-value = 0.6514, var.test: p-value = 0.003067. Il test ci dà un valore del p-value di gran lunga > al 5% il risultato, quindi, ci porta ad accettare l'ipotesi di uguaglianza delle medie; cosa evidente anche dai boxplot sopra.

Il test di uguaglianza della varianza nelle due macchine di produzione ci porta a non accettare l'ipotesi di omogeneità nei due gruppi.

Di seguito viene riportata la tabella riassuntiva con i dati aggiornati e non aggiornati del primo raggruppamento. Inoltre si possono notare i valori dei test applicati alle quattro caratteristiche di interesse soprascritti.

### Tabella riassuntiva prima analisi

<b>Dati non aggiornati</b>		<b>Titolo</b>	<b>Tenacità</b>	<b>Allungamento</b>	<b>Estraibili</b>
<b>MC</b>	<b>Numerosità</b>	<b>med/sd</b>	<b>med/sd</b>	<b>med/sd</b>	<b>med/sd</b>
<b>9</b>	88	2,22/0,00795	39,08/5,665	29,41/3,274	0,654/0,0121
<b>19</b>	77	2,262/0,00361	40,45/4,76	29,36/1,465	0,626/0,0037
<b>Dati</b>	165				
<b>Media</b>		2.24	39.72	29,39	0,641
<b>Sd</b>		0.00634	5.683	2,4166	0,00834
<b>var.test(p-value)</b>		0,00052	0,439	0,00042	0
<b>t.test(p-value)</b>		0,000381	0,000174	0,824	0,0475
<b>Wilcox.test(p-value)</b>		0	0	0,86	0,043

<b>Dati aggiornati</b>		<b>Titolo</b>	<b>Tenacità</b>	<b>Allungamento</b>	<b>Estraibili</b>
<b>MC</b>	<b>Numerosità</b>	<b>med/sd</b>	<b>med/sd</b>	<b>med/sd</b>	<b>med/sd</b>
<b>9</b>	94	2,22/0,087	39,08/2,36	29,48/1,78	0,65/0,105
<b>19</b>	115	2,28/0,126	39,87/2,466	28,45/1,65	0,65/0,0787
<b>Dati</b>	209				
<b>Media</b>		2.25	39,51	28,91	0,65
<b>Sd</b>		0,113	2,447	1,783	0,0916
<b>var.test(p-value)</b>		0.000316	0.6722	0.4485	0.003067
<b>t.test(p-value)</b>		7,59E-05	0.01870	2,65E-05	0.6514

### Considerazioni

Per la caratteristica Titolo si nota come tra le due macchine vi sia una diversità sia in media che in varianza. Tenacità e Allungamento denotano una variabilità in media ma non in varianza, ovvero nonostante la media per entrambe le caratteristiche sia diversa nelle due macchine di produzione, i test ci inducono a considerare che la varianza sia la stessa. Al contrario si può dire per la caratteristica Estraibili, ovvero i test ci portano a considerare stessa media per le due macchine ma varianza diversa.

I risultati dei test applicati quindi evidenziano delle diversità per le caratteristiche considerate nelle due macchine di filatura. La causa di tali differenze può essere dovuta alle diversità di marcia operativa nonché strutturale delle due macchine di produzione.

### 4.5.2 Analisi dei dati raggruppati per colore

Ciò che si vuole valutare con l'analisi seguente è la presenza o meno di una relazione tra il colore del prodotto e le caratteristiche in esame. A tal proposito vengono riportati i risultati dei test basati su due diversi raggruppamenti. Il primo prevede l'utilizzo dei dati non aggiornati, quindi con una bassa numerosità, mentre il secondo considera i dati aggiornati ma con un criterio di raggruppamento diverso. In questo modo vengono confrontati i risultati dei

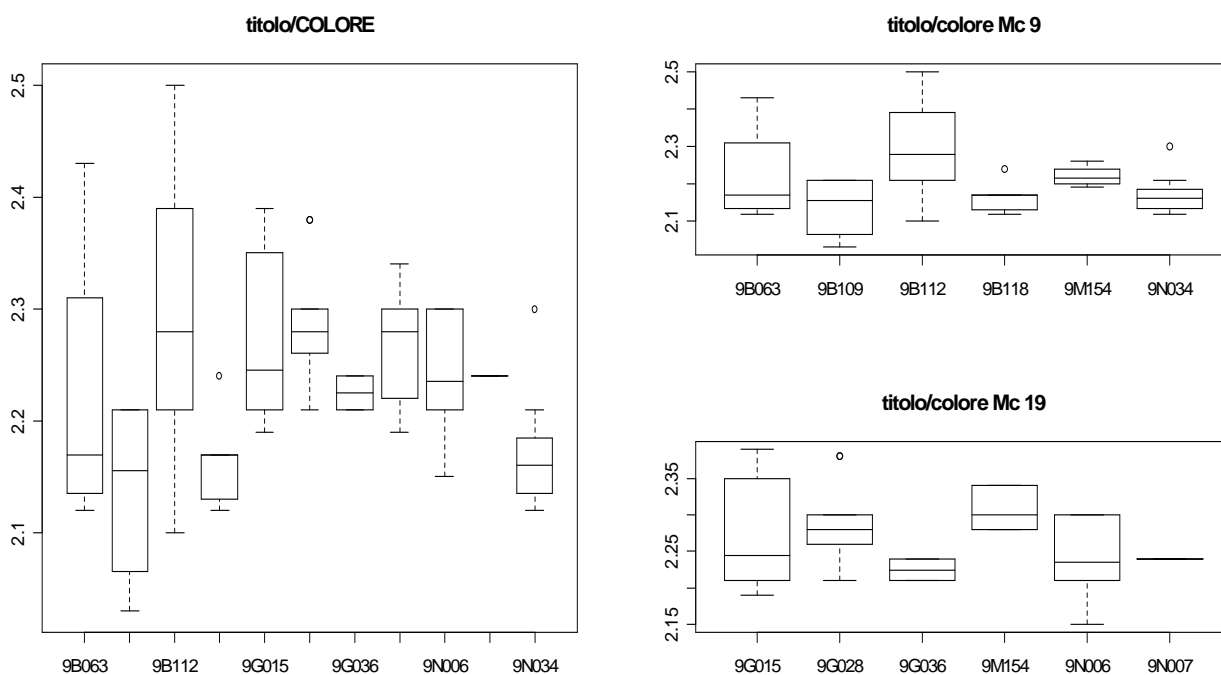
test sui due raggruppamenti ed è perciò possibile fare delle considerazioni più specifiche nell'eventuale dipendenza delle caratteristiche con il colore.

#### 4.5.2.1 Primo raggruppamento, versione dei dati non aggiornata.

Dei dati presenti nel dataset non aggiornato, sono stati considerati e raggruppati i record con una numerosità  $\geq 4$  per ogni colore. Sono stati ottenuti undici gruppi di colore diverso<sup>26</sup> per un totale di 77 osservazioni, su cui viene fatta l'analisi della varianza e della media.

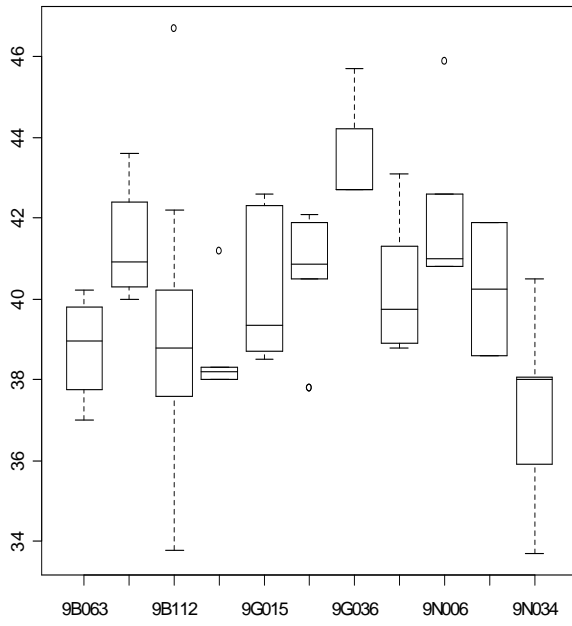
L'analisi riguarda dapprima tutto il gruppo di osservazioni senza tener conto della macchina di filatura, poi è stata eseguita la stessa operazione tenendo conto delle due macchine diverse di produzione.

Di seguito vengono riportati i boxplot dei dati suddivisi per colore (lato sinistro) e poi suddivisi oltre che per colore anche per macchina di filatura (lato destro).

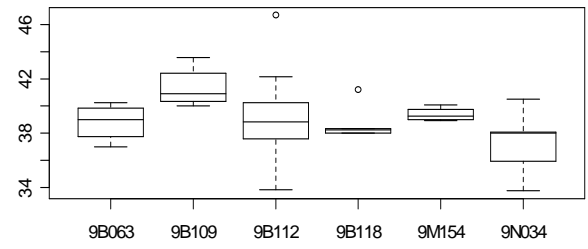


<sup>26</sup> Nel raggruppamento che tiene conto della diversità di macchina, vi è un colore comune, quindi nella rappresentazione complessiva i gruppi sono solo undici.

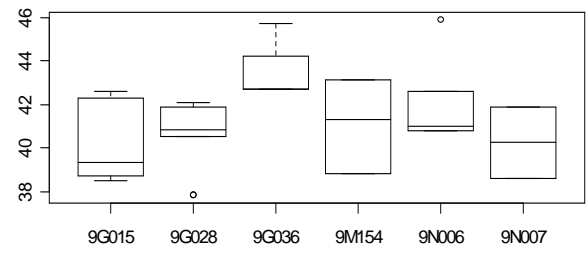
**tenacità/colore**



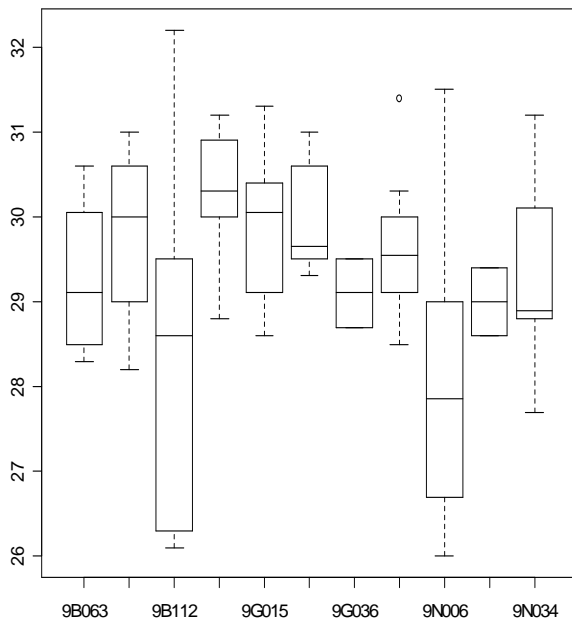
**tenacità/colore Mc 9**



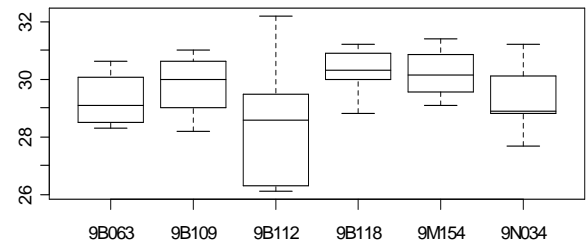
**tenacità/colore Mc 19**



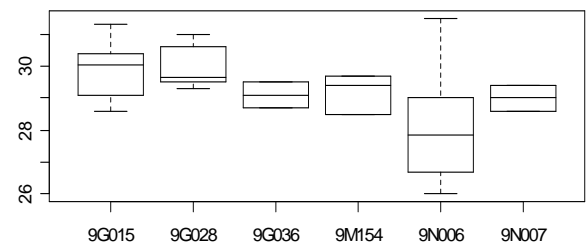
**allungamento/colore**

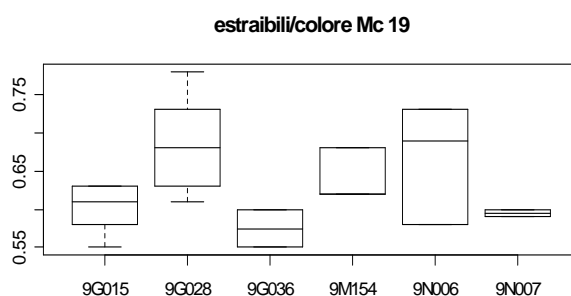
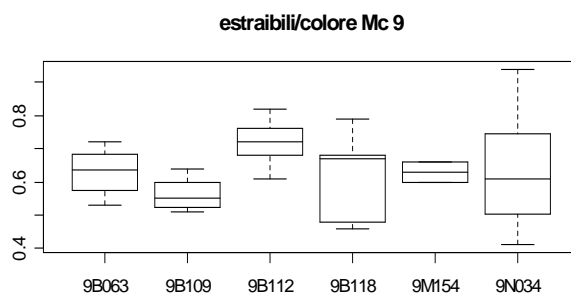
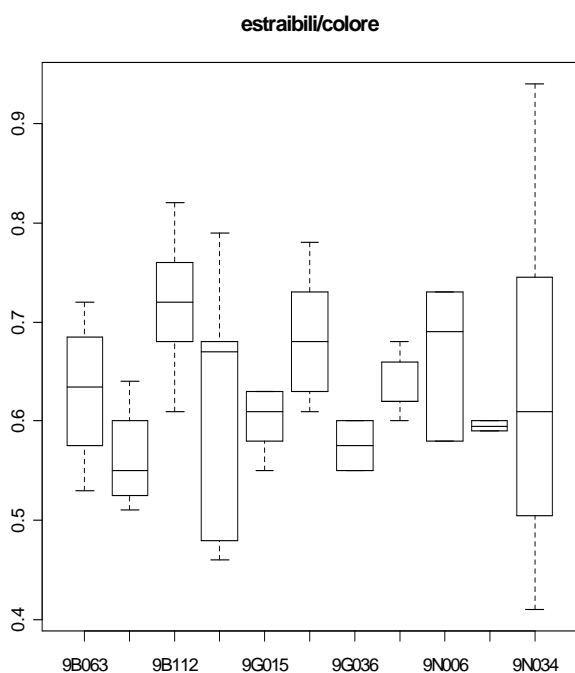


**allungamento/colore Mc 9**



**allungamento/colore Mc 19**





**Sintesi dei risultati e misure di rilevanza**

Dati ragg. per colore		<b>Titolo</b>	<b>Tenacità</b>	<b>Allungamento</b>	<b>Estraibili</b>
MC	Dim.	med/var	med/var	med/var	med/var
<b>9</b>	33	2,215/0,012	38,95/6,416	28,406/2,319	0,644/0,0145
<b>19</b>	44	2,268/0,0035	2,268/3,676	29,43/1,3	0,635/0,00356
<b>Dati</b>	<b>77</b>				
<b>Media</b>		2,246	40,1	29,42	0,639
<b>Var</b>		0,0077	6	1,712	0,0081
<b>Norm(Shapiro)</b>	p-value	0,183	0,051	0,00477	0,04433

**Considerazioni**

Presi singolarmente i dati (vettori) possono considerarsi normali per titolo e tenacità mentre al contrario si può dire per allungamento ed estraibili. Se consideriamo la separazione dei dati per le 2 macchine di produzione i risultati portano per quasi tutte le caratteristiche alla non normalità (esclusa Mc 9 per allungamento ed estraibili).

Norm (shapiro test p-value)	raggruppamento per colore	
	<b>9</b>	<b>19</b>
<b>Titolo</b>	0.00754	0.03506
<b>Tenacità</b>	0,0389	0,0257
<b>Allungamento</b>	0,2304	0,006186
<b>Estraibili</b>	0,9691	0,00528



Visti i risultati ottenuti precedentemente, per quanto riguarda la trasformazione di dati, continuiamo le opportune verifiche con l'uso di test parametrici e non dove richiesto.

La tabella di seguito riporta i valori dei p-value dei vari test<sup>27</sup> applicati alle caratteristiche di interesse per il vettore completo (senza tener conto della macchina di produzione).

	<b>titolo</b>	<b>tenacita</b>	<b>allungamento</b>	<b>estraibili</b>
<b>aov(p-value)</b>	0,003739	0,004561		
<b>bartlett.test</b>	0	0,2642		
<b>kruskal.test(p-value)</b>			0,1301	0,002117
<b>fligner.test(p-value)</b>			0,03154	0,03852

I risultati riportati in tabella denotano una differenza in media ed in varianza per la caratteristica titolo ed estraibili, mentre per la caratteristica tenacità, i test riportano una differenza in media ma non in varianza ed infine si può dire che per la caratteristica allungamento la varianza risulta statisticamente significativa e la media uguale.

Visti i risultati relativi alla diversità dei valori per le due macchine di produzione e per completezza, viene riportata l'analisi separata ritenendo di non conoscere la distribuzione dei dati. La tabella sottostante riporta i valori dei p-value dei test di uguaglianza della media e della varianza per ogni caratteristica separatamente per ogni macchina.

<b>kruskal/fligner(p-value)</b>	<b>titolo</b>	<b>tenacita</b>	<b>allungamento</b>	<b>estraibili</b>
<b>9</b>	0,1644/0,1494	0,08537/0,3201	0,3601/0,2651	0,1573/0,3573
<b>19</b>	0,1056/0,003918	0,03496/0,6633	0,02691/0,0134	0,001195/0,0629

### Conclusione prima analisi

I risultati ottenuti dalle analisi ci portano ad accettare uguale media e varianza per titolo, allungamento ed estraibili nella mc.9. Per la caratteristica tenacità invece si distingue l'uguaglianza in varianza ma non in media.

Nella Mc 19, per la caratteristica allungamento ed estraibili, non si ha l'uguaglianza in media ed in varianza, mentre per titolo si ha la stessa media ma non la varianza e viceversa per la caratteristica tenacità.

In conclusione, sulla base delle considerazioni fatte, si può dire che generalmente anche il colore incide nella variabilità dei dati, in particolar modo per la Macchina 19.

<sup>27</sup> Appendice D.

### 4.5.2.2 Secondo raggruppamento, versione dati aggiornati

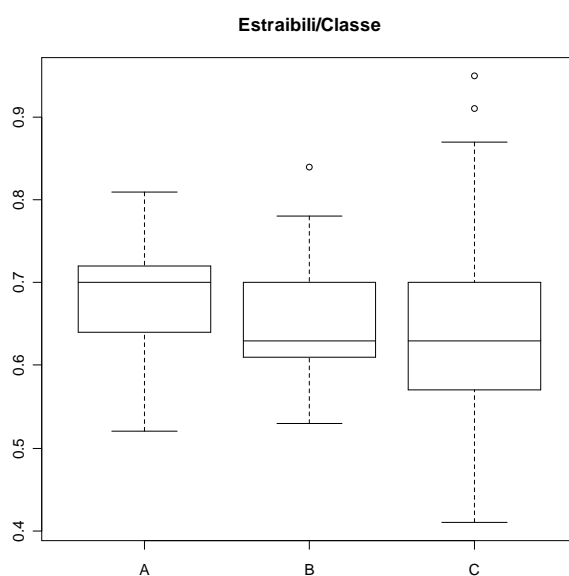
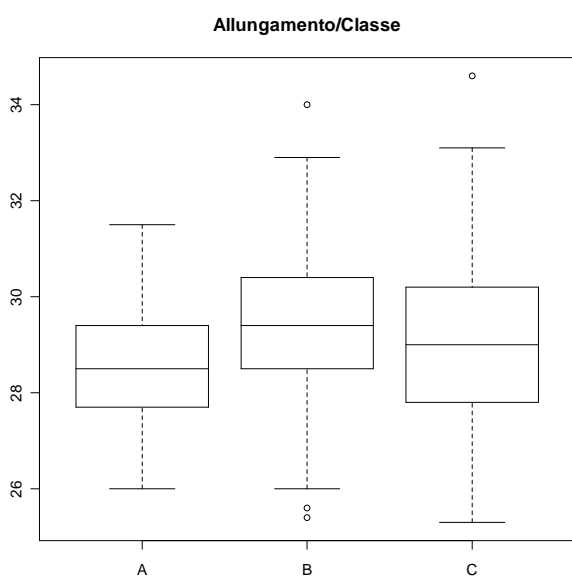
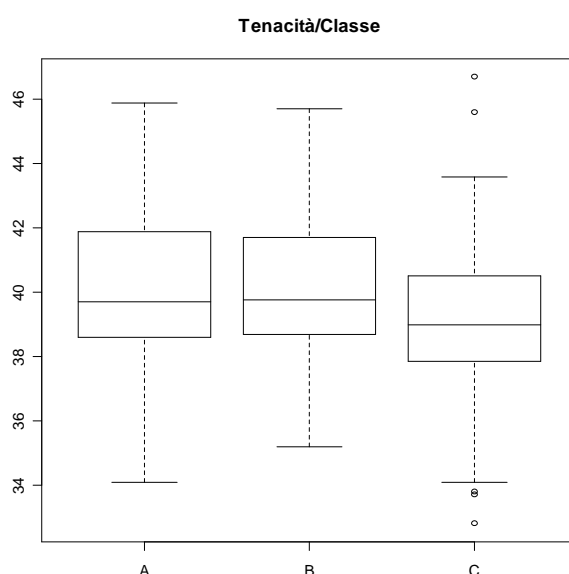
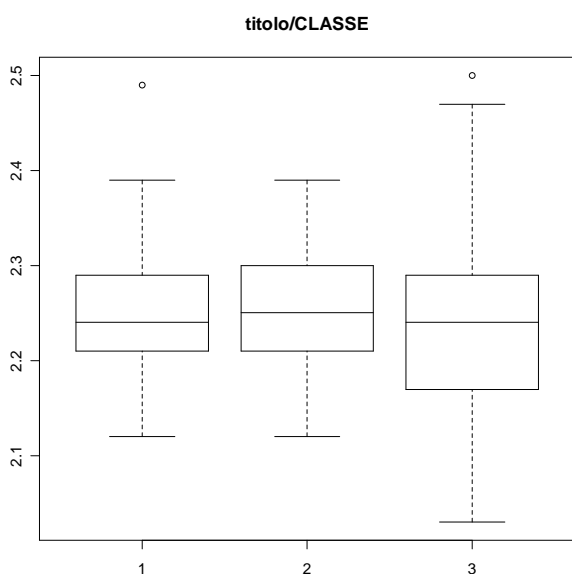
I dati aggiornati sono stati suddivisi e classificati (sulla base della quantità di miscela madre applicata al colore) in 3 classi :

Classe 1 o indicata con la lettera A = bassa concentrazione

Classe 2 o indicata con la lettera B = media concentrazione

Classe 3 o indicata con la lettera C = alta concentrazione

#### Analisi grafica



Le considerazioni per i test statistici sono state eseguite con l'utilizzo del Barlett.test per la verifica di omoschedasticità mentre viene utilizzato il comando "aov"<sup>28</sup> per l'analisi delle medie nelle tre classi.

Di seguito vengono riportati i valori dei p-value dei test:

### Suddivisione per classi

test per uguaglianza delle medie	p-value	Mc9	Mc19
<b>titolo</b>	0,204	0,2386	0,641
<b>tenacita</b>	0,0081	0,3584	0,186
<b>allungamento</b>	0,01942	0,196	0,0403
<b>estraibili</b>	0,0673	0,3808	0,0567

Bartlett.test omoschedasticità	p-value	Mc9	Mc19
<b>titolo</b>	9,49E-11	0,0489	0
<b>tenacita</b>	0.852	0,0035	0,1466
<b>allungamento</b>	0.03089	0,692	0,175
<b>estraibili</b>	0.0002177	0,186	0,0368

### Conclusioni seconda analisi

Per la caratteristica **titolo** la suddivisione nelle tre classi di colore, evidenziata nel test, ci induce a considerare un uguale valore in media ma non in varianza. Per la caratteristica **tenacità**, se non si considera la diversità di macchina, i test indicano differenze in media ma non in varianza nelle tre classi. Le cose cambiano invece, quando si inizia ad osservare i valori tenendo conto delle due macchine di produzione in cui si denota una diversità in media ma non in varianza per la Mc 19, mentre anche in varianza per la Mc 9.

Il test effettuato per la caratteristica **allungamento** ci porta a considerare le tre classi diverse sia in media che in varianza trascurando la diversificazione delle due macchine, la situazione opposta si verifica tenendo conto delle macchine di produzione (anche se si è al limite dell'accettazione per la Mc 19).

Infine, per la caratteristica **estraibili**, il risultato del test ci porta a considerare media uguale (al limite) e varianza diversa, mentre se consideriamo la separazione tra le due macchine si ottiene media e varianza uguali.

<sup>28</sup> Appendice D.

## 4.6 Considerazioni finali

Sulla base dei risultati ottenuti nelle analisi effettuate finora, le caratteristiche osservate non evidenziano generalmente forti relazioni con il colore tuttavia, visti i risultati dei test, si consiglia di approfondire lo studio con una numerosità maggiore.

La significatività dei test applicati invece, sottolinea come vi sia una differenza sia in media che in varianza tra le due macchine di produzione. Tale fatto dev'essere tenuto presente nell'identificazione come fonte di variabilità del processo produttivo e ci indirizza a scelte specifiche nelle considerazioni del paragrafo sottostante. Viene svolta infine un'analisi complementare per verificare se tra i fattori di variabilità emersi nello studio vi sia interazione. L'analisi effettuata non denota la presenza di interazioni (valore congiunto).

## 4.7 Studio delle componenti di variabilità del dato

Questo paragrafo descrive e propone un modello che cerca di spiegare le fonti di variabilità riscontrate nell'analisi statistica soprascritta, in modo da evidenziarne le componenti principali e definire così un'indice di variabilità globale del dato.

Come riportato nell'introduzione al controllo statistico di processo, siamo interessati a quella parte di variabilità dovuta a cause identificabili o determinabili, non naturali, che sono fonte primaria di un processo fuori controllo.

Lo studio del dato e di come esso sia stato ottenuto evidenzia due aspetti principali:

- il primo è strettamente legato al processo fisico di produzione;
- il secondo è legato alla rilevazione del dato stesso nel laboratorio.

Nel primo caso si denotano tutti quei fattori legati al processo produttivo che sono fonte di variabilità intrinseca e parzialmente controllabile.

Nel secondo caso invece, si evidenzia la presenza di una fonte di variabilità legata alla modalità di rilevazione del dato, ovvero una componente di variabilità riconducibile al risultato (finale) del test di laboratorio.

Le analisi statistiche effettuate sulla base dei dati presenti, definiscono nella componente di variabilità del processo, un valore statisticamente significativo causato dalla diversità di macchina di produzione (diversità strutturali di progettazione).

Dalla lettura del capitolo tre invece, si possono fare considerazioni tali da permettere il calcolo della componente di variabilità riconducibile al metodo di rilevazione della misura analitica ottenuta in laboratorio.

Infine nel modello conclusivo bisogna considerare quella componente di variabilità casuale che non viene spiegata fondamentalmente dalle altre due fonti e che viene considerata come componente di errore.

Il modello finale quindi si può considerare strutturato nel seguente modo:

**Dato = componente di variabilità + componente di variabilità + componente errore**  
*dovuta al processo produttivo    riconducibile alla rilevazione*  
*(test di lab)*



# Capitolo 5

## Carte di controllo

### 5.1 Nozioni di base

La carta di controllo é una delle principali tecniche adottate per il controllo statistico di un processo produttivo. Per carta di controllo si intende uno strumento statistico che consente di registrare misure sperimentali di uno o più caratteri rappresentativi della qualità di un processo produttivo, al fine di constatare se, nel corso del tempo, il processo stesso assume direzioni che possono modificare un prefissato livello di qualità giudicato soddisfacente.

Alla base delle considerazioni statistiche nella costruzione delle carte di controllo vi è la relazione con le procedure di verifica d'ipotesi. Tali procedure sono di rilevante utilità in molti problemi di controllo statistico della qualità e del processo.

Un'**ipotesi statistica** è una proposizione riguardante i valori dei parametri di una distribuzione di probabilità. La formulazione di un sistema di ipotesi avviene definendo l'**ipotesi nulla** che di solito viene scritta come  $H_0$  e l'**ipotesi alternativa**  $H_1$  (unilaterale o bilaterale).

Per effettuare una verifica d'ipotesi si osserva un campione casuale della popolazione allo studio, si calcola sulla base delle risultanze campionarie un'appropriata statistica test atta alla verifica in oggetto e, mediante tale statistica, si procede a rifiutare o ad accettare l'ipotesi nulla.

Nel caso specifico delle carte di controllo il legame con i test d'ipotesi può essere illustrato attraverso la seguente formulazione:

**$H_0$  : nessun cambiamento ← processo in controllo**

**$H_1$  : cambiamento in un istante temporale ← processo fuori controllo.**

In relazione alla procedura del test si possono verificare due tipi di errori. Se l'ipotesi nulla è rifiutata quando è vera diciamo che si è verificato un errore di primo tipo; se l'ipotesi nulla non è rifiutata (o accettata) quando essa non è vera, si commette un errore del secondo tipo.

Le probabilità relative a questi due tipi di errori sono indicate con  $\alpha$  e  $\beta$ , rispettivamente.

**$\alpha$  = P{errore del I tipo} = P{rifiutare  $H_0$ | $H_0$  è vera}**

**$\beta$  = P{errore del II tipo} = P{accettare  $H_0$ | $H_0$  è falsa}**

Nell'ambito del controllo di qualità,  $\alpha$  viene detto anche **rischio del produttore** (probabilità che unità conformi siano rifiutate).  $\beta$  invece, viene chiamato anche **rischio del cliente** (probabilità di accettare un lotto di bassa qualità).

In generale le procedure di verifica d'ipotesi comportano di specificare principalmente un valore della probabilità assegnata all'errore del primo tipo  $\alpha$ , e poi progettare o definire la procedura di verifica in modo che sia il più piccola possibile la probabilità dell'errore  $\beta$ .

Questo implica, nell'ambito delle applicazioni del controllo qualità, scegliere in modo diretto il rischio  $\alpha$ , mentre il valore del rischio  $\beta$  risulta generalmente una funzione della dimensione campionaria e quindi viene controllato in modo indiretto.

Più ampia è la dimensione dei campioni impiegati nella verifica, più piccolo è il rischio  $\beta$ . Il modo consueto per presentare i risultati di una verifica d'ipotesi è quello di indicare che l'ipotesi nulla sia o non sia rifiutata rispetto ad un prefissato valore  $\alpha$  che è detto livello di significatività. Per superare difficoltà che possono insorgere nelle conclusioni del test formale, è stato adottato ampiamente nella pratica l'approccio del P-valore.

Con il termine P-valore (*p-value*) si intende la probabilità che la statistica test assuma un valore che è almeno come estremo uguale a quello osservato quando l'ipotesi nulla  $H_0$  è vera. Quindi il p-valore contiene molte informazioni relative all'evidenza di non validità di  $H_0$  così, il responsabile delle decisioni può trarre una conclusione per ogni specifico livello di significatività.

La definizione del p-value può essere quindi riassunta come:

il più piccolo livello di significatività che conduce al rifiuto dell'ipotesi nulla  $H_0$ .

E' abitudine chiamare una statistica test (e i dati corrispondenti) significativa quando l'ipotesi nulla è rifiutata, quindi il p-value è il più piccolo livello  $\alpha$  per il quale i dati osservati sono significativi.

Una tipica carta di controllo (tra le varie tipologie a seconda della caratteristica di interesse nel processo) è costituita da una **linea centrale (LC)** e **limiti di controllo superiore e inferiore (UCL Upper Control Limit e LCL Lower Control Limit)**. La linea centrale indica il punto dove dovrebbe posizionarsi la caratteristica del processo produttivo (target) se non fossero presenti fonti di variabilità anomale, mentre i limiti di controllo vengono (individuati sulla base di considerazioni statistiche) scelti in modo tale che, se il processo è sotto controllo, quasi tutti i valori campionari cadranno al loro interno senza alcuna conseguente azione correttiva nel processo. Se invece un punto cade al di fuori di tali limiti si evidenzia la presenza di un processo fuori controllo e quindi si renderanno necessarie opportune azioni



correttive o di indagine sul processo per individuare ed eliminare le cause che hanno portato l'insorgere di quei fattori specifici.

## 5.2 Caratteristiche di un disegno di controllo

Nel disegno di uno schema di controllo bisogna considerare:

- specificazione dei limiti di controllo
- scelta della dimensione dei campioni
- frequenza di campionamento

La definizione dei limiti di controllo è uno dei passaggi più critici nella progettazione di una carta (disegno). Quanto più i limiti vengono posizionati lontano dalla linea centrale della carta di controllo, tanto minore sarà il rischio di I tipo, ovvero il rischio che un punto si posizioni al di fuori dei limiti, indicando una situazione di fuori controllo, quando invece nessun fattore specifico si è manifestato. D'altro canto l'ampliare i limiti di controllo comporta un aumento del rischio di II tipo, ovvero il rischio che un punto cada all'interno dei limiti di controllo quando di fatto il processo non è sotto controllo. Se i limiti vengono invece avvicinati alla linea centrale si otterrà l'effetto opposto: il rischio di I tipo aumenterà mentre quello di II tipo diminuirà.

La dimensione del campione nel disegno dello schema consente di individuare piccoli spostamenti o regolazioni nel processo. Infatti, la probabilità di individuare uno scostamento aumenta all'aumentare della dimensione campionaria; per scegliere la dimensione campionaria ottimale bisogna avere presente qual è lo scostamento del processo che si vuole individuare più velocemente.

Oltre alla numerosità del campione, bisogna determinare la frequenza di campionamento. La situazione ottimale sarebbe quella di poter esaminare grandi campioni di frequente, ma tale situazione è sicuramente poco accettabile dal punto di vista economico.

Uno strumento utile per calcolare l'ottimale dimensione campionaria e la frequenza di campionamento è la lunghezza media delle sequenze **ARL** (*Average Run Length*).

L'ARL è il numero medio di punti che devono essere osservati prima che un punto cada al di fuori dei limiti di controllo.

$$ARL = \frac{1}{p} \quad \text{dove } p \text{ è la probabilità che un punto superi i limiti di controllo.}$$

I requisiti di uno schema di controllo sono:

- segnali raramente un falso allarme
- segnali il più presto possibile un effettivo fuori controllo.

Sulla base di questi requisiti e considerando la misura di efficienza di uno schema di controllo si può concludere che:

**Processo in controllo** → **ARL più grande possibile**

**Processo fuori controllo** → **ARL più piccola possibile**

### **5.3 Fasi operative per il calcolo del periodo di controllo**

#### **5.3.1 Fase 1 e fase 2**

Il controllo della qualità da effettuare durante un processo produttivo presuppone la conoscenza di standards ideali riferiti alle caratteristiche del prodotto. Ogni carattere del prodotto viene trattato come una variabile casuale ed il controllo futuro ha lo scopo di verificare se ci sono scostamenti rilevanti dagli stessi standards. Quello che possiamo controllare è la misura media del carattere in un campione e la sua varianza. Se la media e la varianza non sono note devono essere stimate. È opportuno che le stime vengano effettuate mentre il processo produttivo è ritenuto soddisfacente, ossia mentre è in controllo. Una prima fase di lavoro consiste pertanto nell'effettuare un campionamento in un periodo che verrà chiamato "periodo di base". Se il periodo di base prescelto fosse fuori controllo si correrebbe il rischio di costruire una carta di controllo "ottimistica", ovvero che induce ad accettare lotti futuri di produzione quando invece sarebbero da scartare.

La seconda fase operativa prevede il riscontro della carta costruita nel periodo di base con previsioni di osservazioni future e come si comporta in presenza di valori anomali che generano ovvero condizioni di fuori controllo. In questo modo si verifica l'efficienza e l'efficacia della carta di controllo. Tale descrizione trova impiego nella costruzione delle carte di controllo multivariate (in particolare per la carta  $T^2$  di Hotelling) e per le carte univariate EWMA, tuttavia il metodo può essere generalizzato nella definizione del disegno della carta.

#### **5.3.2 Caso studio**

Per la costruzione delle carte di controllo, in relazione a quanto detto sopra, è stato analizzato un particolare periodo "di base" di produzione in cui l'andamento di filatura delle due macchine risulta meno soggetto a variazioni dovute ad interventi operativi o a difetti presenti nel corso del processo produttivo. In particolare sono state consultate le tabelle di marcia operativa e la presenza di difetti (codificati e rilevati secondo un ordine ponderato in base alla gravità/entità del difetto) durante la lavorazione per entrambe le macchine nel corso del periodo interessato. In questo modo è stato ottenuto un indice che definisce l'andamento operativo in relazione ai legami delle caratteristiche di studio con il procedimento di

produzione. Tale periodo viene considerato per la costruzione dei limiti di controllo della carta EWMA, T<sup>2</sup> di Hotelling ed MEWMA.

## 5.4 Carte di controllo per misure singole

### 5.4.1 Carte singole

La carta di controllo singola permette di accertare che la distribuzione di una data caratteristica misurabile rimanga costante nel tempo, si tratta ovvero di monitorare sia la media che la varianza. Le carte per misure singole permettono di sorvegliare contemporaneamente questi due parametri. Tali carte trovano maggior impiego nei casi in cui le variabili oscillano poco e la deviazione standard è molto piccola rispetto alla globalità della produzione.

Quando i parametri  $\mu$  e  $\sigma$  (media e varianza) sono ignoti devono essere stimati sulla base di un campione; ciò che viene controllato quindi è il valore campionario del parametro che identifica la distribuzione della caratteristica in esame.

In particolare le carte per misure singole prevedono che il parametro da monitorare venga campionato singolarmente. Una stima della varianza si basa sui range mobili (MR). Nello studio della variabilità si procede alla costruzione di una serie di misure dette “escursioni mobili” date dalla differenza tra ogni valore e quello precedente.

Date  $m$  osservazioni, i parametri di interesse vengono stimati da:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_m}{m}$$

$$\overline{MR} = \frac{MR_1 + MR_2 + \dots + MR_m}{m} \quad ; \quad \text{dove } MR_t = |x_t - x_{t+1}|$$

Ne segue che i limiti di controllo della carta per misure singole e per la carta sulle Escursioni Mobili (MR) sono rispettivamente:

$$\begin{aligned} UCL &= \bar{x} + L \frac{\overline{MR}}{d_2} \\ LC &= \bar{x} \\ LCL &= \bar{x} - L \frac{\overline{MR}}{d_2} \end{aligned}$$

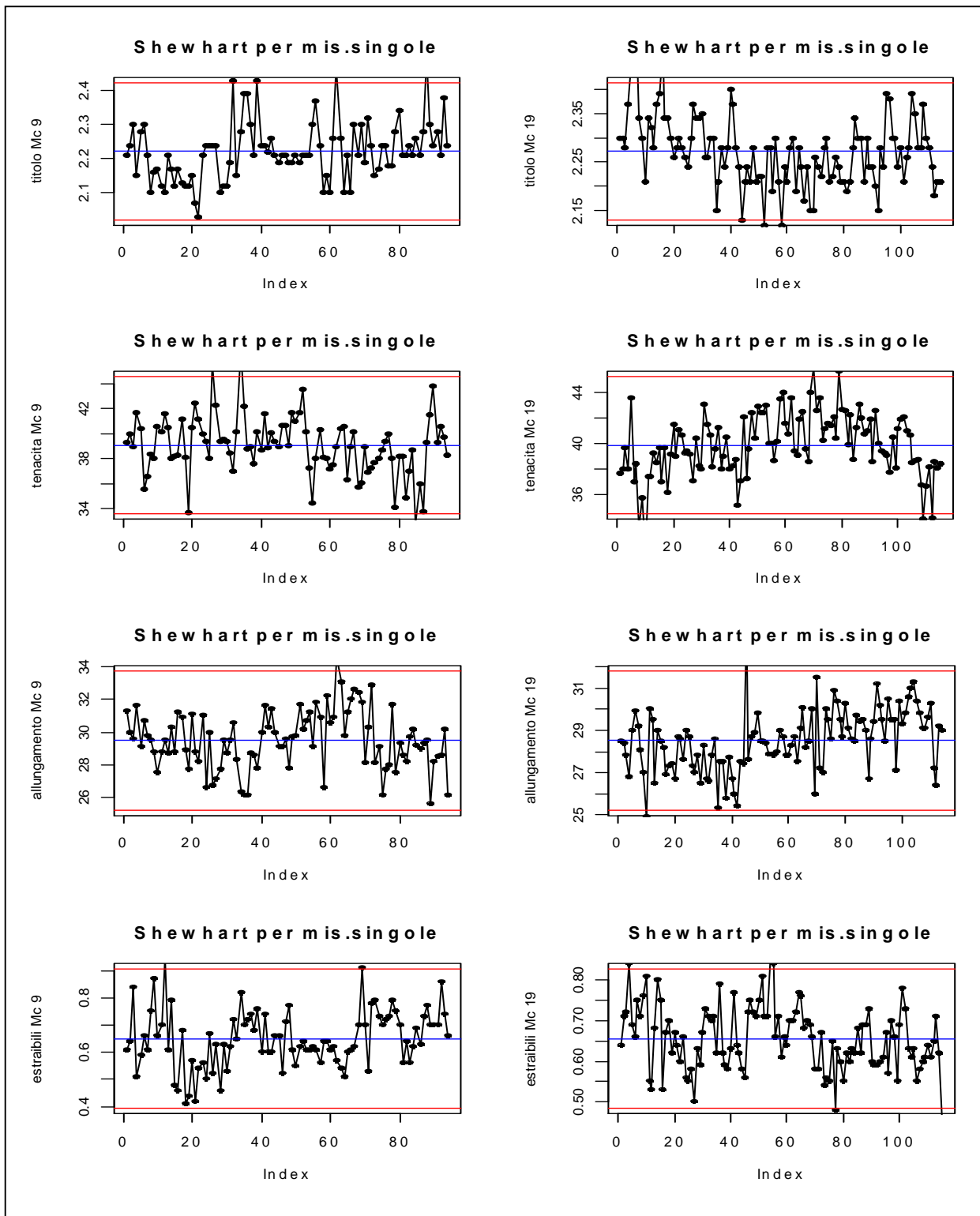
$$\begin{aligned} UCL &= \overline{MR} + Ld_3 \frac{\overline{MR}}{d_2} \\ LC &= \overline{MR} \\ LCL &= \overline{MR} - Ld_3 \frac{\overline{MR}}{d_2} \end{aligned}$$

Le costanti  $d_2$  e  $d_3$  sono costanti tabulate in base alla numerosità campionaria  $n$ .

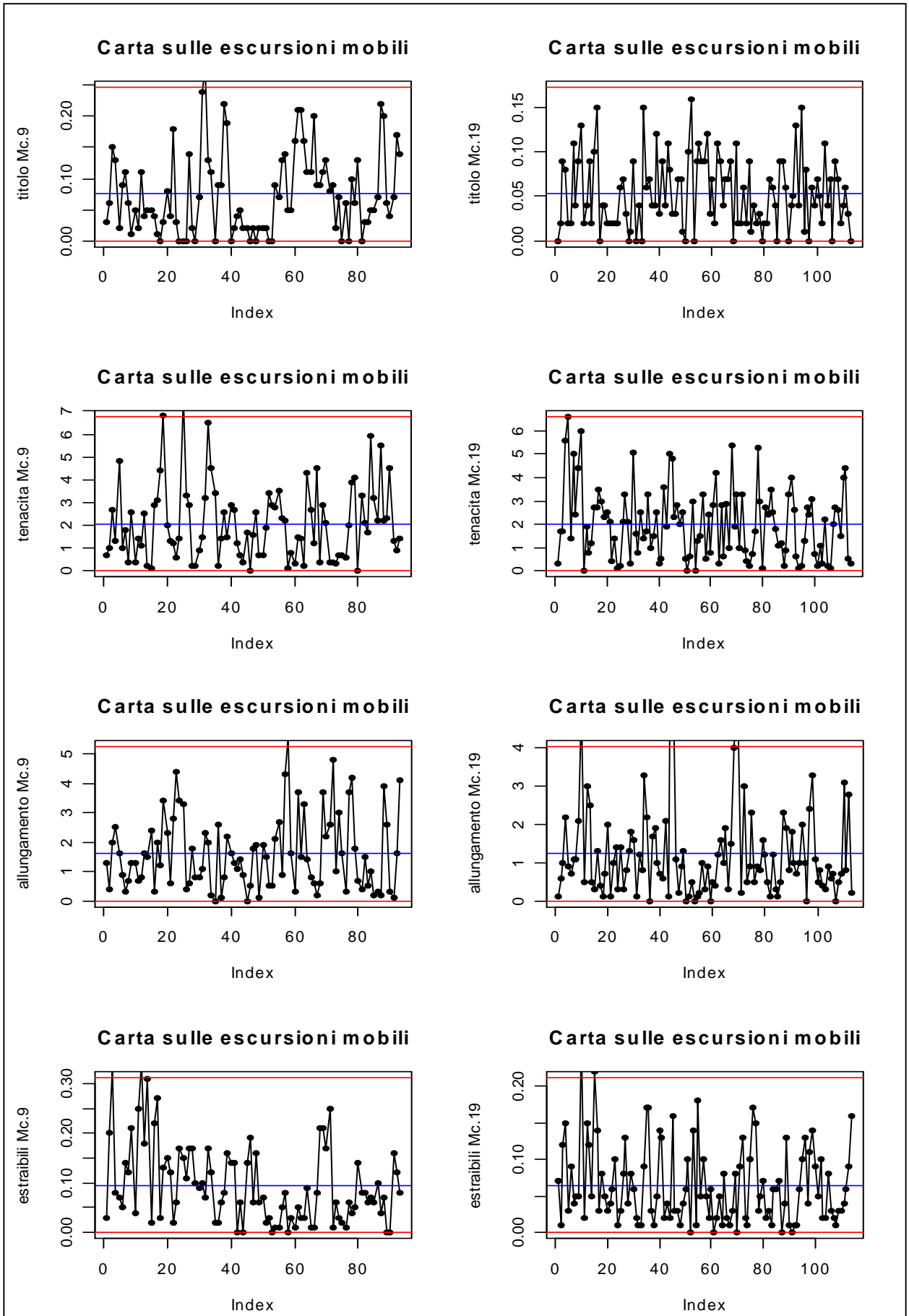
Sostituendo  $L = 3$  si può dire che l’intervallo contiene il 99,73% delle osservazioni e quindi un valore al di fuori evidenzierà una causa sistematica.

### 5.4.2 Costruzione carte per misure singole

Di seguito viene riportata la carta Carta Shewhart per le singole misure con la stima della varianza basata su MR(2) e la Carta sulle escursioni mobili divisa per le due macchine di filatura<sup>29</sup>.



<sup>29</sup> Le funzioni utilizzate ed i dettagli sono riportati in Appendice E.



### 5.4.3 Carte EWMA

Le carte Shewhart illustrate precedentemente hanno un grosso limite: esse fanno uso solamente dell'informazione sul processo contenute nell'ultimo istante di osservazione.

Le carte EWMA, carte a medie mobili pesate esponenzialmente (Roberts, 1959), costituiscono a tal proposito un'utile alternativa, fanno infatti uso delle informazioni precedenti rendendo sensibile la carta a piccole modifiche del livello medio del processo.

La statistica di controllo consiste in una media mobile ponderata delle osservazioni passate, dove viene assegnato a ciascun valore un peso  $\lambda$ , e si definisce nel seguente modo:

$$z_i = \lambda x_i + (1 - \lambda)z_{i-1}$$

dove  $0 < \lambda \leq 1$  è una costante e il valore iniziale  $z_0 = \bar{x}$  è la stima del valore di riferimento ignoto  $\mu_0$ . Se sostituiamo ricorsivamente gli  $z_{i-1}$  con i rispettivi valori otteniamo:

$$z_i = \lambda \sum_{j=0}^{i-1} (1 - \lambda)^j x_{i-j} + (1 - \lambda)^i z_0$$

I pesi  $(1 - \lambda)^j$  decrescono in progressione geometrica e la loro somma è pari all'unità.

Se gli  $x_i$  sono determinazioni di variabili casuali indipendenti di varianza comune  $\sigma^2$ , la varianza della statistica EWMA è pari a:

$$\sigma_{z_i}^2 = \sigma^2 \left( \frac{\lambda}{2 - \lambda} \right) [1 - (1 - \lambda)^{2i}]$$

Il termine:  $1 - (1 - \lambda)^{2i} \rightarrow 1$  per  $i \rightarrow \infty$ ; ne consegue che il valore asintotico della varianza è:

$$\sigma_{z_i}^2 = \sigma_0^2 \left( \frac{\lambda}{2 - \lambda} \right)$$

I limiti di controllo per la carta EWMA sono:

$$UCL = \bar{x} + L\sigma \sqrt{\frac{\lambda}{2 - \lambda}} [1 - (1 - \lambda)^{2i}]$$

$$LC = \bar{x}$$

$$LCL = \bar{x} - L\sigma \sqrt{\frac{\lambda}{2 - \lambda}} [1 - (1 - \lambda)^{2i}]$$

Al crescere di  $i$  i limiti di controllo della carta EWMA si stabilizzano intorno ai valori asintotici:

$$UCL = \bar{x} + L\sigma \sqrt{\frac{\lambda}{2 - \lambda}}$$

$$LC = \bar{x}$$

$$LCL = \bar{x} - L\sigma \sqrt{\frac{\lambda}{2 - \lambda}}$$

Come stima di  $\sigma$  possiamo usare  $\overline{MR}/d_2$ . Dove  $\overline{MR}$  è la media dei range mobili.

La scelta dei parametri  $L$  ed  $\lambda$  è importante, con il primo si stabilisce l'ampiezza dei limiti di controllo, con il secondo si stabilisce il peso che si vuol dare all'informazione passata.

Diversi studi di carattere teorico sono stati svolti da Crowder (1987), Lucas e Saccucci (1990) dai quali si ricavano tabelle dei valori di  $L$  ed  $\lambda$ .

In generale si è trovato che per valori di  $\lambda$  compresi nell'intervallo 0,05-0,25 la carta EWMA dà buone prestazioni nel riuscire a segnalare piccoli spostamenti del livello medio del processo.

Si è inoltre riscontrato che per  $\lambda > 0,1$  utilizzare  $L = 3$  costituisce una scelta particolarmente ragionevole (Montgomery, 1996).

Data la struttura delle EWMA, si possono ricavare molte proprietà: è adatta per l'individuazione di salti di entità limitata, è robusta rispetto alla non normalità delle singole misure, considera la correlazione tra osservazioni passate, evidenzia trend di deriva anche con poche osservazioni, è ideale infine per carte per misure singole.

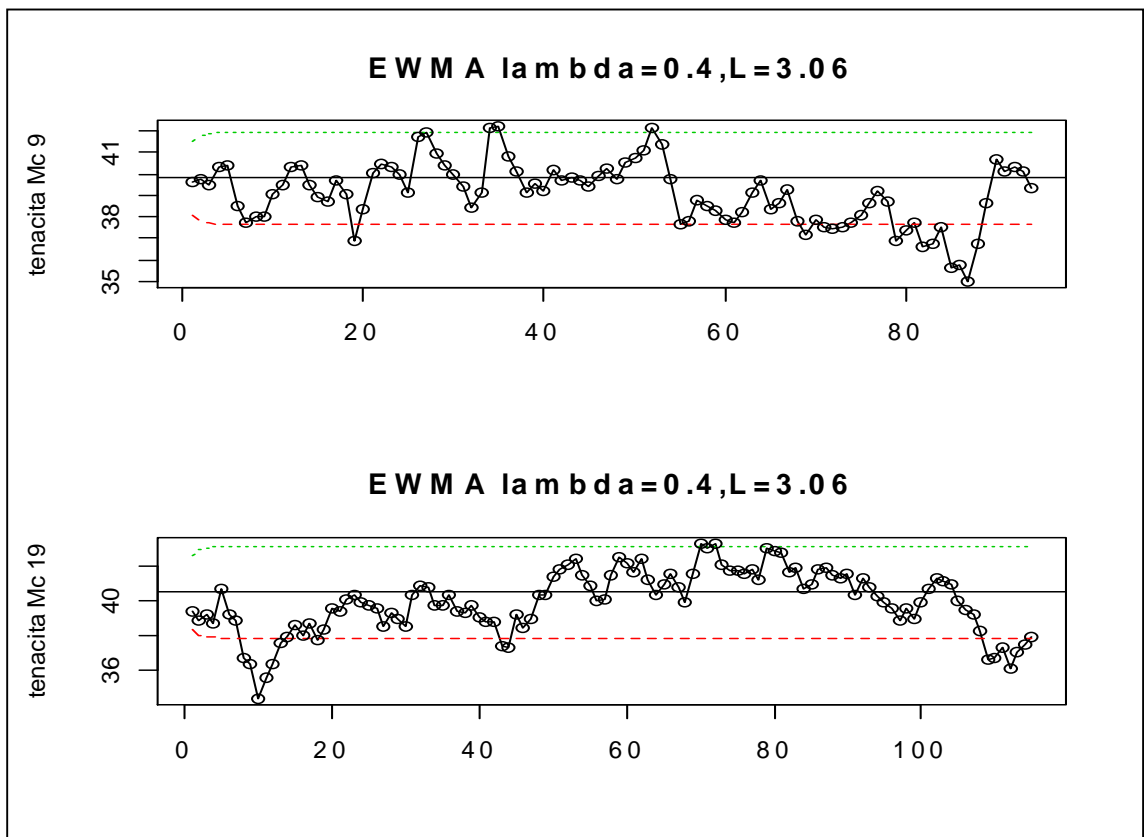
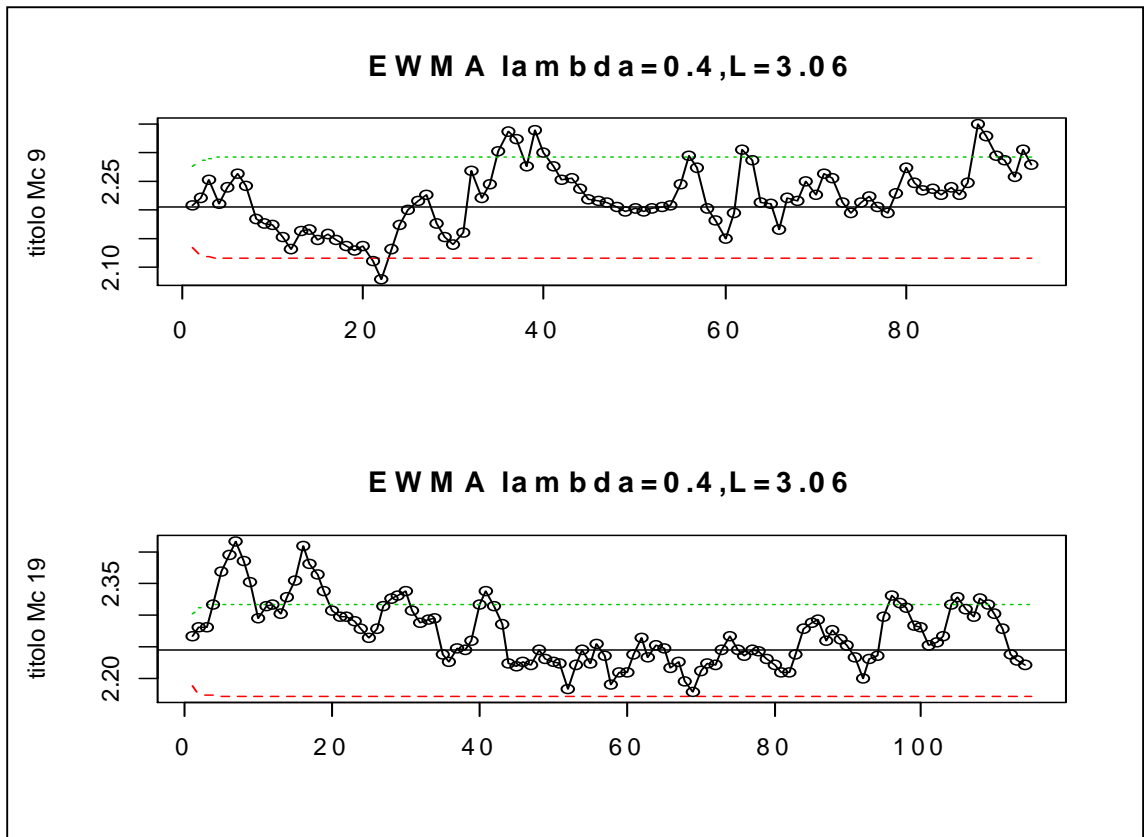
Un altro vantaggio delle carte EWMA è che necessita di un solo grafico dove sono riportati i valori della statistica  $z_i$  in funzione del numero di campioni  $i$ , la linea centrale e i limiti di controllo superiore e inferiore assumeranno i valori di LC, UCL, LCL rispettivamente.

Quando il processo è in controllo la statistica campionaria fluttua attorno al valore medio. La procedura di ponderazione introduce una certa inerzia: la carta EWMA non reagisce in maniera veloce a bruschi cambiamenti, i suoi valori si spostano lentamente verso un nuovo livello seguendo il cambiamento nel processo, l'individuazione di un cambiamento si basa sull'identificazione di un trend nel valore della statistica e non solo sulla segnalazione di fuori controllo.

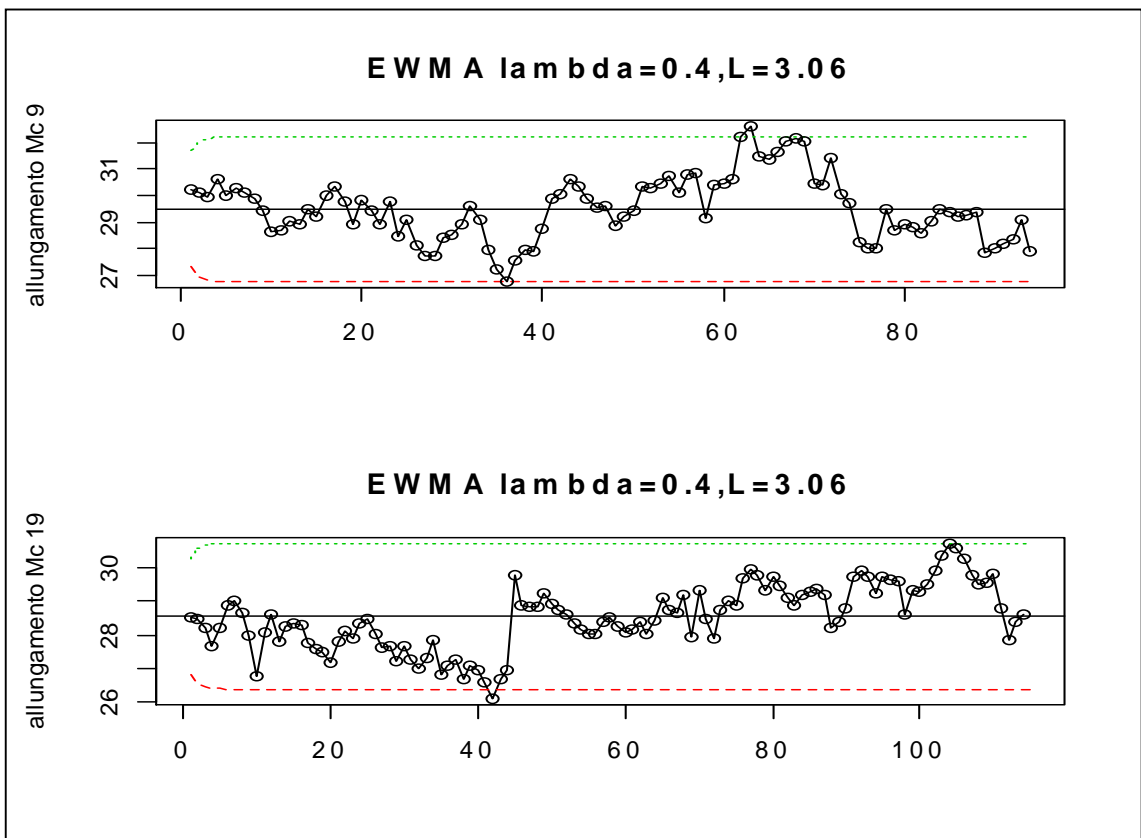
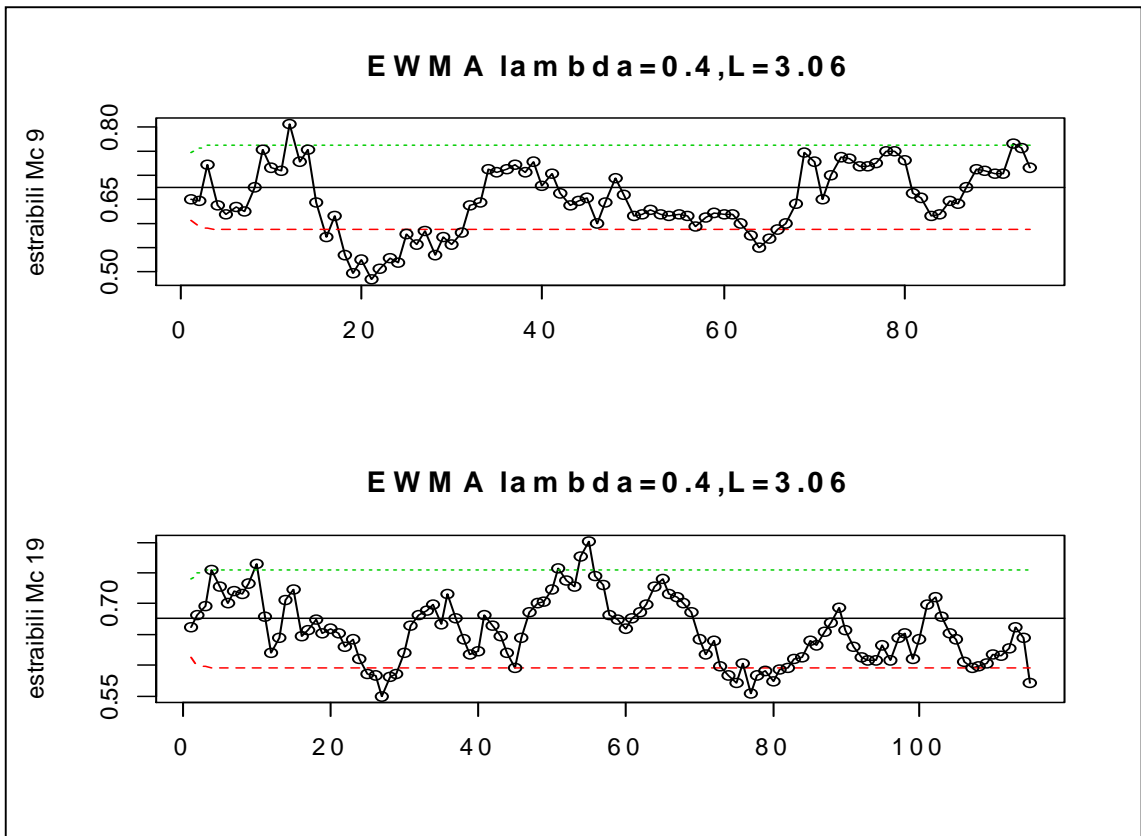
#### 5.4.4 Caso studio

Segue la costruzione delle carte EWMA con definiti valori di  $\lambda$  ed  $L$  per le caratteristiche di interesse. I valori dei limiti di controllo sono stati assegnati considerando un particolare periodo di controllo della macchina di filatura a cui si sono aggiunti i valori completi del vettore, mentre i valori di  $\lambda$  ed  $L$  sono stati scelti dopo opportune verifiche (valori tabulati)<sup>30</sup>.

<sup>30</sup> Crowder (1989), Lucas e Saccucci (1990), Brook e Evans (1972)







## 5.5 Controllo multivariato

### 5.5.1 Introduzione

In molte situazioni nel controllo statistico di processo le variabili coinvolte sono molte, diviene necessario quindi tenere sotto controllo quelle caratteristiche che influenzano il processo.

I problemi nei quali interessa tenere monitorate più variabili sono spesso denominati problemi di **controllo multivariato della qualità**.

Tra le varie procedure e carte di controllo nell'ambito del multivariato quella che sicuramente ha trovato molti impieghi è la statistica  $T^2$  di Hotelling. Tuttavia visti i problemi insorti di normalità dei dati e come confronto, viene trattata anche la carta MEWMA (versione multivariata della carta di controllo EWMA).

### 5.5.2 La distribuzione $T^2$ di Hotelling

La distribuzione T-quadrato di Hotelling (chiamata così secondo Harold Hotelling) è una generalizzazione della variabile casuale  $t$  di Student e utilizzata nei test di ipotesi multivariati.

La statistica T-quadrato di Hotelling è definita come segue:

Siano  $x_1, \dots, x_n$

$p \times 1$  vettori colonna di numeri reali e

$\bar{x} = (x_1 + \dots + x_n) / n$  le loro medie.

Sia  $W = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})' / (n-1)$

la matrice non negativa data dalla loro varianza. Considerando  $\mu$  un vettore colonna  $p \times 1$  noto (in applicazione dei valori medi ipotizzati per la popolazione), la statistica T-quadrato di Hotelling è data da  $T^2 = (\bar{x} - \mu)' W^{-1} (\bar{x} - \mu)$

Se  $x \sim N_p(\mu, V)$  è una variabile casuale con una distribuzione normale multivariata e  $Q \sim W_p(m, V)$  è distribuita come una variabile casuale di Wishart<sup>31</sup>, con  $x$  e  $Q$  indipendenti, allora  $T^2$  è distribuita come una variabile casuale T-quadrato di Hotelling.

Si può dimostrare che se  $x_1, \dots, x_n \sim N_p(\mu, V)$ , sono indipendenti e  $\bar{x}$  e  $W$  sono definiti come sopra allora  $W$  è distribuita come una variabile casuale di Wishart con  $m = n - 1$  gradi di libertà ed è indipendente da  $\bar{x}$  e

$$\bar{x} \sim N_p(\mu, V/n)$$

<sup>31</sup> Appendice A.

Inoltre, se entrambe le distribuzioni sono non-singolari, si può dimostrare che:

$$\frac{m-p+1}{pm} T^2 \sim F_{p, m-p+1}$$

dove  $F$  è la variabile casuale  $F$  di Snedecor.

### 5.5.3 Costruzione della carta $T^2$ di Hotelling

#### FASE I

Il metodo tradizionale usato per la prima fase di costruzione della carta si deve ad Hotelling (1947). Se il campionamento di base (fase uno) viene effettuato con  $m$  estrazioni dal solito vettore  $x = (px1)$  normale di medie  $\mu$  e varianze  $\Sigma$  allora è possibile controllare la qualità nel periodo di base utilizzando la statistica (caso con dati individuali, numerosità campione  $n = 1$ ):  $T^2 = (x - \bar{x})^T S^{-1} (x - \bar{x})$

(dove  $S$  rappresenta la matrice di varianze e covarianze stimata nel periodo di base) che si distribuisce approssimativamente come:  $\frac{p(m-1)}{(m-p)} F_{(p, m-p)}$  di Snedecor

Il limite superiore di controllo diventa:

$$UCL = \frac{p(m-1)}{(m-p)} F_{\alpha} \quad \text{con } F_{\alpha} \text{ } \alpha\text{-esimo percentile della } F_{(p, m-p)} \text{ di Snedecor.}$$

Si tratta di distribuzione approssimata in quanto i vettori  $p$ -variati  $\bar{x}$  non sono indipendenti dalla matrice  $S$ , tuttavia, tale dipendenza influisce sempre meno al crescere della numerosità campionaria “ $m$ ”.

L’errore del I tipo ( $\alpha$ ) è la probabilità, in questo caso, che un vettore di dati  $x = (px1)$  (o la relativa statistica  $T^2$ ) cada fuori dal limite di controllo, nonostante il processo sia in controllo.

Il periodo di base, quindi le relative stime di  $S$  e  $x$  sono da ritenere accettabili se il numero di vettori anomali non supera l’ $\alpha$  percentuale della dimensione  $m$  del campione. Vale a dire che se il livello  $\alpha$  viene fissato pari al 5% allora si possono accettare al massimo 1,25 dati anomali su  $m = 25$ .

Se dovesse accadere per esempio che più di un dato venga considerato segnale di “fuori controllo” (con  $\alpha = 5\%$ ), si prosegue l’analisi con un esame particolareggiato di tutto il contesto di rilevazione dei dati anomali e si conclude che:

- a) tutti o alcuni dei dati anomali risentono di cause esterne accidentali che nulla hanno a che vedere con le caratteristiche richieste al processo. In tal caso è sufficiente eliminare quei dati, ricalcolare le statistiche  $T^2$  sui dati che restano (quindi le nuove stime di  $S$  e  $\bar{x}$ ) e ripetere il controllo;
- b) non avendo riscontrato cause esogene che alterassero i risultati, il processo è realmente fuori controllo e quindi non può essere preso come periodo-base per la prima fase di costruzione della carta. Si dovrà rivedere l'impianto di produzione per poi effettuare un nuovo campionamento.

## FASE II

Prendendo come riferimento i valori del vettore medio  $\bar{x}$  e la matrice di varianza e covarianza  $\hat{S}$  calcolati sulla base dei dati in controllo ottenuti nel periodo di base, si calcola la statistica  $T^2$  per ognuno dei campioni provenienti dal processo e indipendenti da quelli considerati in precedenza.

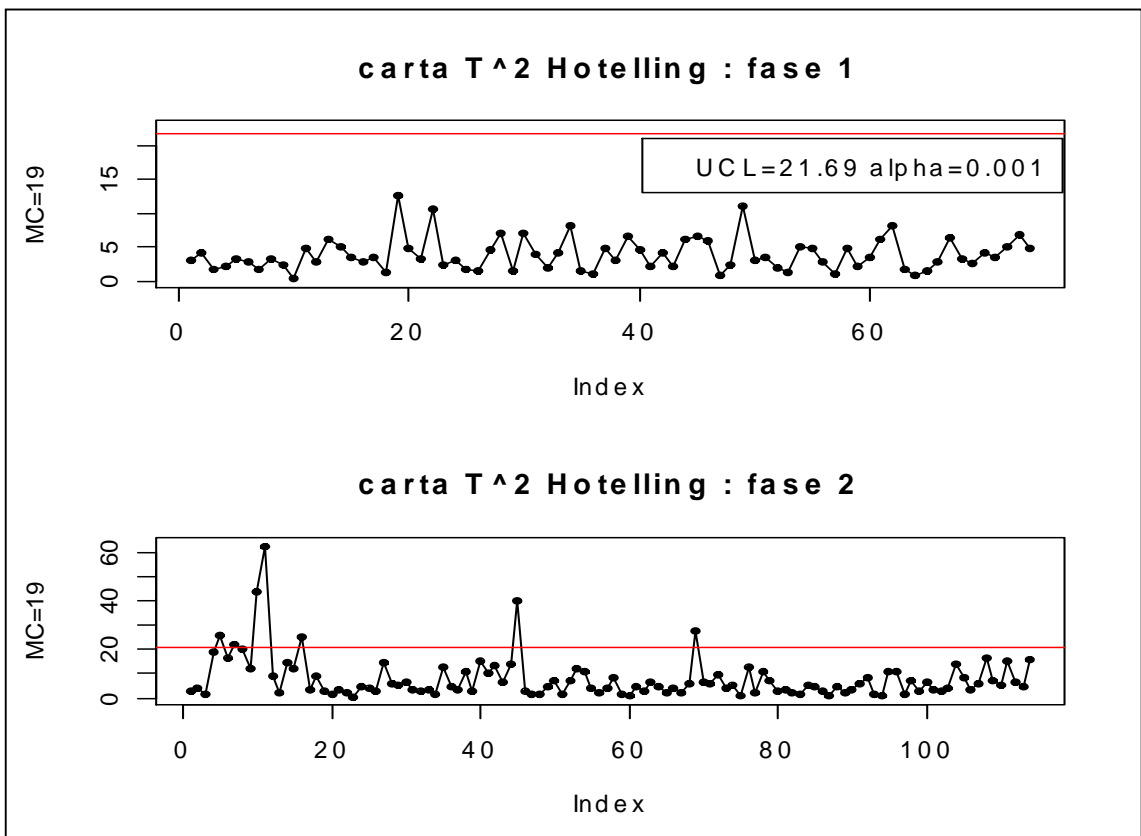
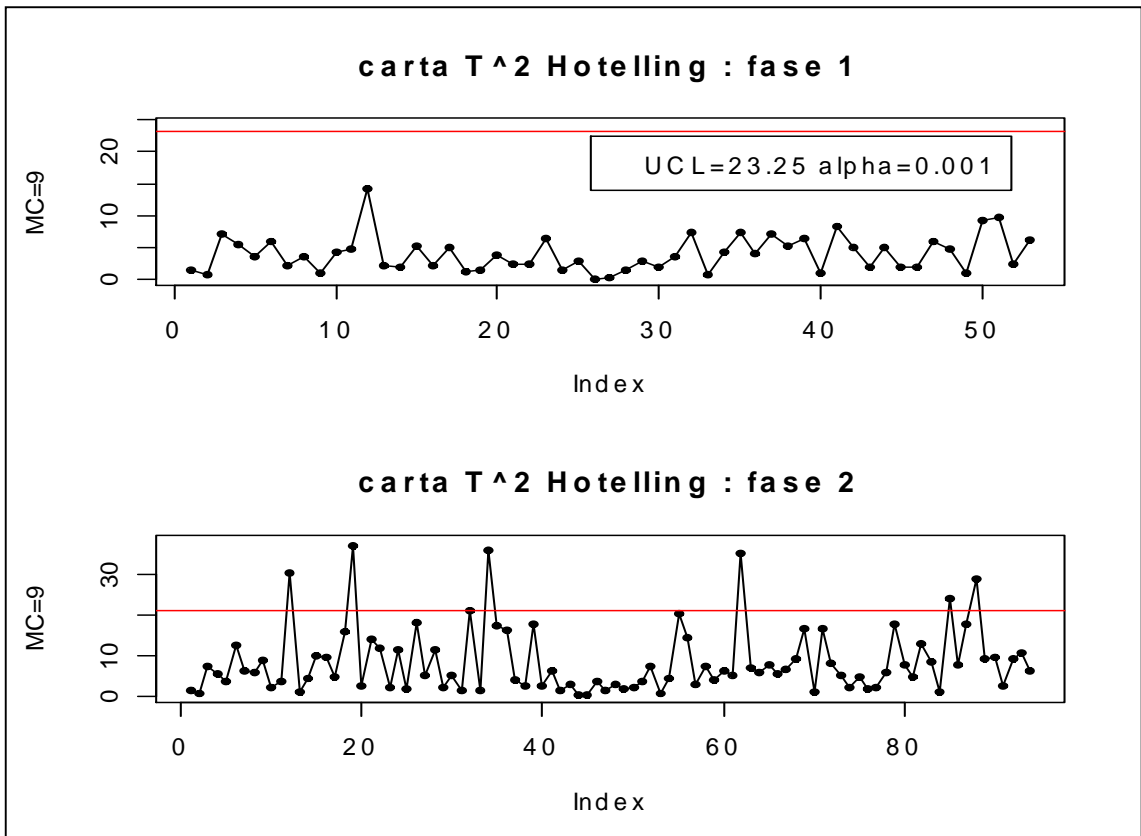
Il limite superiore di controllo per la statistica  $T^2$  è dato da:

$$UCL = \frac{p(m+1)(m-1)}{m^2 - mp} F_{\alpha, p, m-p}$$

### 5.5.4 Caso studio

Di seguito vengono riportate le carte suddivise per macchina di produzione e nelle due fasi operative.

Il valore del percentile scelto nel calcolo dell'UCL viene indicato nel grafico vicino alla corrispondente linea nella sola prima fase operativa.



### 5.5.5 La carta MEWMA

C.A.Lowry et al. (1992) hanno presentato nel dettaglio l'estensione della carta EWMA univariata al caso di più variabili, indicandola come “carta **MEWMA**” (*Multivariate Exponentially Weighted Moving Average*).

Sia  $x_1, x_2, \dots, x_i$  una sequenza di vettori di  $p$  variabili casuali distribuiti normalmente con vettore delle medie  $\mu_i$  e matrice di dispersione  $\Sigma$ .

Si assuma nota la matrice di dispersione, sebbene di fatto essa venga stimata durante un periodo di base ritenuto in controllo, come illustrato precedentemente.

I vettori  $p$ -variati possono rappresentare singole osservazioni oppure possono essere medie campionarie del gruppo (sottogruppi) di osservazioni per ogni istante campionario  $\{i=1,2,\dots\}$ .

Si assuma l'indipendenza nella sequenza dei campionamenti e sia noto e posto pari a zero il vettore “ottimale” delle medie  $\mu_0$  (target del processo quando è in controllo).

L'obiettivo della carta di controllo è quello di rilevare scostamenti significativi della media del processo nel corso del tempo, dal vettore ottimale.

La carta MEWMA si basa sulla seguente quantità:

$$Z_i = \Lambda x_i + (1 - \Lambda)Z_{i-1} \quad (1)$$

dove  $\Lambda$  rappresenta la matrice diagonale con i valori  $\lambda_1, \lambda_2, \dots, \lambda_p$ ; si assume che

$$\lambda_1 = \lambda_2 = \dots = \lambda_p = \lambda \text{ e si pone il valore } Z_0 = 0.$$

La quantità rappresentata sulla carta è:

$$T_i^2 = Z_i^T \Sigma_{Z_i}^{-1} Z_i$$

In cui la matrice di covarianza è:

$$\Sigma_{Z_i} = \frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2i}] \Sigma$$

È da notare che la statistica  $T_i^2 = Z_i^T \Sigma_{Z_i}^{-1} Z_i$  coincide esattamente con la  $T^2$  di Hotelling quando si pone  $\lambda = 1$ , ossia quando nella media mobile rappresentata dalla (1) si trascura completamente l'informazione passata ( $x_{i-1}, x_{i-2}, \dots$ ) dando importanza solo all'osservazione più recente  $x_i$ .

### 5.5.6 ARL e determinazione dei parametri per la carta

Le potenzialità della carta MEWMA si possono evidenziare con un'analisi dell'ARL (Average Run Length). Sia:

$$\delta = [(\mu - \mu_0)^T \Sigma^{-1} (\mu - \mu_0)]^{1/2} = (\mu^T \Sigma^{-1} \mu)^{1/2}$$

il parametro di non centralità, sotto l'ipotesi che  $\mu_0 = 0$ .

Partendo da un valore desiderato dell'ARL dato lo stato di "in controllo" del processo, che indicheremo con  $ARL/\delta = 0$ , si hanno in corrispondenza più coppie  $(\lambda, L)$  a cui seguono andamenti diversi dell'ARL per valori di  $\delta > 0$ .

In particolare, Lowry et al (1992) propongono dei valori per alcune coppie  $(\lambda, L)$  illustrando l'andamento dell'ARL con diverse numerosità di variabili. Dallo studio svolto da Lowry et al si evidenzia che a valori piccoli delle coppie  $(\lambda, L)$  corrisponde una maggior efficienza nella rilevazione di piccoli scostamenti.

Sulla base di quanto detto sopra, si è in grado di tabulare coppie  $(\lambda, L)$  "ottime" per ogni valore dello scostamento  $\delta$  da controllare. L'aggettivo "ottimo" sta a significare che viene scelta la coppia  $(\lambda, L)$  che minimizza l'ARL in corrispondenza di uno specifico scostamento per un dato valore  $ARL/\delta = 0$ . Alcune di queste coppie vengono illustrate nella tabella 1.

Le carte di controllo MEWMA sono particolarmente indicate per rilevare piccole deviazioni ( $\delta = 0.5, \delta = 1.0$ ) dal vettore medio. La tabella sottostante, tratta dal lavoro di Lowry et al. (1992), evidenzia inoltre l'attitudine dello schema MEWMA a rilevare scostamenti  $\delta$  di varia dimensione, per diversi valori di  $ARL/\delta = 0$  e per alcuni valori di "p" (variabili sotto controllo).

TABELLA 1: Coppie  $(\lambda, L)$  ottime per schema di controllo MEWMA con  $ARL = 200/\delta = 0$   
 $\lambda^*$  = intervallo di valori "λ" ugualmente ottimi.

$\delta$		<b>p = 2</b>	<b>p = 3</b>	<b>p = 4</b>
0.5	$\lambda$	0.06	0.06	0.06
	L	7.70	9.80	11.66
	ARL{min}	26.5	30.0	32.3
	$\lambda^*$	0.04-0.06	0.04-0.06	0.04-0.08
1.0	$\lambda$	0.16	0.16	0.14
	L	9.35	11.52	13.34
	ARL{min}	9.95	11.0	12.0
	$\lambda^*$	0.10-0.20	0.10-0.20	0.10-0.16
1.5	$\lambda$	0.24	0.22	0.20
	L	9.90	11.96	13.84
	ARL{min}	5.47	6.02	6.45
	$\lambda^*$	0.18-0.30	0.20-0.30	0.18-0.30
2.0	$\lambda$	0.34	0.30	0.28
	L	10.17	12.31	14.25
	ARL{min}	3.53	3.87	4.20
	$\lambda^*$	0.24-0.34	0.24-0.34	0.22-0.34

Molto spesso quindi, viene scelto il valore  $\lambda = 0.1$  che, come si può rilevare dalla tabella 1 è adatto per scostamenti limitati. Attraverso metodi di simulazione numerica si possono ricavare i valori dell'ARL ed è inoltre possibile stimare i limiti di controllo grazie alla relazione lineare che lega approssimativamente le coppie  $[\ln(ARL/\delta = 0), L]$ .

Lowry et al. (1992) raccomandano di impostare una carta MEWMA operando nel modo seguente:

- 1) decidere soggettivamente quale scostamento si vuole controllare. È opportuno approfittare delle caratteristiche dello schema a media mobile andando a controllare scostamenti limitati;
- 2) fissare quindi il valore  $\lambda$  ottimo per lo scostamento deciso al punto (1), tenendo conto del numero di variabili coinvolte (vedi tabella 1);
- 3) fissare il limite di controllo "L" sulla base dell'ARL che si desidera quando il processo è in controllo.

La dipendenza di ARL dal parametro di non centralità  $\delta$ , che sta alla base delle considerazioni fatte finora, è stata dimostrata nel lavoro di Lowry et al. con il risultato che variazioni sul vettore delle medie influenzano l'ARL solo attraverso il parametro  $\delta$ . Se variazioni del vettore  $\mu$  non spostano il parametro  $\delta$  allora l'ARL non viene in alcun modo modificato (Crosier 1988).

Nella tabella seguente, sempre tratta da Lowry et al., vengono riportati i valori dell'ARL nel caso con  $p = 4$  variabili, i corrispondenti valori dei limiti di controllo con il parametro  $\lambda = 0.1$  ed un ARL in controllo di 50, 100, 500, 1000.

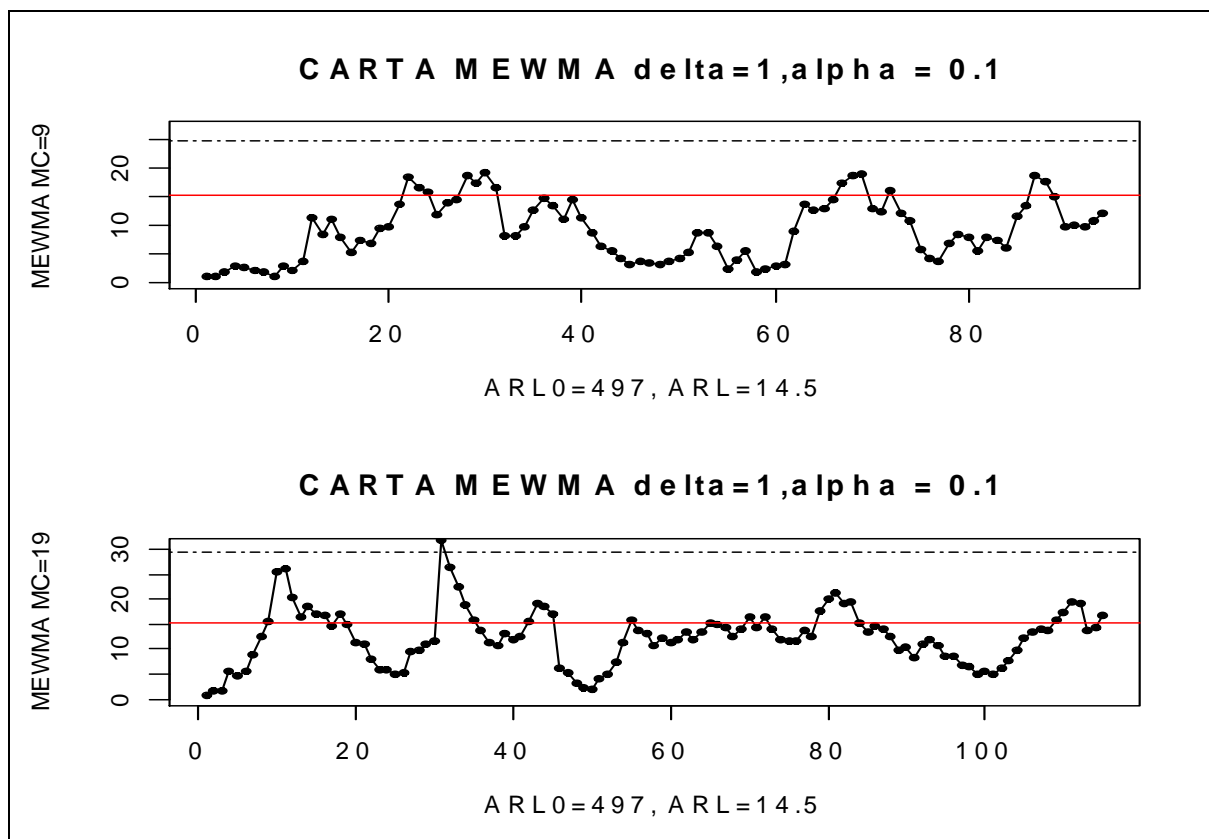
TABELLA 2

<b>p = 4 ; <math>\lambda = 0.1</math></b>				
<b><math>\delta</math></b>	<b>L=8,70</b>	<b>L=10,80</b>	<b>L=15,16</b>	<b>L=16,94</b>
0	50,7	101	497	995
0,5	18,8	26,1	52,3	68
1	8,62	10,4	14,5	16,5
1,5	5,54	6,42	8,2	8,99
2	4,16	4,7	5,79	6,23
2,5	3,31	3,74	4,51	4,81
3	2,81	3,12	3,74	3,97



### 5.5.7 Costruzione delle carte MEWMA nel caso studio

Di seguito vengono riportate le carte MEWMA per le due macchine di filatura, ottenute con la scelta di opportuni valori della coppia di parametri precedentemente specificati.



### 5.5.8 Interpretazione dei segnali fuori controllo

Una delle difficoltà che si presentano con le carte di controllo multivariate è costituita dall'interpretazione di una segnalazione di fuori controllo. Il problema principale si concentra nello stabilire quali delle  $p$  variabili hanno provocato tale situazione. Una soluzione pratica è quella di rappresentare graficamente le carte univariate nel monitoraggio di ogni singola variabile oppure Alt (1985) suggerisce di adottare una soluzione variando i limiti di controllo (Bonferroni). Ciò porta a ridurre il numero di falsi allarmi associati all'uso simultaneo di più carte univariate.

Un altro approccio molto utile nella diagnostica dei fuori controllo è quello di decomporre la statistica  $T^2$  in componenti che riflettano il contributo di ciascuna delle singole variabili (Mason e Young, 1995).

## 5.6 Caratteristiche e confronti

La carta MEWMA rispetto alle usuali carte Shewhart univariate (viste nel paragrafo 5.4.1) rileva più rapidamente le deviazioni con  $\delta < 3$  (scostamenti standard) e tale scostamento rende lo schema a media mobile efficiente (Lowry et al. 1992). Le carte MEWMA hanno la caratteristica di attenuare le informazioni che provengono dalle osservazioni più recenti, dal momento che un certo “peso” viene destinato alle informazioni passate; tale fatto è noto come “problema dell'inerzia”.

Una delle caratteristiche principali tra le due carte multivariate, analizzate nell'elaborato, consiste nella violazione delle ipotesi di normalità dei dati. La distribuzione  $T^2$  è sensibile alla normalità e con essa all'indice di Skewness (non alla kurtosis) e il problema peggiora al crescere di  $n$ .

Altro svantaggio della distribuzione  $T^2$  consiste nella difficoltà (specie quando il numero di variabili sotto controllo è elevato  $n > 10$ ) di interpretare i segnali di fuori controllo.

Tuttavia ha il vantaggio di essere costruita con un metodo di validità generale fondato su una statistica semplice e di avere una valutazione rigorosa della probabilità di falso allarme.

Al contrario della carta  $T^2$ , la carta MEWMA è ragionevolmente robusta all'ipotesi di normalità purché  $\lambda$  sia abbastanza piccolo ma di contro è poco adatta se il processo in controllo ha cicli o trend e non è molto efficiente nella segnalazione di discontinuità isolate. In conclusione l'utilizzo della carta MEWMA viene consigliato nei casi in cui si vogliono rilevare piccoli shift della media.

# Conclusioni

I risultati ottenuti nell'analisi statistica hanno evidenziato alcuni aspetti generalmente presenti in molte realtà produttive. Primo tra tutti la presenza e l'incidenza di valori anomali che rendono difficile e fuorviante l'interpretazione dei grafici e dei test statistici usualmente utilizzati (tra cui la verifica della normalità in distribuzione).

L'inclusione o l'eliminazione di tali valori dev'essere preventivamente concordata con il responsabile di produzione ed il management.

Un altro aspetto emerso e non trascurabile è la considerazione di distribuzione normale dei dati anche quando i test applicati rendono tale scelta discutibile ma per alcuni versi necessaria. Nel nostro caso con l'utilizzo delle consuete trasformazioni i risultati ottenuti nei test non portano a migliorare la situazione in distribuzione. In tal senso un'utile soluzione al problema dev'essere svolta preventivamente nello studio del controllo statistico del processo ed in particolare nelle procedure di campionamento. L'analisi di un corretto ed adeguato piano di campionamento può di sicuro ridurre, se non eliminare, le problematiche riguardanti la distribuzione dei dati.

Un altro aspetto da approfondire emerso con l'analisi statistica è la presenza di una relazione tra le caratteristiche esaminate ed il colore del prodotto. I test applicati nei due raggruppamenti non evidenziano particolarmente forti relazioni, tuttavia si consiglia di ripetere l'analisi con una numerosità maggiore per poter avanzare ipotesi più precise.



# Appendice

## A: Variabile casuale di Wishart

La variabile casuale di Wishart (in onore a John Wishart), è una delle variabili casuali per matrici non-negative di variabili casuali (matrici stocastiche). Queste distribuzioni sono di grande importanza per la stima delle matrici di covarianza nell'ambito della statistica multivariata.

### A1 - Definizione della distribuzione di Wishart

La variabile casuale di Wishart viene definita come segue. Posto  $X_1 \sim N_p(0, V)$  dove  $X_1$  è un vettore colonna  $p \times 1$  di variabili casuali (vettore stocastico) distribuite come una variabile casuale normale, il cui valore atteso è un vettore colonna  $p \times 1$  con tutti i valori pari a zero e la cui varianza è una matrice non negativa  $V p \times p$ . Si ottiene così

$$E[X_1] = 0$$

$$e: \text{var}[X_1] = E[(X_1 - \mu)(X_1 - \mu)'] = E[X_1 X_1'] = V$$

dove la trasposta di una matrice  $A$  viene indicata con  $A'$ .

Posto inoltre che:  $X_1, X_2, \dots, X_n$  siano indipendenti e identicamente distribuite, una variabile casuale di Wishart è una matrice stocastica  $p \times p$

$$S = \sum_{i=1}^n X_i X_i'$$

dove  $S$  ha tale distribuzione di probabilità scrivendo  $S \sim W_p(V, n)$

L'intero positivo  $n$  corrisponde ai gradi di libertà. Se  $p = 1$  e  $V = 1$  allora questa è una variabile casuale Chi quadro. Un noto statistico (George Seber) fa notare che la distribuzione di Wishart non è chiamata "Chi quadrato multivariata" in quanto la distribuzione marginale degli elementi non diagonali non sono distribuiti come una Chi quadrato. Seber preferisce riservare il termine "multivariata" per i casi in cui tutti i marginali univariati sono della stessa famiglia.

### A2 - Stimatore della distribuzione normale multivariata

La v.c. di Wishart è la variabile casuale dello stimatore di massima verosimiglianza della matrice delle covarianze di una variabile casuale gaussiana multivariata. Tale derivazione è sorprendentemente sottile e elegante. Essa coinvolge, da una parte, il teorema spettrale e, dall'altra, la ragione per la quale può essere meglio interpretare uno scalare come la traccia di una matrice  $1 \times 1$  piuttosto che come un semplice scalare.

## B: Test di normalità

### Shapiro test

Molto spesso lo statistico è chiamato a verificare se i dati raccolti provengono o meno da una popolazione normale. Il test di Shapiro-Wilk è considerato uno dei test più potenti per la verifica della normalità, soprattutto per piccoli campioni. La verifica della normalità avviene confrontando due stimatori alternativi della varianza  $S^2$ :

uno stimatore non parametrico basato sulla combinazione lineare ottimale della statistica d'ordine di una variabile aleatoria normale al numeratore, e il consueto stimatore parametrico, ossia la varianza campionaria, al denominatore. I pesi per la combinazione lineare ( $a_i$ ) sono disponibili su apposite tavole. La statistica  $W$  può essere interpretata come il quadrato del coefficiente di correlazione in un diagramma quantile-quantile (QQ plot).

$$W = \frac{\left( \sum_{i=1}^n a_i x_i \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Il comando per effettuare il test di normalità in questione in ambiente R è `shapiro.test()`: esso restituisce come risultato il valore della statistica  $W$  e il relativo p-value.

Di seguito viene riportata la tabella riassuntiva dei p-value di altri test applicati per la verifica di normalità dei dati in esame; in grassetto si possono notare i valori di accettazione del test.

TABELLA TEST NORMALITÀ

TEST	Titolo	Tenacita	Allungamento	Estraibili
1) <code>sf.test()</code> effettua il test di Shapiro-Francia	0,0032	0,0264	<b>0,078</b>	0,002
2) <code>ad.test()</code> effettua il test di Anderson-Darling	0,0002	0,032	0,041	0,0025
3) <code>cvm.test()</code> effettua il test di Cramer-Von Mises	0,0002	<b>0,0984</b>	0,0418	0,00249
4) <code>lillie.test()</code> effettua il test di Lilliefors	0	0,0062	0,0064	0,00045
5) <code>pearson.test()</code> effettua il test del chi-quadro di Pearson	0	<b>0,327</b>	0,0034	0

## C: Trasformazione dei dati

Tra le varie funzioni che si possono trovare in R per le trasformazioni di dati vi è quella di BoxCox che è così definita:

$$y^\lambda = \begin{cases} (y^\lambda - 1) / \lambda & \text{se } y \neq 0 \\ \log(y) & \text{se } y = 0 \end{cases}$$

dove  $y$  è un vettore di  $n$  osservazioni strettamente positive e  $\lambda \in [-2, 2]$ .

Tale funzione equivale alla famiglia esponenziale ed il parametro  $\lambda$  include nel suo intervallo i vari casi possibili di trasformate: logaritmica, inversa e radice quadrata.

Nel pacchetto in R "MASS" si può trovare la funzione `boxcox()` che consente, attraverso l'uso di sequenze iterative, di esplicitare il valore  $\lambda$  che massimizza la funzione soprascritta.

Nella tabella sottostante vengono riportati i valori dei p-value con il test di Shapiro applicati alla trasformata:

### Shapiro Test

Trasformate	Titolo	Tenacità	Estraibili
$\lambda = -2$	0,0255	0	0
$\lambda = -1.5$	0,0232	0	0
$\lambda = -1$	0,019	0,00047	0
$\lambda = -0.5$	0,014	0,0021	0,0028
$\lambda = 0,5$	0,0056	0,0335	0,021
$\lambda = 1,5$	0,0016	0,042	0
$\lambda = 2$	0,00075	0,0356	0
Log	0,0093	0,0073	0,0226
no trasf.	0,0031	0,03357	0,0023

## D: Verifiche d'ipotesi sulla media e varianza

### D1 - Verifica di ipotesi sulla media

La verifica di ipotesi sulla media della popolazione da cui proviene un campione e nel caso in cui la varianza della popolazione sia sconosciuta viene effettuato tramite il test t di student. Tale test consente di verificare molte ipotesi:

- se la media della popolazione da cui proviene un campione è pari ad un valore prefissato
- se la media delle popolazioni da cui sono estratti due campioni sono uguali tra di loro
- se un certo trattamento effettuato su un campione ne ha modificato la media.

La quantità  $t$  è il valore osservato della statistica test;  $df$  sono i gradi di libertà della distribuzione della statistica test sotto  $H_0$  (ipotesi nulla),  $p$ -value è il livello di significatività osservato.

Il comando `t.test` in R consente inoltre di specificare il valore dell'ipotesi nulla ed il livello di significatività del test. Di seguito si possono osservare i valori dell'output di R nel caso studio:

```
>t.test(titolo[MC==9],titolo[MC==19])
Welch Two Sample t-test
data: titolo[MC == 9] and titolo[MC == 19]
t = -4.4728, df = 178.284, p-value = 1.372e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.07218077 -0.02798721
sample estimates:
mean of x mean of y
2.222021 2.272105
```

### D2 - Verifica d'ipotesi per la varianza

Avendo a disposizione 2 campioni provenienti da 2 popolazioni diverse, ci si pone il problema di verificare se le varianze delle popolazioni da cui provengono sono tra di loro uguali. Tale test in R viene identificato con il comando `var.test`. ed è anch'esso basato su un test del rapporto di verosimiglianza nella verifica delle ipotesi. Anche per questo test è possibile cambiare il livello di significatività del test ed implementare l'ipotesi alternativa da bilaterale ad unilaterale.

```
> var.test(tenacita[MC==9],tenacita[MC==19])
F test to compare two variances
data: tenacita[MC == 9] and tenacita[MC == 19]
F = 0.9183, num df = 93, denom df = 114, p-value = 0.6722
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6244564 1.3608351
sample estimates:
ratio of variances
 0.9182539
```

Nel caso in cui si volesse applicare un test sulla uguaglianza delle varianze di più di due popolazioni, bisogna utilizzare il Bertlett.test il cui comando è `bartlett.test`.

Tale test, assumendo che la distribuzione della variabile risposta nei gruppi sia normale (con medie e varianze non necessariamente uguali) consente, attraverso il rapporto di verosimiglianza, la verifica dell'ipotesi nulla di omoschedasticità, ossia di uguaglianza delle varianze.



```
>bartlett.test(titolo,CLASSE)
Bartlett test of homogeneity of variances
data: titolo and CLASSE
Bartlett's K-squared = 13.3784, df = 2
```

### D3 - Verifica dell'ipotesi di uguaglianza della media: caso del confronto globale

Avendo a disposizione  $k$  campioni provenienti da  $K$  popolazioni diverse ci si pone il problema di verificare se le medie delle popolazioni da cui provengono tali campioni siano tra di loro uguali, effettuando un test globale tra i  $k$  campioni considerati.

Si tratta di un test noto come **analisi della varianza**.

Tale test ha il seguente comando:

```
>aaa<-oav(y~g) oppure >aaa<-aov(y~g,data=...)
>summary(aaa)
```

in cui si avrà che:

- il vettore  $y$  rappresenta le osservazioni effettuate
- il vettore  $g$  è un factor che rappresenta il campione di appartenenza.

Il test si basa su una scomposizione della varianza del modello impostato e la quantità indicata con  $\Pr(>F)$  coincide con il p-value del test F (distribuzione F di Snedecor) .

### D4 - Verifica di ipotesi su variabili con distribuzione libera

Nel caso in cui si presuppone di non conoscere o di non avere alcuna informazione circa la popolazione di origine dei campioni, la verifica di ipotesi sulla media della popolazione da cui proviene un campione è effettuato tramite il test di Wilcoxon. Tale test consente di verificare molte ipotesi:

- se la media della popolazione da cui proviene un campione è pari ad un valore prefissato
- se la media delle popolazioni da cui sono estratti due campioni sono uguali tra di loro
- se un certo trattamento effettuato su un campione ne ha modificato la media

la sintassi è:

```
>wilcox.test(x,y=null,alt="two.sided",mu=0,paired=F, var.equal=T,conf.level=0.95)
```

### Verifica dell'ipotesi di uguaglianza della media: caso del confronto globale

Quando si hanno a disposizione  $k$  campioni provenienti da  $k$  popolazioni diverse e ci si pone il problema di verificare se le medie delle popolazioni da cui provengono tali campioni siano tra di loro uguali senza però conoscerne la distribuzione viene utilizzato il test

```
>kruskal.test(x, g, ...)
oppure >kruskal.test(formula, data, subset,na.action,...)
```

in cui si avrà che:

- il vettore  $y$  rappresenta le osservazioni effettuate
- il vettore  $g$  è un factor che rappresenta il campione di appartenenza
- il valore di formula rappresenta la formula intesa nel seguente modo  $y\sim g$

### Test sulla uguaglianza delle varianze varianza di più popolazioni

Avendo a disposizione  $k$  campioni provenienti da  $k$  popolazioni diverse ci si pone il problema di verificare se le varianze delle popolazioni da cui provengono tali campioni siano tra di loro uguali. Ciò può essere svolto con il test di Fligner. Tale test ha il seguente comando:

`>fligner.test(y, g, ...)`

oppure il seguente comando:

`>fligner.test(formula, data, subset, na.action, ...)`

in cui:

- il vettore `y` rappresenta le osservazioni effettuate
- il vettore `g` è un factor che rappresenta il campione di appartenenza
- il valore di `formula` rappresenta la formula intesa nel seguente modo `y~g`
- il test può essere applicato anche nel caso di solo due popolazioni.

## E: Funzioni utilizzate in R

### E1 - Grafici per analisi preliminare con Shapiro test:

```
f<-function(dati) #misure singole (vettore){
par(mfrow=c(2,2))
hist(dati,nclass=10,prob=T)
lines(density(dati),col=2)
qqnorm(dati)
qqline(dati,col=2)
boxplot(dati)
plot(dati)
shapiro.test(dati)
}
```

### E2 - CARTA SHEWHART MISURE SINGOLE

```
"carta.mis.sin"<-function(dati,L,da="dati",tit="Carta Shewhart per le singole misure-stima di
sigma basata su MR(2)"){
"wsing"<-function(dati) #leggi osservazioni con n=1
{
MR2<-abs(diff(dati)) #calcola le m escursioni mobili
MR2
}
"vaMR2"<-function(dati) {
wi<-wsing(dati)
MRmedio<-mean(wi) #calcola MR medio
cat("MRmedio = ", round(MRmedio,digits=4), "\n")
d2<-c(1.128,1.693,2.059,2.326,2.534,2.70,2.847,2.970,3.078)
d3<-c(0.853,0.888,0.880,0.864,0.848,0.833,0.820,0.808,0.80)
stimaMR<-MRmedio/d2[1] #stima di sigma basata su MR(2)
cat("stimaMR = ", round(stimaMR,digits=4), "\n")
vaMR<-d2[1]*stimaMR #stima del valore atteso di MR(2)
cat("vaMR = ", round(vaMR,digits=4), "\n")
sigmaMR<-stimaMR*d3[1] #stima della deviazione standard di MR(2)
cat("sigmaMR = ", round(sigmaMR,digits=4), "\n")
risMR<-round(c(MRmedio,vaMR,sigmaMR,stimaMR),digits=3)
}
muw<-mean(dati)
sqmw<-vaMR2(dati)[4]
LCL<-muw-(L*sqmw)
LC<-muw
UCL<-muw +(L*sqmw)
cat("LCL = ", round(LCL,digits=4), "\n")
cat("LC = ", round(LC,digits=4), "\n")
cat("UCL = ", round(UCL,digits=4), "\n")
plot(dati,ylim=c(LCL,UCL),ylab=da,main=tit,type="l")
points(dati,pch=20)
abline(a=LCL,b=0,col="red")
abline(a=LC,b=0,col="blue")
abline(a=UCL,b=0,col="red")
}
```

### E3 - CARTA SULLE ESCURSIONI MOBILI

```
"carta.mis.sing.var"<-function(dati,L,da="",tit="Carta sulle escursioni mobili")
{
  "wsing"<-function(dati) #leggi osservazioni con n=1
  {
    MR2<-abs(diff(dati)) #calcola le m escursioni mobili
    MR2
  }
  "vaMR2"<-function(dati)
  {
    wi<-wsing(dati)
    MRmedio<-mean(wi) #calcola MR medio
    cat("MRmedio = ", round(MRmedio,digits=4), "\n")
    d2<-c(1.128,1.693,2.059,2.326,2.534,2.70,2.847,2.970,3.078)
    d3<-c(0.853,0.888,0.880,0.864,0.848,0.833,0.820,0.808,0.80)
    stimaMR<-MRmedio/d2[1] #stima di sigma basata su MR(2)
    cat("stimaMR = ", round(stimaMR,digits=4), "\n")
    vaMR<-d2[1]*stimaMR #stima del valore atteso di MR(2)
    cat("vaMR = ", round(vaMR,digits=4), "\n")
    sigmaMR<-stimaMR*d3[1] #stima della deviazione standard di MR(2)
    cat("sigmaMR = ", round(sigmaMR,digits=4), "\n")
    risMR<-round(c(MRmedio,vaMR,sigmaMR,stimaMR),digits=3)
  }
  stat<-wsing(dati) #carica le escursioni mobili MR(2)i
  muw<-vaMR2(dati)[1] #carica l'MR medio
  sqmw<-vaMR2(dati)[3] #deviazione standard di MR(2)
  LCL<-muw-L*sqmw
  if (LCL < 0) LCL<-0
  LC<-muw
  UCL<-muw+L*sqmw
  plot(stat,ylim=c(LCL,UCL),main=tit,type="l",ylab=da)
  cat("LCL = ", round(LCL,digits=4), "\n")
  cat("LC = ", round(LC,digits=4), "\n")
  cat("UCL = ", round(UCL,digits=4), "\n")
  points(stat,pch=20)
  abline(a=LCL,b=0,col="red")
  abline(a=LC,b=0,col="blue")
  abline(a=UCL,b=0,col="red")
}
```

### E4 - CARTE EWMA

```
"cartaEWMA"<-function(dati,la,L,n,tit="CARTA EWMA",tito="dati")
{
  media<-mean(dati)
  sd<-sd(dati)
  "ewma"<-function(dati,la)
  {
    k<-length(dati)
    x<-c(rep(0,k))
```

```

for (t in 1:length(dati))
{
if (t==1)
x[1]<-la*dati[1]+(1-la)*media
else
x[t]=la*dati[t]+(1-la)*x[t-1]
}
t<-t+1
x
}
"limiti"<-function(dati,la,L,n)
{
k<-length(dati)
sqm<-sd/sqrt(n)
lst<-c(rep(0,k))
lit<-c(rep(0,k))
for (t in 1:k)
{
lit[t]<-media-(L*sqm*sqrt((la/(2-la))*(1-(1-la)^(2*t))))
lst[t]<-media+(L*sqm*sqrt((la/(2-la))*(1-(1-la)^(2*t))))
}
t<-t+1
li<-media-(L*sqm*sqrt((la/(2-la))))
ls<-media+(L*sqm*sqrt((la/(2-la))))
rbind(cbind(lit,lst),c(li,ls))
}
wewma<-ewma(dati,la)
lim<-limiti(dati,la,L,n)
lim2<-lim[1:length(dati),]
matplot(cbind(wewma,lim2),type="l",lty=1:3,main=tit,ylab=tito)
points(wewma)
abline(a=media,b=0)
}

```

## E5 - CARTA T<sup>2</sup> MULTIVARIATA PER VALORI SINGOLI

Funzione di calcolo della statistica T<sup>2</sup> Hotelling nella **fase 1**, costruzione della carta.

Numerosità campione = 1; Dati = matrice m x p ; m = campioni, p = variabili.

NB: I dati devono ritenersi idonei per un processo in controllo.

```
T2_fase1<-function(dati,p,m,alpha=.01,ti="T",tit="carta T^2 Hotelling : fase 1"){
```

```
S<-matrix(var(dati),nrow=p,byrow=T)
```

```
for(i in 1: m){
```

```
diff<-matrix(t(dati[i,]-apply(dati,MARGIN=2,FUN=mean)))
```

```
T[i]<-t(diff)%*%solve(S)%*%(diff)
```

```
}
```

```
UCL<-(p*(m-1)/(m-p))*qf(1-alpha,p,m-p)
```

```
plot(T,type="l",lty=1:3,ylim=c(0,UCL+1),main=tit,ylab=ti)
```

```
points(T,pch=20)
```

```
abline(a=UCL,b=0,col=2)
```

```
DA<-list("statistica T2 Hotelling"=T,"limite superiore UCL"=UCL)
```

```
return(DA)}
```

Costruzione della carta nella **fase 2**

```
#dati completi = dati2; dati in controllo = dati 1
T2_fase2<-function(dati1,dati2,p,m,alpha=.01,ti="T",tit="carta T^2 Hotelling : fase 2"){
S<-matrix(var(dati1),nrow=p,byrow=T)
for(i in 1: m){
diff<-matrix(t(dati2[i,]-apply(dati1,MARGIN=2,FUN=mean)))
T[i]<-t(diff)%*%solve(S)%*%(diff)
}
UCL<-((p*(m+1)*(m-1))/(m^2-m*p))*qf(1-alpha,p,m-p)
plot(T,type="l",lty=1:3,ylim=c(0,UCL+1),main=tit,ylab=ti)
points(T,pch=20)
abline(a=UCL,b=0,col=2)
DA<-list("statistica T2 Hotelling"=T,"limite superiore UCL"=UCL)
return(DA)
}
```

**E6 - CARTA DI CONTROLLO MEWMA**

```
cartaMEWMA<-function(dati,n,lambda,UCL,ti="MEWMA",titolo="CARTA MEWMA"){
S<-var(dati)
mewma<-function(dati)
{
la<-lambda
x<-c(rep(0,n))
for (t in 1:length(dati))
{
if (t==1)
x[1]<-la*dati[1]+(1-la)*0
else
x[t]=la*dati[t]+(1-la)*x[t-1]
}
t<-t+1
x
}
Z<-apply(dati,MARGIN=2,FUN=mewma)
M<-c(rep(0,n))
for(i in 1:n){
laa[i]<-(lambda/(2-lambda))*(1-(1-lambda)^(2*i))
M[i]<-t(Z[i,])%*%solve(laa[i]*S)%*%Z[i,]
}
M
#UCL<-c(8.7,10.8,15.16,16.94) valori di riferimento
#UCL<-mean(M)+3*sd(M) valori di esempio
plot(M,type="l",lty=1:3,ylim=c(0,UCL+1),main=titolo,ylab=ti)
points(M,pch=20)
abline(a=UCL,b=0,col=2)
li<-list("statistica MEWMA"=M,"limite superiore UCL"=UCL)
return(li)
}
```

---

# Bibliografia

- Douglas C Montgomery (2000) – “*Controllo statistico della qualità*”, McGraw Hill.
- D.C. Montgomery (1996) – “*Introduction To Statistical Quality Control*”, 3<sup>a</sup> edizione By J. Wiley & Sons.
- Amor Messaoud, Claus Weihs and Franz Hering – “*A Nonparametric Multivariate Control Chart Based on Data Depth*”, Department of Statistics, University of Dortmund, Germany.
- Mason R.L., Chou Y.M. and Young J.C.(October 2001) – “*Applying Hotelling’s  $T^2$  Statistic to Batch Processes*”, Journal of Quality Technology, Vol.33, No.4.
- UNI EN (2000) – “*Sistemi di gestione della qualità, requisiti*”, normativa UNI EN ISO 9001/Vision 2000.
- S.L.R. Ellison (LGC, UK), M. Rosslein (EMPA, Switzerland), A. Williams (UK) – “*Quantifying uncertainty in analytical measurement*”, EURACHEM / CITAC Guide CG 4. Second edition (2000), Rapporti ISTISAN 03/30, Istituto Superiore di Sanità.
- Stefano Polesello (Ottobre 2002) – “*La validazione dei metodi analitici*”, articolo tratto dal convegno: “Le tre giornate del controllo ambientale”, Morgan Edizioni Tecniche, rivista “Laboratorio 2000”.
- Mario Bona (1993) – “*Statistical methods for the textile industry*”, Texilia, Istituto per la Tradizione e la Tecnologia Tessile S.p.A.
- Camil Fuchs & S.Kenett (1998) – “*Multivariate Quality Control: theory and applications*”, New York, Dekker M.
- Keats, Montgomery (1991) – “*Statistical Process Control in Manufacturing*”, New York, Dekker M, Quality and reliability.
- Mason R., Tracy N. & Young J. (1995) – “*Decomposition of  $T^2$  for Multivariate Control Chart Interpretation*”. Journal of Quality Technology, Vol. 27, No. 2, 99-108.
- Andreoli C., Fretti F. (2004)– “*Le fibre chimiche*”, quaderni di tecnologia tessile, fondazione ACIMIT.
- ACSA stab. Pto.Marghera (1966) – “*Misure delle proprietà fisico-meccaniche*”.
- ISO/IEC 17025 – “*General requirements for the competence of calibration and testing laboratories*”, Geneva:International Organization for Standardization; 1999. (Trad. it.: UNI CEI EN ISO/IEC 17025. *Requisiti generali per la competenza dei laboratori di prova e di taratura*. Milano: Ente Nazionale Italiano e di Unificazione; 2000).

- International Organization for Standardization.- “*Guide to the expression of uncertainty in measurement*”. Geneva: ISO; 1993. (Corretta e ristampata nel 1995). (Trad. it.: UNI CEI ENV 13005:2000. *Guida all’espressione dell’incertezza di misura*. Milano: Ente Nazionale Italiano di Unificazione; 2000).
- EURACHEM – “*Quantifying uncertainty in analytical measurement*”. London, Laboratory of the Government Chemist; 1995.
- BISFA (2000) – “*Testing methods for acrylic staple fibres*”, The International Bureau for the Standardisation of Man-Made Fibres.
- Lowry et al.- “A Multivariate Exponentially Weighed Moving Average Control Chart”, *Technometrics* 34 (1992, pag 46-53).
- Capizzi Giovanna - materiale del corso di “*Metodi Statistici di Controllo Qualità*” (2003/04).
- Leardini - materiale ed appunti del corso di “*Controllo Statistico della Qualità (certificazione)*” (2004/05).
- Stevanato R.: comunicazioni e documentazioni fornite da incontri di collaborazione, Montefibre P.to Marghera
- Lau Meng Cheng, Yuwaldi Away, Mohammad Khatim Hasan – “*The Algorithm And Design For Real-Time Hotelling’s T2 And Mewma Control Chart In Mspc*”, Department of Industrial Computing Faculty of Information Science and Technology National University of Malaysia.
- Barry M. Wise, N. Lawrence Ricker and David J. Veltkamp – “*Upset and sensor failure detection in multivariate processes*”, Center for Process Analytical Chemistry and Department of Chemical Engineering University of Washington Seattle.
- Peter R. Goulding, Barry Lennox, David J. Sandoz, Keith J. Smith and Ognjen Marjanovic - “*Fault detection in continuous processes using multivariate statistical methods*”, Control Technology Centre, School of Engineering, University of Manchester, Manchester, UK.

## **Tesi consultate**

- Marascalchi Martina (2002-2003) – “*Il controllo statistico di processo su dati multivariati*”.
- Varotto Enrico (2004-2005) – “*Cenni di controllo statistico della qualità*” .
- Iones Bertolo (1996-1997) – “*Metodi multivariati per il controllo di un processo produttivo*”.
- Lucio Carniello (1992-1993) – “*Carte di controllo multivariate*”.



# Ringraziamenti

A conclusione del corso di studi svolto in questi anni desidero esprimere la mia gratitudine a tutti coloro che mi hanno sostenuto ed in qualche modo incoraggiato nel raggiungimento di questo non facile traguardo.

Un ringraziamento va ai colleghi di lavoro ed in particolare al gruppo di reparto che con la loro disponibilità e collaborazione mi hanno permesso di frequentare i corsi e sostenere gli esami.

Più in generale vorrei ringraziare la società Montefibre che nella figura del dott. Zandonella e del tutor dott. Marinetti hanno reso possibile la realizzazione del presente elaborato.

Un sentito grazie alla Prof.ssa Capizzi ed agli amici dei corsi di studio, in particolare Goran.

Infine ma non per questo meno importante, vorrei ringraziare la mia famiglia che mi ha sempre sostenuto e Laura che mi ha incoraggiato ed aiutato nella stesura finale.