



# **UNIVERSITA' DEGLI STUDI DI PADOVA**

FACOLTA' DI SCIENZE STATISTICHE

CORSO DI LAUREA IN

STATISTICA E TECNOLOGIE INFORMATICHE

TESI DI LAUREA TRIENNALE

**SPC IN PRESENZA DI COMPONENTI DI VARIAZIONE GERARCHICA**

Relatore: Ch.ma Prof.ssa Giovanna Capizzi

Laureando: Mirko Zordan

ANNO ACCADEMICO: 2007/2008

## Sommario

CAPITOLO 1.....	3
INTRODUZIONE E FINALITA' DELLA TESI.....	3
1.1    INTRODUZIONE.....	3
1.2    FINALITA' DELLA TESI.....	5
CAPITOLO 2.....	7
ANALISI DELLA VARIANZA IN PRESENZA DI COMPONENTI GERARCHICHE DI VARIAZIONE.....	7
2.1    DESCRIZIONE DEL PROCESSO.....	7
2.2    MODELLI DI VARIAZIONE.....	8
2.2.1    Componenti di varianza annidate – 2 livelli di annidamento.....	8
2.2.2    Modelli con 3 e più livelli di annidamento.....	10
2.3    ESEMPIO – DUE LIVELLI DI ANNIDAMENTO.....	11
2.4    ESEMPIO – TRE LIVELLI DI ANNIDAMENTO.....	12
CAPITOLO 3.....	13
CARTE DI CONTROLLO PER IL MONITORAGGIO DELLA MEDIA E DELLE COMPONENTI DI VARIANZA.....	13
3.1    DESCRIZIONE DEL PROCESSO E MODELLO.....	13
3.2    COSTRUZIONE E DISEGNO DELLE CARTE DI CONTROLLO SHEWHART.....	14
3.2.1    Carta di controllo Shewhart per la media.....	14
3.2.2    Carta di controllo Shewhart per la componente di varianza $\sigma_E^2$ .....	15
3.2.3    Carta di controllo Shewhart per la componente di varianza $\sigma_B^2$ .....	15
3.2.4    Esempio di disegno delle carte Shewhart per il controllo della media e delle componenti di varianza.....	16
3.3    COSTRUZIONE E DISEGNO DELLE CARTE DI CONTROLLO CUSUM.....	19
3.3.1    Carta di controllo CUSUM per la media.....	19
3.3.2    Carta di controllo CUSUM per la componente di varianza $\sigma_E^2$ .....	20
3.3.3    Carta di controllo CUSUM per la componente di varianza $\sigma_B^2$ .....	20
3.3.4    Esempio di disegno delle carte CUSUM per il controllo della media e delle componenti di varianza.....	21
3.4    CARTA DI CONTROLLO EWMA PER LA MEDIA.....	24
CAPITOLO 4.....	27
CONFRONTO TRA I PROFILI ARL DELLE CARTE SHEWHART, CUSUM ED EWMA IN PRESENZA DI COMPONENTI GERARCHICHE DI VARIAZIONE.....	27
APPENDICE.....	29
ISTRUZIONI UTILIZZATE PER I GRAFICI DELLE CARTE NEGLI ESEMPI.....	29
FUNZIONI UTILIZZATE.....	30
BIBLIOGRAFIA.....	47



# CAPITOLO 1.

## INTRODUZIONE E FINALITA' DELLA TESI

### 1.1 INTRODUZIONE

#### *Definizione di qualità*

La qualità è diventata uno dei fattori più importanti nel processo di decisione del consumatore, quando si tratta di scegliere tra beni o servizi in concorrenza tra loro. Di conseguenza, comprendere e migliorare la qualità è un elemento decisivo per raggiungere il successo nell'impresa produttiva, per la crescita e il miglioramento della propria posizione nel mercato.

Il termine qualità è definito in molti modi. Alcuni interpretano la qualità in relazione a una o più caratteristiche che il prodotto o servizio deve possedere:

- Prestazione
- Affidabilità
- Durata
- Manutenibilità
- Aspetti formali (aspetto esteriore del prodotto)
- Funzionalità
- Livello di qualità percepito
- Conformità alle normative.

La definizione tradizionale del termine qualità si basa sul presupposto che beni e servizi devono soddisfare le richieste di coloro che li utilizzano:

#### **QUALITA' SIGNIFICA ESSERE APPROPRIATO PER L'USO.**

Ci sono due aspetti generali dell'essere appropriato per l'uso:

1. la qualità del progetto;
2. la conformità alle normative.

Tutti i beni e i servizi vengono prodotti con vari gradi o livelli di qualità. Queste diversità sono intenzionali. Ad esempio tutte le autovetture hanno come obiettivo basilare quello di essere un mezzo di trasporto sicuro per chi le acquista, tuttavia esse differiscono per dimensione, equipaggiamenti, modello e prestazioni. Queste differenze sono il risultato di progettazioni intenzionali che comprendono anche il tipo di materiali usati nella costruzione, le specifiche caratteristiche dei componenti, l'affidabilità ottenuta mediante lo sviluppo ingegneristico dei motori e di altri accessori ed equipaggiamenti.

La conformità alle normative è relativa a quanto il prodotto risulta conforme alle specifiche che sono richieste dal progetto. Essa è influenzata da parecchi fattori quali la scelta dei processi di lavorazione, la preparazione e supervisione della forza lavoro, il tipo di sistema utilizzato per l'accertamento della qualità, l'accuratezza con cui le procedure del sistema di accertamento della qualità vengono eseguite.



La definizione di qualità è stata associata più agli aspetti della conformità alle normative che a quelli della qualità del progetto portando a dare meno importanza al cliente e più a un approccio alla qualità in “conformità alle specifiche”. Vi è inoltre una convinzione diffusa che la qualità sia un problema che riguarda unicamente la fabbricazione.

Si preferisce dunque la più recente definizione del termine qualità:

### **LA QUALITA' E' INVERSAMENTE PROPORZIONALE ALLA VARIABILITA'.**

Questa definizione implica che la diminuzione della variabilità nelle caratteristiche importanti del prodotto implica un aumento della qualità del prodotto stesso. Questa definizione conduce anche alla seguente definizione del miglioramento della qualità:

### **IL MIGLIORAMENTO DELLA QUALITA' E' LA RIDUZIONE DELLA VARIABILITA' NEL PROCESSO PRODUTTIVO E NEL PRODOTTO.**

Molte organizzazioni trovano difficile e costoso fornire agli acquirenti prodotti che abbiano caratteristiche di qualità che siano sempre uguali in ciascun elemento prodotto, o che siano al massimo livello dell'aspettativa dell'acquirente. La causa principale di ciò è la variabilità, dalla cui presenza consegue che non ci sono mai due prodotti identici. Se la variazione è piccola può non avere nessuna importanza per l'acquirente, ma se la variazione è grande allora l'acquirente può ritenere che l'elemento sia indesiderabile e inaccettabile. Le origini di questa variabilità comprendono differenze nei materiali, diversità nelle prestazioni e nell'operatività degli strumenti di produzione e differenze nella modalità con cui gli operatori compiono il proprio lavoro.

Poiché la variabilità può essere descritta solamente in termini statistici, i metodi statistici hanno un ruolo centrale negli sforzi per il miglioramento della qualità. L'apparato di metodi e di tecniche statistiche atte allo scopo sono note con il nome di Controllo Statistico di Processo (SPC, Statistical Process Control).

#### ***Carte di controllo per variabili***

La carta di controllo è una delle principali tecniche adottate per il controllo statistico di un processo. Quando sono presenti fonti di variabilità anomale, queste vengono segnalate dalla carta, indicando la necessità di effettuare indagini per rimuovere e correggere gli eventuali errori intervenuti.

Le carte di controllo per caratteristiche del processo misurate su scala numerica (lunghezza, ampiezza, volume, ...) prendono il nome di carte di controllo per variabili. Di una variabile, in un contesto di SPC è solitamente necessario poter controllare sia la media sia la variabilità. Il monitoraggio può essere effettuato con carte di Shewhart (in riferimento alla metodologia sviluppata da W. A. Shewhart), carte CUSUM (CUMulative SUM, a somme cumulate) e carte EWMA (Exponentially Weighted Moving Average, a medie mobili pesate esponenzialmente), ognuna con possibili modifiche e opportuni accorgimenti a seconda delle necessità e caratteristiche del processo (vedi Montgomery, 2001).

## 1.2 FINALITA' DELLA TESI

Questa tesi si focalizza sui metodi per la sorveglianza di un processo la cui varianza totale può essere decomposta in specifiche componenti. Questo accade, ad esempio, nel caso della fabbricazione di circuiti integrati o nella produzione di wafer nelle industrie dei semiconduttori dove le differenti fonti di variabilità (tipicamente quelle lot-to-lot, wafer-to-wafer-within-lot e within wafer) sono associate a diverse cause assegnabili (speciali) identificabili e ineliminabili, per cui è essenziale separarle, stimarle e sorvegliarle separatamente. L'utilità di questo modo di procedere è la seguente: quando si hanno diverse cause che contribuiscono alla varianza totale del processo, l'uso di un'unica statistica che misuri la varianza totale del processo non è un approccio pratico ed efficiente in quanto alla presenza di un segnale di fuori controllo, che indica l'esistenza di una causa speciale, non si avrebbero indicazioni sulla natura di tali cause. In tale situazione l'operatore dovrebbe analizzare tutte le possibili cause speciali, il che sarebbe verosimilmente un lavoro costoso oltre che tedioso. Diversamente, se la varianza totale del processo può essere decomposta nelle sue componenti rilevanti e queste componenti vengono monitorate separatamente, l'operatore avrà maggiori indicazioni su dove agire per il ripristino delle condizioni operative standard.

Prima dell'implementazione delle carte di controllo bisogna in ogni caso sforzarsi di analizzare e ridurre il più possibile queste componenti di variazione. Solamente dopo questo passo infatti, nell'eventualità che risultino non eliminabili, è ragionevolmente giustificata l'inclusione di queste componenti di variazione in una carta di controllo.

Nel **capitolo 2** descriveremo una metodologia per la modellazione e la stima delle componenti di varianza.

Nel **capitolo 3** verranno presentate le carte Shewhart e CUSUM per processi con componenti gerarchiche di variazione (Chang e Gan, 2004 e Yashchin, 1994, rispettivamente). L'applicazione di queste carte sarà illustrata prendendo in esame dati reali provenienti da un processo di assemblaggio di moduli di memoria. In queste carte i limiti di controllo sono determinati con l'inclusione dell'inevitabile variabilità sistematica così come della variabilità naturale. Successivamente illustreremo il disegno di una carta EWMA da applicare in presenza di componenti di variazione gerarchica.

Nel **capitolo 4** attueremo uno studio di simulazione per esplorare l'efficienza delle tre carte.



## CAPITOLO 2

### ANALISI DELLA VARIANZA IN PRESENZA DI COMPONENTI GERARCHICHE DI VARIAZIONE

In questo capitolo, prendendo come esempio un processo di drogaggio del polisilicio attuato nelle industrie dei semiconduttori, illustreremo diversi modelli di variazione ed una metodologia per la stima delle componenti di varianza.

#### 2.1 DESCRIZIONE DEL PROCESSO

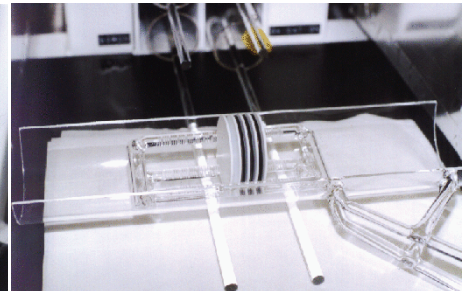
Il polisilicio è il silicio amorfo che una volta drogato viene largamente utilizzato per le sue proprietà di semiconduttore. Il polisilicio viene depositato e drogato in due successivi processi di diffusione. I wafer vengono caricati nei quartz boats (contenitori, *fig.1*) e messi in un forno tubolare (*fig.2*) dove una serie di gas vengono fatti reagire con i wafer. In questo modo vengono creati nuovi layers i quali vengono drogati o temperati. I processi di diffusione necessitano di alte temperature e lunghi tempi di reazione, ma molti wafer possono essere processati simultaneamente.



**Figura 1** - Quartz glass wafer boats for automatic loading systems in the semiconductor and solar industry



**Figura 2**



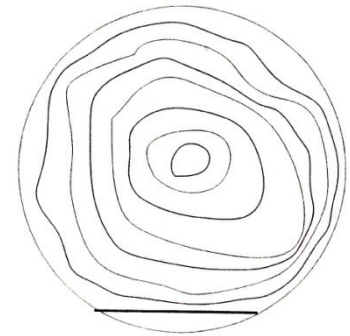
**Figura 3** - Wafer di polisilicio alternati a wafer dopanti in un quartz boat

Per la produzione del polisilicio i wafer vengono innanzi tutto ripuliti dalle impurità. Successivamente il polisilicio viene depositato in un processo di diffusione nel quale le più importanti variabili di processo che influiscono sulle caratteristiche cruciali del polisilicio non drogato (spessore del film e dimensioni della grana) sono la temperatura di reazione, la pressione e il flusso di gas.

Una volta che nei wafer è stato depositato il polisilicio, il drogaggio avviene tramite il riscaldamento dei wafer, i quali sono alternati ad altri wafer sorgenti del dopante (*fig.3*). Questi ultimi producono un gas dopante ad una temperatura predeterminata che si deposita nella grana dei wafer di polisilicio. In questo modo i wafer acquistano la resistività desiderata, caratteristica critica dell'output di questo processo. La resistività subisce l'influenza di diversi fattori: temperatura di reazione, dimensioni della grana, presenza di contaminanti e condizioni del tubo (forno) di diffusione.

La resistività spesso varia in maniera sistematica a seconda della posizione del wafer nel tubo di diffusione. In particolare la resistività dei wafer vicini alla sorgente del gas è diversa dalla resistività dei wafer vicini alla "porta". Per questa ragione la resistività è tipicamente misurata in diversi punti all'interno dello stesso "run" produttivo.

D'altra parte anche la resistività in uno stesso wafer non è omogenea. La "mappa" della resistività in un wafer è simile a quella raffigurata in figura 4.



**Figura 4** - "mappa" tipo della resistività in un wafer

## 2.2 MODELLI DI VARIAZIONE

Supponiamo che la caratteristica di interesse  $X$ , la resistività, abbia una media fissa con una varianza costante e che le misurazioni siano indipendenti una dall'altra. Un semplice modello di variazione è dato da:

$$X_i = \mu + \varepsilon_i \quad i = 1, \dots, n$$

dove  $X_i$  è l' $i$ -esima osservazione di processo, ogni  $\varepsilon_i$  è una osservazione indipendente di una variabile casuale normalmente distribuita con media 0 e varianza  $\sigma_E^2$ ,  $\mu$  è la media del processo e  $n$  è il numero di misurazioni nel campione.

In questo modello abbiamo una singola componente di varianza:  $\sigma_E^2$ . Nel caso in esame il processo presenta una struttura della varianza più complessa. La resistività infatti varia in maniera sistematica sia tra wafer diversi di uno stesso run di diffusione, sia tra punti diversi di uno stesso wafer.

Un modello realistico per la variazione del processo è basato su una struttura nidificata della varianza.

### 2.2.1 Componenti di varianza annidate - 2 livelli di annidamento

Una struttura della varianza alternativa è quella annidata. E' chiaro infatti che misurazioni multiple nello stesso wafer tenderanno ad essere simili una all'altra e per questo vengono dette annidate entro il wafer. Questo semplice tipo di annidamento è descritto dal seguente modello a due livelli:

$$X_{ij} = \mu + B_i + \varepsilon_{j(i)} \quad (2.1)$$

dove

- $B_i \sim N(0, \sigma_B^2)$ ;  $i = 1, \dots, r$
- $\varepsilon_{j(i)} \sim N(0, \sigma_E^2)$ ;  $j = 1, \dots, n$

$X_{ij}$  è la  $j$ -esima osservazione di  $X$  fatta per l' $i$ -esimo wafer,  $n$  è il numero di misurazioni in ogni wafer e  $r$  è il numero dei wafer sui quali vengono effettuate le misurazioni.

Questo è un modello annidato bilanciato in quanto il numero di misurazioni è lo stesso per ogni wafer.

Le componenti di varianza del modello (2.1) sono due:  $\sigma_B^2$  e  $\sigma_E^2$ . L'effetto dell'annidamento è che la varianza totale di una singola misurazione è data dalla somma delle componenti di varianza di ogni livello dell'annidamento:

$$\sigma_T^2 = \sigma_B^2 + \sigma_E^2$$

Per il nostro modello, che presenta una struttura annidata a due livelli e dove le misurazioni sono annidate nei wafer, la varianza della media di  $X$  per wafer è:

$$\sigma_{\bar{B}}^2 = \sigma_B^2 + \frac{\sigma_E^2}{n} \quad (2.2)$$

dove con  $\bar{B}$  si intende la media delle misurazioni per un particolare wafer.

La varianza delle medie di  $X$ , per wafer, può essere stimata da:

$$\hat{\sigma}_{\bar{B}}^2 = \frac{1}{(r-1)} \sum_{i=1}^r (\bar{X}_i - \bar{X}_{..})^2 \quad (2.3)$$

dove

- $\bar{X}_i$  rappresenta la media per il wafer  $i$ ,  $\bar{X}_i = \frac{1}{n} \sum_{j=1}^n X_{ij}$
- $\bar{X}_{..}$  rappresenta la media di tutte le osservazioni,  $\bar{X}_{..} = \frac{1}{rn} \sum_{i=1}^r \sum_{j=1}^n X_{ij}$

La componente di varianza  $\sigma_E^2$  può essere stimata mediando le stime delle varianze entro i wafer:

$$\hat{\sigma}_E^2 = \frac{1}{r} \sum_{i=1}^r \sum_{j=1}^n \frac{(X_{ij} - \bar{X}_i)^2}{n-1} \quad (2.4)$$

Riarrangiando l'equazione (2.2) si ottiene la stima della componente di varianza  $\sigma_B^2$ :

$$\hat{\sigma}_B^2 = \hat{\sigma}_{\bar{B}}^2 - \frac{\hat{\sigma}_E^2}{n} \quad (2.5)$$

### **Significatività delle componenti di varianza**

Un quesito da porsi durante la valutazione di un modello con componenti di varianza riguarda la significatività statistica di tali componenti. Per il modello (2.1), questo equivale a saggiare le seguenti ipotesi:

$$H_0: \sigma_B = 0$$

$$H_1: \sigma_B \neq 0$$

Se l'ipotesi nulla è rifiutata significa che i wafer sono una fonte significativa di variazione. Diversamente, se l'ipotesi nulla non è rifiutata, i wafer possono essere ignorati come fonte di variazione. Questi test d'ipotesi si basano sull'ANOVA (ANalysis Of VAriance, analisi della varianza).

### 2.2.2 Modelli con 3 e più livelli di annidamento

I modelli annidati possono avere più di due livelli di annidamento. Si può ad esempio ipotizzare che le misurazioni siano annidate entro i wafer, i quali sono a loro volta annidati entro i lotti, ovvero:

$$X_{ijk} = \mu + L_i + B_{j(i)} + \epsilon_{k(ij)} \quad (2.6)$$

dove

- $L_i \sim N(0, \sigma_L^2)$ ;  $i = 1, \dots, l$
- $B_{j(i)} \sim N(0, \sigma_B^2)$ ;  $j = 1, \dots, r$
- $\epsilon_{k(ij)} \sim N(0, \sigma_E^2)$ ;  $k = 1, \dots, n$

$\mu$  è la media del processo,  $l$  è il numero dei lotti misurati nel campione,  $r$  è il numero dei wafer sui quali vengono effettuate le misurazioni in ogni lotto ed  $n$  è il numero di misurazioni effettuate in ogni wafer.

Questo modello ha dunque tre componenti di varianza:  $\sigma_L^2$ ,  $\sigma_B^2$  e  $\sigma_E^2$ .

Come in tutti i modelli a componenti di variazione, l'effetto dell'annidamento è che la varianza totale di una singola misurazione è data dalla somma delle componenti di varianza di ogni livello dell'annidamento:

$$\sigma_T^2 = \sigma_L^2 + \sigma_B^2 + \sigma_E^2$$

La componente di varianza  $\sigma_E^2$  è stimata calcolando la varianza entro ogni wafer in ogni lotto e mediando queste  $lr$  stime:

$$\hat{\sigma}_E^2 = \frac{1}{lr} \sum_{i=1}^l \sum_{j=1}^r \sum_{k=1}^n \frac{(X_{ijk} - \bar{X}_{ij.})^2}{n-1} \quad (2.7)$$

Per ottenere la stima della componente di varianza  $\sigma_B^2$  si deve innanzi tutto stimare la varianza delle medie per wafer, calcolata con:

$$\hat{\sigma}_B^2 = \frac{1}{l} \sum_{i=1}^l \sum_{j=1}^r \frac{(\bar{X}_{ij.} - \bar{X}_{i..})^2}{r-1}$$

Ora, la componente di varianza dei wafer  $\sigma_B^2$  è stimata tramite:

$$\hat{\sigma}_B^2 = \hat{\sigma}_B^2 - \frac{\hat{\sigma}_E^2}{n} \quad (2.8)$$

Per ottenere la stima della componente di varianza  $\sigma_L^2$  si deve innanzi tutto stimare la varianza delle medie per lotto, calcolata con:

$$\hat{\sigma}_L^2 = \sum_{i=1}^l \frac{(\bar{X}_{i..} - \bar{X}_{...})^2}{l-1}$$

Ora, la componente di varianza dei lotti è stimata tramite:

$$\hat{\sigma}_L^2 = \hat{\sigma}_L^2 - \frac{\hat{\sigma}_B^2}{r} - \frac{\hat{\sigma}_E^2}{rn} \quad (2.9)$$

Per modelli con maggiori livelli di annidamento, il procedimento da adottare è lo stesso: ottenere una stima per la componente di varianza di livello più basso e usare questa per derivare quella della componente di livello più alto.

### 2.3 ESEMPIO - DUE LIVELLI DI ANNIDAMENTO

Tre misurazioni della resistività vengono effettuate in ognuno di sei wafer. La *figura 5* mostra graficamente il risultato di queste misurazioni.

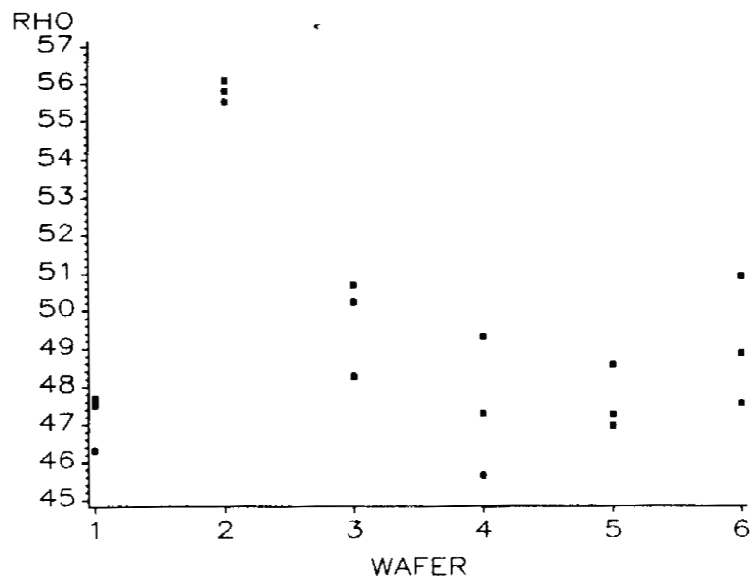


Figura 5

Si osserva che le misurazioni entro i wafer tendono ad essere più simili tra loro rispetto alle misurazioni sugli altri wafer. La tabella ANOVA sottostante mostra i risultati della stima delle componenti di varianza.

Variance Source	Degrees of Freedom	Sum of Squares	F Value	Pr > F	Error Term
TOTAL	17	178.499361			
WAFER	5	159.863028	20.5873	0.000017	ERROR
ERROR	12	18.636333			

Variance Source	Mean Square	Variance Component	Percent of Total
TOTAL	10.499962	11.692887	100.0000
WAFER	31.972606	10.139859	86.7182
ERROR	1.553028	1.553028	13.2818

Il p-value indica senza dubbio che i wafer sono una fonte di variazione statisticamente significativa.



## 2.4 ESEMPIO - TRE LIVELLI DI ANNIDAMENTO

Due misurazioni della resistività vengono effettuate in ogni wafer. I wafer selezionati casualmente da ogni lotto sono tre e i lotti presenti nel campione sono undici. La *figura 6* mostra graficamente il risultato delle misurazioni.

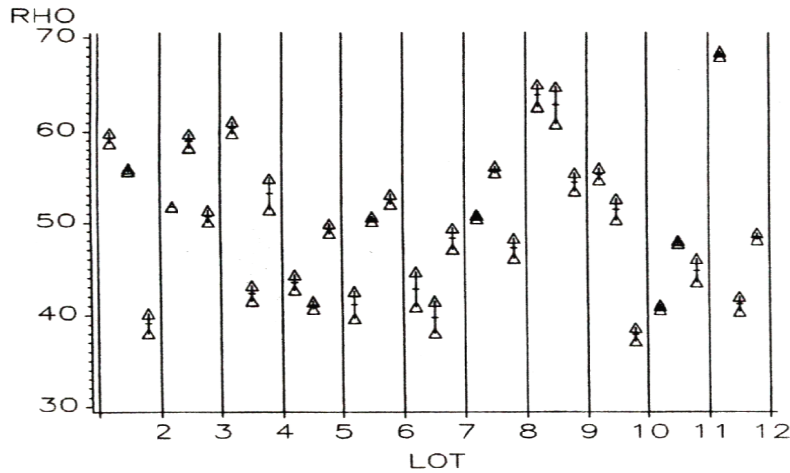


Figura 6

La tabella ANOVA sottostante mostra i risultati della stima delle componenti di varianza.

Variance Source	Degrees of Freedom	Sum of Squares	F Value	Pr > F	Error Term
TOTAL	65	4025.487062			
LOT	10	1453.333712	1.27299	0.303499	WAFER ERROR
WAFER ERROR	22	2511.673500	62.2936	0.000000	
ERROR	33	60.479850			

Variance Source	Mean Square	Variance Component	Percent of Total
TOTAL	61.930570	63.194249	100.0000
LOT	145.333371	5.194399	8.2197
WAFER	114.166977	56.167127	88.8801
ERROR	1.832723	1.832723	2.9000

I p-value indicano che dobbiamo considerare significative le componenti di varianza  $\sigma_L^2$  e  $\sigma_B^2$ .

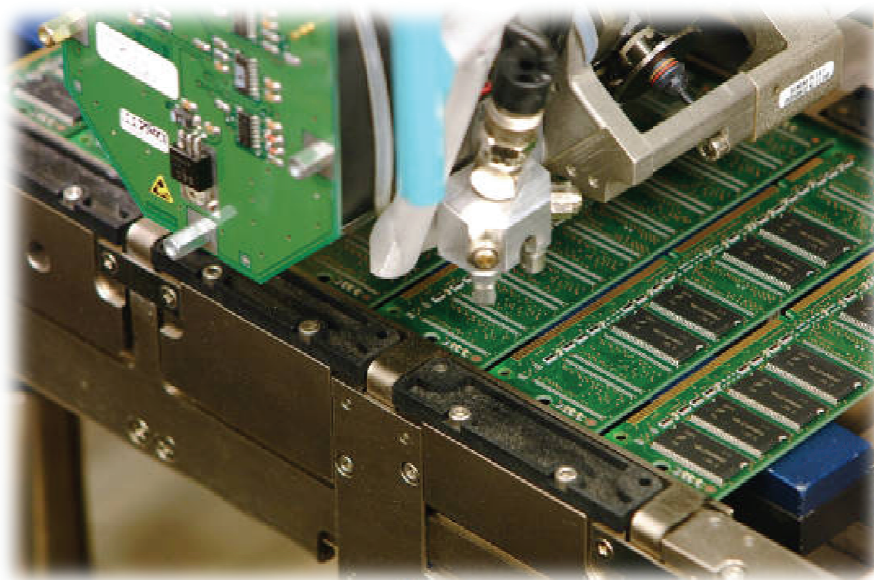
## CAPITOLO 3

### CARTE DI CONTROLLO PER IL MONITORAGGIO DELLA MEDIA E DELLE COMPONENTI DI VARIANZA

In questo capitolo verranno presentate le carte Shewhart, CUSUM ed Ewma per processi con componenti gerarchiche di variazione (Chang e Gan, 2004; Yashchin, 1994). In riferimento al modello (2.1), nella carta CUSUM per la sorveglianza di  $\sigma_B^2$ , Yashchin (1994) utilizza una approssimazione della distribuzione di  $\hat{\sigma}_B^2$ , mentre, nella carta Shewhart per la sorveglianza di  $\sigma_B^2$ , Chang e Gan (2004) ne propongono la distribuzione esatta. Al fine di rendere omogenea la trattazione le carte verranno applicate considerando il modello (2.1) e facendo uso dei risultati esatti sulla distribuzione di  $\hat{\sigma}_B^2$  forniti da Chang e Gan. I risultati approssimati di Yashchin verranno utilizzati nel caso di più di due livelli di annidamento.

#### 3.1 DESCRIZIONE DEL PROCESSO E MODELLO

Nel processo di assemblaggio dei moduli di memoria che prendiamo come esempio, una lega è stampata sulla basetta per circuiti stampati (in seguito PCB, Printed Circuit Board) prima che i circuiti integrati (in seguito IC, Integrated Circuits) siano montati sulla stessa. In questo processo lo spessore della lega determina la saldabilità degli IC: uno spessore omogeneo della lega in una PCB è desiderabile. Il controllo di processo prevede che si effettuino  $n$  misurazioni dello spessore della lega in ognuna di  $r$  postazioni casuali e differenti .



**Figura 7** – PCB con otto banchi di memoria sui quali vengono saldati nove IC

Queste misurazioni possono essere rappresentate per mezzo di un modello annidato a due livelli ad effetti casuali:

$$X_{ij} = \mu + B_i + \varepsilon_{ij} \quad i = 1, 2, \dots, r; \quad j = 1, 2, \dots, n$$

dove:

- $\mu$  è la media del processo;
- $B_i$  è l'effetto casuale dell'  $i$ -esima postazione,  $B_i \sim N(0, \sigma_B^2)$ ;
- $\varepsilon_{ij}$  è l'effetto casuale della  $j$ -esima misurazione effettuata nell' $i$ -esima postazione,  $\varepsilon_{ij} \sim N(0, \sigma_E^2)$ .

### 3.2 COSTRUZIONE E DISEGNO DELLE CARTE DI CONTROLLO SHEWHART

In questo paragrafo verrà illustrata la costruzione delle carte di controllo Shewhart per la sorveglianza della media e delle componenti di varianza. Successivamente le carte verranno disegnate per una serie di campioni ottenuti tramite simulazione.

#### 3.2.1 Carta di controllo Shewhart per la media

Sia  $\bar{X}_{..}$  la media di tutte le osservazioni. In controllo  $E[\bar{X}_{..}] = \mu$ , parametro del processo da sottoporre in controllo, e

$$\begin{aligned} \text{Var}(\bar{X}_{..}) &= \text{Var}(\mu) + \text{Var}\left(\frac{1}{r} \sum_{i=1}^r B_i\right) + \text{Var}\left(\frac{1}{rn} \sum_{i=1}^r \sum_{j=1}^n \varepsilon_{ij}\right) \\ &= \frac{1}{r^2} r \sigma_B^2 + \frac{1}{r^2 n^2} rn \sigma_E^2 \\ &= \frac{1}{r} \sigma_B^2 + \frac{1}{rn} \sigma_E^2 \end{aligned} \quad (3.1)$$

Assumendo la normalità delle osservazioni  $X_{ij}$ , abbiamo:

$$\bullet \text{ UCL} = \mu + Z_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{\sigma_B^2}{r} + \frac{\sigma_E^2}{rn}\right)} \quad (3.2)$$

$$\bullet \text{ CL} = \mu \quad (3.3)$$

$$\bullet \text{ LCL} = \mu - Z_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{\sigma_B^2}{r} + \frac{\sigma_E^2}{rn}\right)} \quad (3.4)$$

Dove  $\mu$ ,  $\sigma_B^2$  e  $\sigma_E^2$  saranno stimati da campioni provenienti dal processo in controllo ed  $\alpha$  è un conveniente valore del tasso di falsi allarmi.

### 3.2.2 Carta di controllo Shewhart per la componente di varianza $\sigma_E^2$

Per la sorveglianza della componente di varianza  $\sigma_E^2$ , i limiti di controllo Shewhart vengono calcolati considerando la distribuzione del rapporto  $\frac{\hat{\sigma}_E^2}{\sigma_E^2}$ .

La variabile casuale  $\frac{r(n-1)\hat{\sigma}_E^2}{\sigma_E^2} \sim X_{r(n-1)}^2$ , dove  $X_v^2$  rappresenta la distribuzione chi-quadro con  $v$  gradi di libertà. I limiti di controllo per la carta Shewhart possono essere determinati facilmente e accuratamente per un dato tasso di falsi allarmi  $\alpha$  come segue:

$$\bullet \text{ UCL} = \frac{\sigma_E^2 X_{1-\left(\frac{\alpha}{2}\right); r(n-1)}^2}{r(n-1)} \quad (3.5)$$

$$\bullet \text{ CL} = \frac{\sigma_E^2 X_{0,5; r(n-1)}^2}{r(n-1)} \quad (3.6)$$

$$\bullet \text{ LCL} = \frac{\sigma_E^2 X_{\left(\frac{\alpha}{2}\right); r(n-1)}^2}{r(n-1)} \quad (3.7)$$

dove  $\sigma_E^2$  è stimata usando dati provenienti dal processo in controllo con l'equazione (2.4).

### 3.2.3 Carta di controllo Shewhart per la componente di varianza $\sigma_B^2$

La componente  $\hat{\sigma}_B^2$  può essere espressa come la differenza tra due variabili casuali indipendenti, ognuna proporzionale ad una variabile casuale chi-quadro. Dall'equazione (2.5) possiamo scrivere:

$$\hat{\sigma}_B^2 = S - \frac{T}{n}$$

$$\frac{(r-1)S}{\sigma_B^2} \sim X_{(r-1)}^2 \quad ; \quad \frac{r(n-1)T}{\sigma_E^2} \sim X_{r(n-1)}^2 \quad (3.8)$$

Ponendo  $u = \frac{(r-1)}{2}$  e  $v = \frac{r(n-1)}{2}$  le funzioni di densità di probabilità di  $S$  e  $T$  sono rappresentate da  $f_{\sigma^*,u}(s)$  e  $f_{\sigma,v}(t)$  rispettivamente, dove i deponenti  $\sigma$  e  $\sigma^*$  rappresentano la radice quadrata della vera varianza.  $S$  e  $T$  sono indipendenti e hanno funzioni di densità continue, dunque la loro funzione di densità congiunta è continua ed è data dal prodotto delle loro funzioni di densità. La funzione di densità  $Y = \hat{\sigma}_B^2$  si ottiene da:

$$\frac{d}{dy} P(Y \leq y) = \int_{x=-\infty}^{\infty} f_{\sigma^*,u}(x+y) f_{\sigma,v}(nx) dx.$$

Integrando la funzione di densità si ottiene la funzione di ripartizione di  $Y$ :

$$P(Y \leq y) = \int_{w=-\infty}^y \int_{x=-\infty}^{\infty} f_{\sigma^*,u}(x+w) f_{\sigma,v}(nx) dx dw.$$

La funzione di ripartizione permette di ricavare i vari percentili della distribuzione di  $\hat{\sigma}_B^2$ . Una stima accurata (somma dei quadrati degli errori compresa tra  $4 \times 10^{-7}$  e  $5 \times 10^{-4}$ ) di questi percentili, per la determinazione dei limiti di controllo, è ricavata mediante modelli di regressione. Chang e Gan (2004) forniscono diverse tabelle con i coefficienti di regressione per determinare i limiti di controllo per valori di  $\alpha \in \{0.001, 0.002, 0.005\}$  e per diversi valori di  $r$  e  $n$ . Questi modelli sono validi per un rapporto  $\frac{\sigma_B}{\sigma_E}$  compreso nel range (0,05-3).

La costruzione della carta Shewhart per la sorveglianza di  $\sigma_B^2$  si basa sui seguenti tre passi:

- 1) Determinare  $r$  e  $n$  e successivamente stimare il rapporto  $Z = \frac{\sigma_B}{\sigma_E}$  quando  $\sigma_B$  e  $\sigma_E$  sono in controllo;
- 2) Calcolare il limite di controllo superiore UCL, corrispondente al quantile  $(1 - \frac{\alpha}{2})$ , il limite di controllo inferiore LCL e la linea centrale CL attraverso:

$$UCL/LCL/CL = \sigma_E^2 (\beta_0 + \beta_1 z + \beta_2 z^2) \quad (3.9)$$

dove i coefficienti  $\beta_0$ ,  $\beta_1$  e  $\beta_2$  possono essere ottenuti dalle tabelle in Chang e Gan (2004).

Quando  $r$  è un numero pari i limiti di controllo esatti non possono essere ottenuti perché la formula esatta della funzione di ripartizione per  $\hat{\sigma}_B^2$  non può essere calcolata. In questi casi il limite di controllo superiore viene approssimato usando la media degli UCL basati su  $(r-1)$  e  $(r+1)$ . Lo stesso procedimento si applica per LCL e CL.

### 3.2.4 Esempio di disegno delle carte Shewhart per il controllo della media e delle componenti di varianza

Di seguito vengono disegnate le tre carte Shewhart per 50 campioni ottenuti tramite simulazione. Si suppone che, per le misurazioni dello spessore della lega nel processo descritto all'inizio del capitolo, siano selezionate  $r = 5$  postazioni casuali della bassetta e in ognuna di queste vengano effettuate  $n = 2$  misurazioni. Si suppone inoltre che, basandosi su una collezione di campioni provenienti dal processo in controllo,  $\mu$ ,  $\sigma_E$  e  $\sigma_b$  siano stimati pari a 40, 7.135 e 7.014 rispettivamente. Le funzioni per la simulazione dei campioni e per il disegno delle carte si trovano in appendice.

Per un tasso di falsi allarmi  $\alpha=0,005$  (ARL in controllo =  $1/\alpha = 200$ ) i limiti di controllo e la linea centrale della carta Shewhart per la media sono:

- UCL =  $\mu + Z_{1-\frac{\alpha}{2}} \sqrt{(\frac{\hat{\sigma}_B^2}{r} + \frac{\hat{\sigma}_E^2}{rn})} = 40 + 2.8 \sqrt{(\frac{7,014^2}{5} + \frac{7,135^2}{10})} = 50,8$
- CL =  $\mu = 40$
- LCL =  $\mu - Z_{1-\frac{\alpha}{2}} \sqrt{(\frac{\hat{\sigma}_B^2}{r} + \frac{\hat{\sigma}_E^2}{rn})} = 40 - 2.8 \sqrt{(\frac{7,014^2}{5} + \frac{7,135^2}{10})} = 29,2$

Per un tasso di falsi allarmi  $\alpha=0,005$  (ARL in controllo =  $1/\alpha = 200$ ) i limiti di controllo e la linea centrale della carta Shewhart per la componente  $\hat{\sigma}_E^2$  sono:

- $$UCL = \frac{\sigma_E^2 X_{1-\left(\frac{\alpha}{2}\right); r(n-1)}^2}{r(n-1)} = \frac{50,908 X_{0,9975; 5}^2}{5} = \frac{50,908 * 18,38561}{5} = 187,2$$
- $$CL = \frac{\sigma_E^2 X_{0,5; r(n-1)}^2}{r(n-1)} = \frac{50,908 X_{0,5; 5}^2}{5} = \frac{50,908 * 4,35146}{5} = 44,3$$
- $$LCL = \frac{\sigma_E^2 X_{\left(\frac{\alpha}{2}\right); r(n-1)}^2}{r(n-1)} = \frac{50,908 X_{0,0025; 5}^2}{5} = \frac{50,908 * 0,3074818}{5} = 3,1$$

Per un tasso di falsi allarmi  $\alpha=0,005$  (ARL in controllo =  $1/\alpha = 200$ ) la determinazione dei limiti di controllo e della linea centrale della carta Shewhart per la componente  $\hat{\sigma}_B^2$  avviene seguendo i tre passi:

1) Si determinano:

- $r = 5;$
- $n = 2;$
- $z = \frac{\sigma_B}{\sigma_E} = \frac{7,014}{7,135} = 0,983$

2)

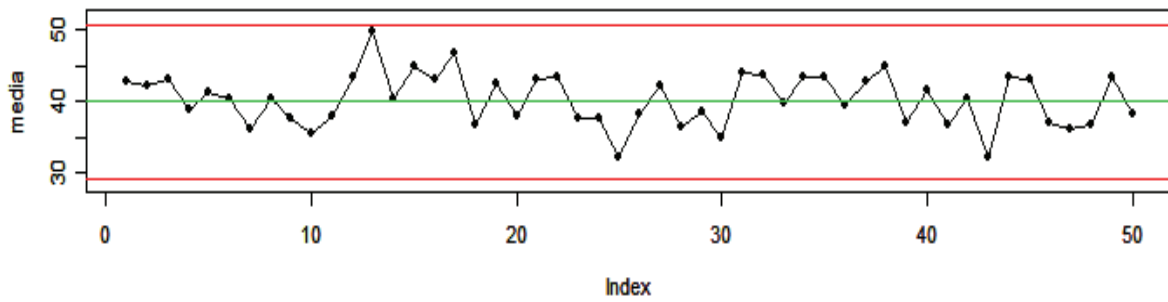
- $UCL = \sigma_E^2 (\beta_0 + \beta_1 z + \beta_2 z^2) = 50,908 (1,4865 - 0,093539 * 0,983 + 3,7335 * 0,983^2) = 254,7$
- $LCL = 0$

3)

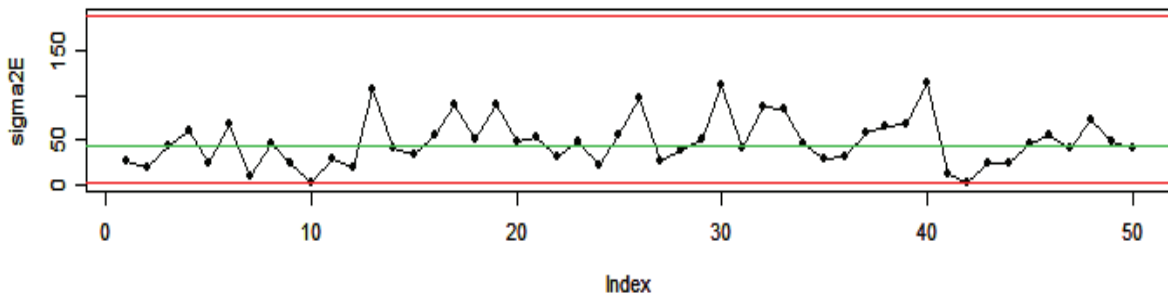
- $CL = \sigma_E^2 (\beta_0 + \beta_1 z + \beta_2 z^2) = 50,908 (-0,0072305 - 0,05503 * 0,983 + 0,85025 * 0,983^2) = 38,7$

Di seguito la rappresentazione grafica delle carte con i campioni simulati.

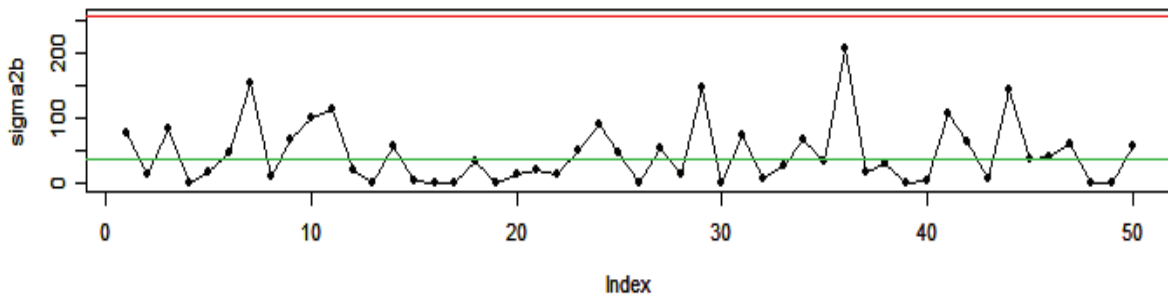
**carta Shewhart per le medie**



**carta Shewhart per sigma2e**



**carta Shewhart per sigma2b**



### 3.3 COSTRUZIONE E DISEGNO DELLE CARTE DI CONTROLLO CUSUM

In questo paragrafo verrà illustrata la costruzione delle carte di controllo CUSUM per la sorveglianza della media e delle componenti di varianza. Successivamente le carte verranno disegnate per gli stessi campioni ottenuti dalla precedente simulazione (vedi paragrafo 3.2.4).

#### 3.3.1 Carta di controllo CUSUM per la media

Lo schema di controllo CUSUM per il monitoraggio di  $\mu$  è implementato con una coppia di schemi:

- schema unilaterale superiore disegnato per l'individuazione di un aumento del valore di  $\mu$ . Lo schema è denotato con  $\{S_i^+\}$  ed è calcolato usando la seguente formula:

$$S_0^+ = 0; \quad S_i^+ = \max[S_{i-1}^+ + (\bar{X}_i - k^+), 0]$$

- schema unilaterale inferiore disegnato per l'individuazione di una diminuzione del valore di  $\mu$ . Lo schema è denotato con  $\{S_i^-\}$  ed è calcolato usando la seguente formula:

$$S_0^- = 0; \quad S_i^- = \max[S_{i-1}^- + (-\bar{X}_i - k^-), 0]$$

Ogni volta che  $S_i^+ > h^+$  oppure  $S_i^- > h^-$  la carta segnala un allarme. I valori  $h^+ \geq 0$  e  $h^- \geq 0$  sono detti soglia superiore e inferiore. Il valore  $k^+$  è detto valore di riferimento superiore ed è tipicamente scelto tra il limite superiore della regione accettabile per  $\mu$  e il limite inferiore della regione inaccettabile per  $\mu$ .

Analogamente il valore  $k^-$  è detto valore di riferimento inferiore ed è tipicamente scelto tra il limite inferiore della regione accettabile per  $\mu$  e il limite superiore della regione inaccettabile per  $\mu$ .

Si determinano quindi le regioni accettabili e inaccettabili ricavando i valori di  $k^+$  e  $k^-$  a seconda del salto di livello in unità di  $\sigma$ .

Dalla formula:

$$S_i^+ = \max[S_{i-1}^+ + (\bar{X}_i - k^+), 0]$$

$$= \max \left[ S_{i-1}^+ + \bar{X}_i - \left( \mu_0 + \frac{|\mu_1 - \mu_0|}{2} \right), 0 \right]$$

Si ricavano:

$$k^+ = \left( \mu_0 + \frac{|\mu_1 - \mu_0|}{2} \right) \quad \text{e} \quad k^- = \left( \mu_0 - \frac{|\mu_1 - \mu_0|}{2} \right)$$

dove  $\mu_0$  rappresenta la media del processo in controllo e  $\mu_1$  rappresenta la media del processo fuori controllo;  $\mu_1$  in particolare è espresso in termini di salto di livello in unità di  $\sigma$ :  $\mu_1 = \mu_0 + \delta\sigma$  dove  $\sigma = \sqrt{\text{Var}(\bar{X}_{..})}$ .

Costruite le statistiche  $\{S_i^+\}$  e  $\{S_i^-\}$ , i valori di  $h^+$  e  $h^-$  per un opportuno ARL in controllo vengono ricavati risolvendo numericamente la relativa equazione integrale (Goel e Wu, 1971; Brook e Evans, 1972; Lucas e Crosier, 1982; Vance, 1986; Waldmann, 1986; Crosier, 1986).



### 3.3.2 Carta di controllo CUSUM per la componente di varianza $\sigma_E^2$

Se la regione accettabile per  $\sigma$  è  $\sigma \leq \sigma_0$  e la regione inaccettabile per  $\sigma$  è  $\sigma \geq \sigma_1$ , con  $(\sigma_0 \leq \sigma_1)$ , lo schema di controllo CUSUM è implementato con un unico schema unilaterale superiore definito come segue:

$$S_0 = 0; \quad S_i = \max[S_{i-1} + (\hat{\sigma}_i^2 - k), 0]$$

dove il valore di riferimento  $k$  è calcolato come segue:

$$k = k_\sigma = \frac{2 \ln\left(\frac{\sigma_1}{\sigma_0}\right)}{\sigma_0^{-2} - \sigma_1^{-2}}$$

La soglia approssimata  $H=h*\sigma$ , con  $\sigma = \sqrt{\text{Var}(\hat{\sigma}_E^2)}$ , può essere calcolata tramite simulazione (vedi appendice)

### 3.3.3 Carta di controllo CUSUM per la componente di varianza $\sigma_B^2$

Lo schema CUSUM per il monitoraggio di questa componente di varianza è più complesso, soprattutto per la presenza del parametro di disturbo  $\sigma_E^2$ . Per semplificare la notazione introduciamo il parametro  $\eta = \frac{\sigma_E^2}{n}$  e denotiamo con  $\sigma_{B0}$  il limite superiore della regione accettabile e con  $\sigma_{B1}$  il limite inferiore della regione inaccettabile, con  $(\sigma_{B0} < \sigma_{B1})$ .

Tratteremo prima il caso in cui  $\eta$  sia noto e successivamente il caso in cui sia incognito.

#### 1. $\eta$ noto.

In questo caso la statistica di controllo per il monitoraggio di  $\sigma_B^2$  è semplicemente la sequenza degli stimatori non distorti  $\{\hat{\sigma}_B^2 - \eta\}$  dove  $\hat{\sigma}_B^2$  è dato dalla (2.3).

Lo schema di controllo CUSUM è definito come segue:

$$S_0 = 0; \quad S_i = \max[S_{i-1} + \hat{\sigma}_B^2 - \eta - k_b(\eta), 0]$$

dove il valore di riferimento  $k_b(\eta)$  è calcolato come segue:

$$k_b(\eta) = -\eta + \frac{\ln\left[\frac{\sigma_{B1}^2 + \eta}{\sigma_{B0}^2 + \eta}\right]}{(\sigma_{B0}^2 + \eta)^{-1} - (\sigma_{B1}^2 + \eta)^{-1}}$$

La soglia approssimata  $H=h*\sigma$ , con  $\sigma = \sqrt{\text{Var}(\hat{\sigma}_B^2)}$ , può essere calcolata tramite simulazione (vedi appendice).

## 2. $\eta$ incognito.

In questo caso la statistica di controllo per il monitoraggio di  $\sigma_B^2$  è data da  $\hat{\sigma}_B^2 = \left\{ \hat{\sigma}_B^2 - \frac{\hat{\sigma}_E^2}{n} \right\}$

Lo schema di controllo CUSUM è definito come segue:

$$S_0 = 0; \quad S_i = \max[S_{i-1} + \hat{\sigma}_b^2 - k_b(\hat{\eta}), 0]$$

dove il valore di riferimento  $k_b(\eta)$  è calcolato come segue:

$$k_b(\hat{\eta}) = -\hat{\eta} + \frac{\ln \left[ \frac{\sigma_{B1}^2 + \hat{\eta}}{\sigma_{B0}^2 + \hat{\eta}} \right]}{(\sigma_{B0}^2 + \hat{\eta})^{-1} - (\sigma_{B1}^2 + \hat{\eta})^{-1}}$$

La soglia approssimata  $H=h*\sigma$ , con  $\sigma = \sqrt{\text{Var}(\hat{\sigma}_B^2)}$ , può essere calcolata tramite simulazione (vedi appendice).

### 3.3.4 Esempio di disegno delle carte CUSUM per il controllo della media e delle componenti di varianza

Di seguito vengono disegnate le tre carte CUSUM per gli stessi 50 campioni ottenuti tramite la precedente simulazione (paragrafo 3.2.4) e sotto le stesse assunzioni. Le funzioni per la simulazione dei campioni e per il disegno delle carte si trovano in appendice.

Per la sorveglianza della media del processo, considerando inaccettabile un salto di livello pari a  $1\sigma$  le aree di rifiuto sono  $|\mu - \mu_0| > \delta\sigma$ . Dal paragrafo 3.3.2 abbiamo che:

$$\text{Var}(\bar{X}_{..}) = \frac{1}{r} \sigma_B^2 + \frac{1}{rn} \sigma_E^2 = \frac{1}{5} 7,014^2 + \frac{1}{5*2} 7,135^2 = 14,93$$

Dunque  $\sigma = \sqrt{\text{Var}(\bar{X}_{..})} = 3,86$ , le aree di rifiuto sono  $|\mu - 40| > 3,86$  e  $\mu_1 = \mu_0 + \delta\sigma = 43,86$

A questo punto i valori di  $k^+$  e  $k^-$  sono pari a:

$$k^+ = \left( 40 + \frac{|43,86-40|}{2} \right) = 41,93$$

$$k^- = \left( 40 - \frac{|43,86-40|}{2} \right) = 38,07$$

Con  $k$  pari a  $0,5 = (\delta/2)$  e per un'ARL in controllo pari a 200, le soglie  $H=h*\sigma$  sono  $\pm 3,5*3,86 = \pm 13,52$

Per la sorveglianza della componente di varianza  $\sigma_E^2$ , considerando accettabile la regione  $\sigma \leq \sigma_0$  ed inaccettabile un salto di livello pari a  $1\sigma$ , l'area di rifiuto è  $\sigma \geq \sigma_1$ . Dal paragrafo 3.2.2 abbiamo che la

variabile casuale  $\frac{r(n-1)\hat{\sigma}_E^2}{\sigma_E^2} \sim X_{r(n-1)}^2$  da cui si ottiene  $\hat{\sigma}_E^2 = \frac{X_{r(n-1)}^2 \sigma_E^2}{r(n-1)}$

$$\begin{aligned} \text{s.e.}(\hat{\sigma}_E^2) &= \sqrt{\text{Var}(\hat{\sigma}_E^2)} = \sqrt{\text{Var}\left(\frac{X_{r(n-1)}^2 \sigma_E^2}{r(n-1)}\right)} = \sqrt{\frac{\sigma^4}{(r(n-1))^2} \text{Var}(X_{r(n-1)}^2)} = \sqrt{\frac{2\sigma^4 r(n-1)}{(r(n-1))^2}} = \\ &= \sqrt{\frac{2(7,135^4)}{5(2-1)}} = 32,197 \end{aligned}$$

$$\text{Dunque } \sigma = 32,197; \quad \sigma_1^2 = \sigma_0^2 + \delta(\text{s.e.}(\hat{\sigma}^2)); \quad \sigma_1 = \sqrt{7,135^2 + 1(32,197)^2} = 9,116$$

la regione inaccettabile per  $\sigma$  è  $\sigma \geq 9,116$  quindi il valore di riferimento  $k$  è pari a:

$$k_\sigma = \frac{2 \ln\left(\frac{9,116}{7,135}\right)}{(7,135)^{-2} - (9,116)^{-2}} = 64,4$$

Con  $k$  pari a 0,5 = ( $\delta/2$ ) e per un'ARL in controllo pari a 200, la soglia approssimata  $H=h*\sigma$  calcolata tramite simulazione è pari a 161.

Per la sorveglianza della componente di varianza  $\sigma_B^2$ , considerando accettabile la regione  $\sigma \leq \sigma_0$  ed inaccettabile un salto di livello pari a  $1\sigma$ , l'area di rifiuto è  $\sigma \geq \sigma_1$ .

$$\text{Si ricordi che } \sigma_B^2 = \sigma_B^2 + \frac{\sigma_E^2}{n} = 74,65$$

Dalla (3.8) abbiamo che  $\hat{\sigma}_B^2$  può essere espressa come la differenza tra due variabili casuali indipendenti, ognuna proporzionale ad una variabile casuale chi-quadro, quindi:

$$\hat{\sigma}_B^2 = \frac{X_{(r-1)}^2 \sigma_B^2}{r-1} \rightarrow \text{Var}(\sigma_B^2) = \text{Var}\left(\frac{X_{(r-1)}^2 \sigma_B^2}{r-1}\right) = \left(\frac{2(r-1)\sigma_B^4}{(r-1)^2}\right) = \left(\frac{2\sigma_B^4}{r-1}\right)$$

$$\hat{\sigma}_E^2 = \frac{X_{r(n-1)}^2 \sigma^2}{r(n-1)} \rightarrow \text{Var}\left(\frac{\hat{\sigma}_E^2}{n}\right) = \text{Var}\left(\frac{X_{r(n-1)}^2 \sigma^2}{n(r(n-1))}\right) = \left(\frac{2r(n-1)\sigma^4}{(r(n-1))^2 n^2}\right) = \left(\frac{2\sigma^4}{(r(n-1))n^2}\right)$$

$$\text{s.e.}(\hat{\sigma}_B^2) = \sqrt{\text{Var}(\hat{\sigma}_B^2)} = \sqrt{\text{Var}(\sigma_B^2) + \text{Var}\left(\frac{\hat{\sigma}_E^2}{n}\right)} = \sqrt{\frac{2 * 74,65^2}{5-1} + \frac{2 * 50,908^2}{(5(2-1))^2}} = 55,186$$

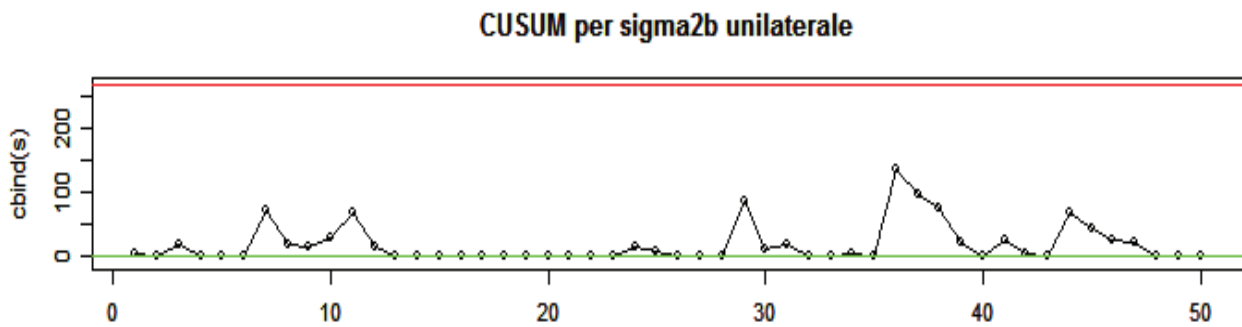
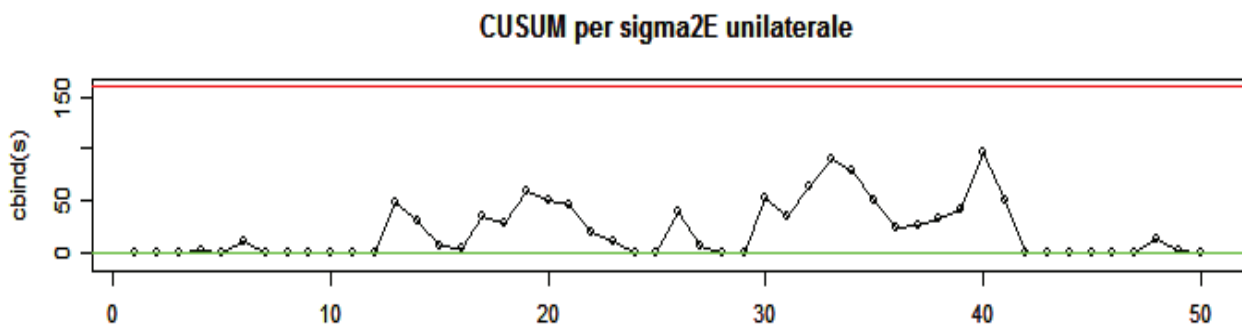
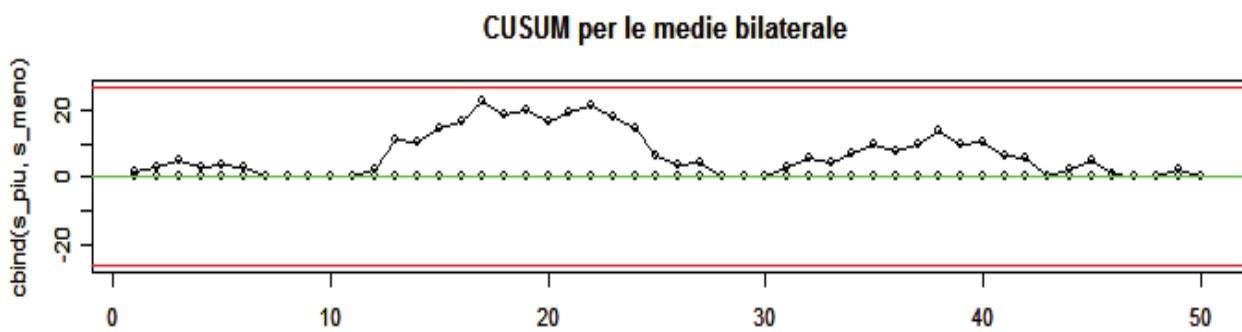
$$\text{Dunque } \sigma_{B1}^2 = \sigma_{B0}^2 + \delta(\text{s.e.}(\hat{\sigma}_B^2)); \quad \sigma_1 = \sqrt{7,014^2 + 1(55,186)^2} = 10,217$$

la regione inaccettabile per  $\sigma$  è  $\sigma \geq 10,217$ . Il valore di riferimento  $k_b(\eta)$  è pari a:

$$\text{con } \eta = \frac{\sigma_E^2}{n} = 25,45 \quad k_b(\eta) = -25,45 + \frac{\ln\left[\frac{10,217^2+25,45}{7,014^2+25,45}\right]}{(7,014^2+25,45)^{-1}-(10,217^2+25,45)^{-1}} = 71,75$$

Con  $k$  pari a  $0,5$  ( $\delta/2$ ) e per un'ARL in controllo pari a  $200$ , la soglia approssimata  $H=h*\sigma$  calcolata tramite simulazione è pari a  $268,15$ .

Di seguito la rappresentazione grafica delle carte con i campioni simulati.



### 3.4 CARTA DI CONTROLLO EWMA PER LA MEDIA

La carta EWMA è stata introdotta da Roberts (1959). Una buona presentazione dell'argomento può poi trovarsi in Crowder (1987a, 1989) e in Lucas e Saccucci (1990).

In questo paragrafo verrà illustrata la costruzione della carte di controllo EWMA per la sorveglianza della media in presenza di componenti di variazione gerarchica. Successivamente la carta sarà disegnata per gli stessi campioni ottenuti dalla precedente simulazione (vedi paragrafo 3.2.4).

La carta EWMA è definita come

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

dove  $0 < \lambda \leq 1$  è una costante e il valore iniziale (richiesto con il primo campione per  $i=1$ ) va fissato al valore di riferimento, ovvero  $z_0 = \mu_0$

Secondo il modello (2.1) abbiamo che:

$$E(X_{ij}) = \bar{X}_{..} \quad ; \quad Var(\bar{X}_{..}) = \frac{1}{r} \sigma_B^2 + \frac{1}{rn} \sigma_E^2; \quad \sigma = \sqrt{\frac{1}{r} \sigma_B^2 + \frac{1}{rn} \sigma_E^2}$$

i limiti di controllo per la carta EWMA per il monitoraggio di  $\mu$  sono:

- UCL:  $\bar{X}_{..} - L\hat{\sigma} \sqrt{\frac{\lambda}{2-\lambda}} [1 - (1 - \lambda)^t]$
- CL:  $\bar{X}_{..}$
- LCL:  $\bar{X}_{..} + L\hat{\sigma} \sqrt{\frac{\lambda}{2-\lambda}} [1 - (1 - \lambda)^t]$

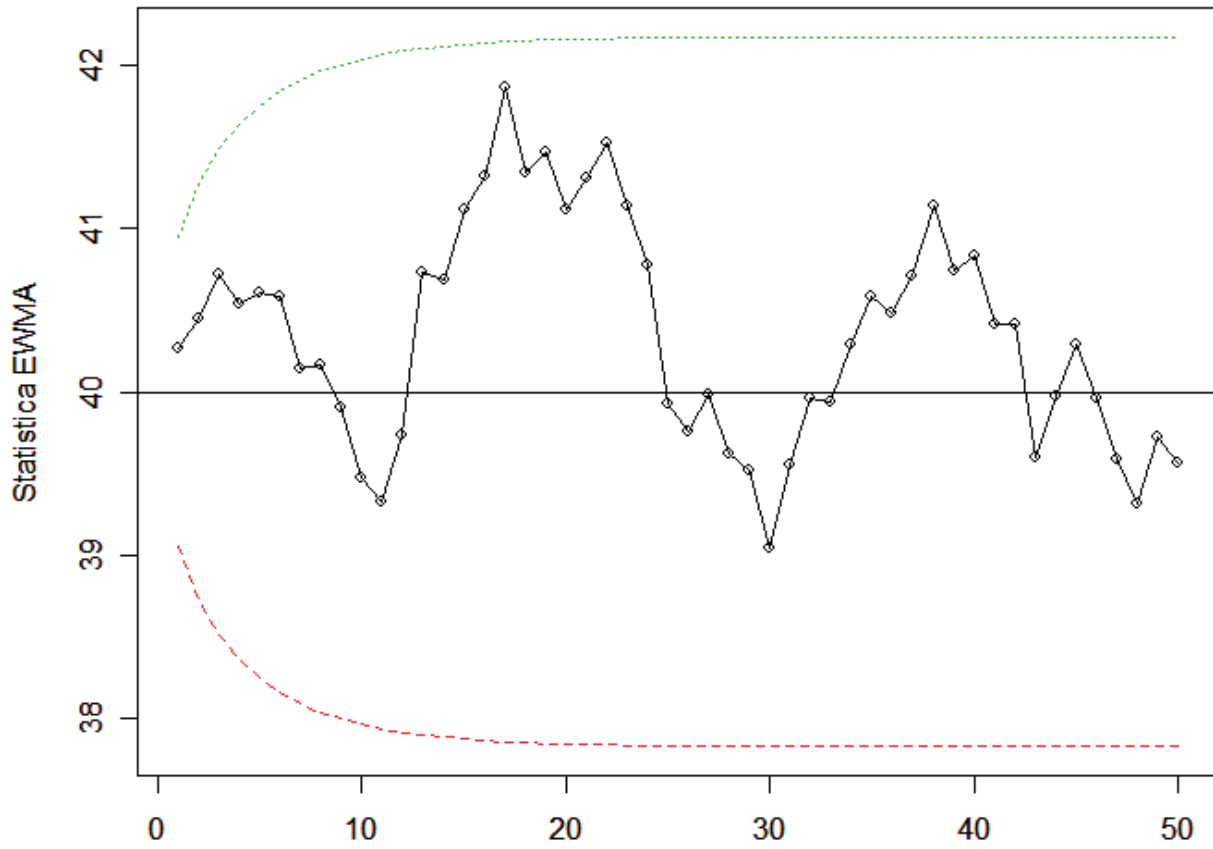
dove  $\hat{\sigma}$  sarà stimata da campioni provenienti dal processo in controllo. Il termine  $1 - (1 - \lambda)^t$  si approssima all'unità al crescere di  $t$  quindi dopo un periodo sufficientemente elevato si possono utilizzare i limiti di controllo asintotici:

- UCL:  $\bar{X}_{..} - L\hat{\sigma} \sqrt{\frac{\lambda}{2-\lambda}}$
- CL:  $\bar{X}_{..}$
- LCL:  $\bar{X}_{..} + L\hat{\sigma} \sqrt{\frac{\lambda}{2-\lambda}}$

Le funzioni per il calcolo delle statistiche EWMA, dei limiti di controllo e per il disegno della carta si trovano in appendice.

Di seguito la rappresentazione grafica della carta con i campioni simulati.

**carta EWMA per le medie**





## CAPITOLO 4

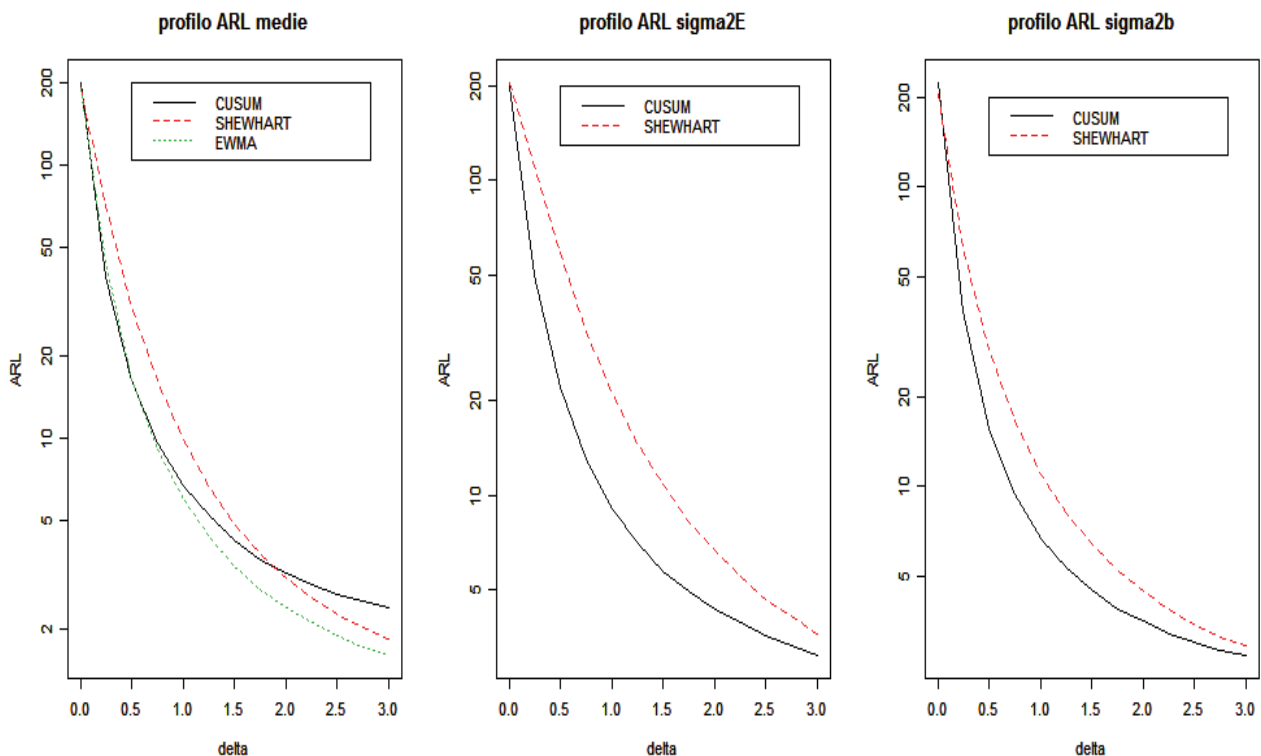
### CONFRONTO TRA I PROFILI ARL DELLE CARTE SHEWHART, CUSUM ED EWMA IN PRESENZA DI COMPONENTI GERARCHICHE DI VARIAZIONE

Con il seguente studio di simulazione si vogliono confrontare i profili ARL delle carte di controllo Shewhart, CUSUM ed EWMA adottate per la sorveglianza della media e delle componenti di varianza del processo di assemblaggio dei moduli di memoria descritto nel capitolo 3.

La simulazione è stata condotta assumendo che tutte le carte abbiano un ARL in controllo pari a 200. Le carte di controllo CUSUM ed EWMA sono state disegnate per essere sensibili ad un salto di livello pari a  $1\sigma$ .

Si sono ricavati i profili della ARL per ognuna delle sette carte (Shewhart e CUSUM per la media e per le componenti di varianza, carta EWMA per la media) mediando 10.000 RL (Run Length), questo per salti di livello dei parametri in unità di  $\sigma$  da 0 a 3.

La funzione utilizzata per la simulazione si trova in appendice mentre di seguito sono rappresentati graficamente i profili dell'ARL per le carte poste a confronto.





Dall'analisi dei profili ARL delle carte per le medie si osserva il risultato più importante della simulazione: la carta EWMA e la carta CUSUM risultano avere lo stesso profilo fino ad un salto di circa  $1\sigma$  mentre per salti più ampi della media l'EWMA si dimostra essere più efficiente, quando solitamente le due carte risulterebbero estremamente simili per qualsiasi salto di livello.

Il profilo della carta Shewhart rimane sempre al di sopra di quello EWMA, mentre risulta al di sotto di quello CUSUM per salti di livello maggiori di  $2\sigma$ . Da queste osservazioni si evince che per la sorveglianza della media del processo la carta EWMA risulta più efficiente rispetto alle altre.

Osservando i profili ARL delle carte Shewhart e CUSUM per la componente di varianza  $\sigma_E^2$  si nota che la carta CUSUM risulta costantemente la più efficace con una differenza importante fino ad un salto di  $2,5\sigma$ . Per salti maggiori la differenza risulta meno marcata. Le stesse osservazioni valgono per i profili ARL delle carte Shewhart e CUSUM per la componente di varianza  $\sigma_B^2$  con la differenza che l'efficienza della carta CUSUM risulta meno marcata per salti del parametro più grandi di  $2\sigma$ .

In conclusione, per la sorveglianza delle componenti di varianza, come atteso, converrà utilizzare carte di controllo CUSUM almeno per salti di livello inferiori a  $2\sigma$ .

## APPENDICE

### ISTRUZIONI UTILIZZATE PER I GRAFICI DELLE CARTE NEGLI ESEMPLI.

#### *Simulazione di 50 campioni*

```
x<-simula_n(50)
```

#### *Grafici Shewhart – paragrafo 3.2.4*

```
lcl<-29.2
cl<-40
ucl<-50.8
media<-x$medie
sigma2E<-x$sigma2e
sigma2b<-x$sigma2b
par(mfrow=c(3,1))
plot(media,type="l",ylim=c(min(media,lcl)-1,max(media,ucl)+1),main="carta Shewhart per le medie")
points(media,pch=19)
abline(h=c(cl,lcl,ucl),col=c(3,2,2))
lcl<-3.1
cl<-44.3
ucl<-187.2
plot(sigma2E,type="l",ylim=c(min(sigma2E,lcl)-1,max(sigma2E,ucl)+1),main="carta Shewhart per sigma2e")
points(sigma2E,pch=19)
abline(h=c(cl,lcl,ucl),col=c(3,2,2))
cl<-38.7
ucl<-254.7
lcl<-0
plot(sigma2b,type="l",ylim=c(min(sigma2b,lcl)-1,max(sigma2b,ucl)+1),main="carta Shewhart per sigma2b")
points(sigma2b,pch=19)
abline(h=c(cl,ucl),col=c(3,2))
```

#### *Grafici CUSUM – paragrafo 3.3.4*

```
sigma2bmedio<-x$sigma2bmedio

h<-xcusum.crit(k=0.25,L0=200,sided="two")
graficoCUSUM(media,delta=0.5,sigma=3.864,mu0=40,h=h,tit="CUSUM per le medie bilaterale")

h<-5 #soglia=161
graficoCUSUMsigma2E(sigma2E,delta=0.5,sigma=32.197,sigma0=7.135,h=h,tit="CUSUM per sigma2E unilaterale")

h<-4.86 #soglia=268.15
eta<-((7.135)^2)/2
graficoCUSUMsigma2b(sigma2bmedio,delta=0.5,sigma=55.186,sigma0=7.014,h=h,eta=eta,tit="CUSUM per sigma2b unilaterale")
```

#### *Grafico EWMA – paragrafo 3.4*

```
par(mfrow=c(1,1))
lambda<-0.1
arl0<-200
L<-xewma.crit(lambda,arl0,sided="two")
ewma_grafico(media,lambda,L,40,3.864,tit="carta EWMA per le medie")
```

## FUNZIONI UTILIZZATE

```
#####  
# simula un campione e ne calcola le stime della media e componenti di varianza  
#  
# media----->media generale  
# sigmab----->componente sigmab  
# sigmaE----->componente sigmaE  
# ngruppi----->numero gruppi per campione  
# nosservazioni-->numero osservazioni per gruppo  
#####  
  
simula <- function(media=40,sigmab=7.014,sigmaE=7.135,ngruppi=5,nosservazioni=2) {  
  x<-matrix(data = NA, nrow = ngruppi, ncol = nosservazioni,  
            byrow = TRUE,dimnames = list(c(1:ngruppi), c(1:nosservazioni)))  
  
  q<-0  
  z<-0  
  media_campione<-0  
  media_righe <-c(rep(0,ngruppi))  
  somma_riga <-c(rep(0,ngruppi))  
  sigma2bmedio <-0  
  sigma2b<-0  
  
  #generazione del campione e calcolo delle medie per gruppo  
  Wi<-rnorm(ngruppi,0,sigmab)  
  for(i in 1:ngruppi){  
    for(j in 1:nosservazioni){  
      Eij<-rnorm(1,0,sigmaE)  
      x[i,j]<-media+Wi[i]+Eij  
      somma_riga[i]<-somma_riga[i]+x[i,j] }  
      media_righe[i]<-somma_riga[i]/nosservazioni  
    }  
  }  
  #stima media campione  
  media_campione<-mean(media_righe)  
  
  for(i in 1:ngruppi){  
    for(j in 1:nosservazioni){  
      q<-q+((x[i,j]-media_righe[i])^2)/(nosservazioni-1) }  
      z<-z+(media_righe[i]-media_campione)^2 }  
  }  
  #stima di sigma2e  
  sigma2e<-q/ngruppi  
  #stima di sigma2bmedio  
  sigma2bmedio<-z/(ngruppi-1)  
  #stima di sigma2b  
  sigma2b<-sigma2bmedio-(sigma2e/nosservazioni)  
  if (sigma2b<0)  
    sigma2b<-0  
  
  list(campione=x,mu=media_campione,sigma2e=sigma2e,sigma2b=sigma2b,sigma2bmedio=sigma2bmedio)  
}
```

```
#####
# simula n campioni e fornisce i vettori della media e delle componenti di varianza
#
# n_campioni---->numero di campioni che si vogliono simulare
# media----->media generale
# sigmab----->componente sigmab
# sigmaE----->componente sigmaE
# ngruppi----->numero gruppi per campione
# nosservazioni-->numero osservazioni per gruppo
#####

simula_n <- function(n_campioni=100,media=40,sigmab=7.014,sigmaE=7.135,ngruppi=5,nosservazioni=2){
  campioni<-list()
  medie<-c(rep(0,n_campioni))
  s2b <-c(rep(0,n_campioni))
  s2e <-c(rep(0,n_campioni))
  sb <-c(rep(0,n_campioni))
  se <-c(rep(0,n_campioni))
  s2bmedio<-c(rep(0,n_campioni))

  for(i in 1:n_campioni){
    campione<-simula(media,sigmab,sigmaE,ngruppi,nosservazioni)
    medie[i] <-campione$mu
    s2b[i] <-campione$sigma2b
    s2e[i] <-campione$sigma2e
    sb[i] <-sqrt(campione$sigma2b)
    se[i] <-sqrt(campione$sigma2e)
    s2bmedio[i] <-campione$sigma2bmedio
    campioni[[i]]<-campione
  }

  list(medie=medie,sigma2b=s2b,sigmab=sb,sigma2e=s2e,sigmae=se,sigma2bmedio=s2bmedio)
}

```

```
#####
# EWMA - CALCOLA I LIMITI DI CONTROLLO
# dati---->medie dei campioni
# lambda-->
# media0-->media del processo in controllo
# L----->
# sigma--->s.e.(media0) -->s=v(1/r sigma2b + 1/rn sigma2E)
#           n=n°osservazioni per campione
#           r=n°gruppi
#
#####
```

```
ewma_limiti<-function(dati,lambda,L,media0,sigma)
{
k<-length(dati)
limite_sup_t<-c(rep(0,k))
limite_inf_t<-c(rep(0,k))

for (t in 1:k){
  limite_inf_t[t]<-media0-(L*sigma*sqrt((lambda/(2-lambda))*(1-(1-lambda)^(2*t))))
  limite_sup_t[t]<-media0+(L*sigma*sqrt((lambda/(2-lambda))*(1-(1-lambda)^(2*t))))
}

t<-t+1
lim_asint_inf<-media0-(L*sigma*sqrt((lambda/(2-lambda))))
lim_asint_sup<-media0+(L*sigma*sqrt((lambda/(2-lambda))))
rbind(cbind(limite_inf_t,limite_sup_t),c(lim_asint_inf,lim_asint_sup))
}
```

```
#####
# CALCOLA I VALORI DELLA STATISTICA EWMA
# dati---->medie dei campioni
# lambda-->
# media0-->media del processo in controllo
#####
```

```
ewma_statistica<-function(dati,lambda,media0)
{
k<-length(dati)
x<-c(rep(0,k))
for (t in 1:length(dati))
{
  if (t==1)
  x[1]<-lambda*dati[1]+(1-lambda)*media0
  else
    x[t]=lambda*dati[t]+(1-lambda)*x[t-1]
}
t<-t+1
x
}
```

```
#####
# DISEGNA IL GRAFICO EWMA
# dati---->medie dei campioni
# lambda-->
# media0-->media del processo in controllo
# L----->xewma.crit(lambda,ARL0,sided="two")
# sigma-->s.e.(media0) -->s=v(1/r sigma2b + 1/rn sigma2E)
#          n=n°osservazioni per campione; r=n°gruppi
#####
ewma_grafico<-function(dati,lambda,L,media0,sigma,tit) {
wewma<-ewma_statistica(dati,lambda,media0)
lim<-ewma_limiti(dati,lambda,L,media0,sigma)
lim2<-lim[1:length(dati),]
matplot(cbind(wewma,lim2),ylab="Statistica EWMA", type="l",lty=1:3,main=tit)
points(wewma)
abline(a=media0,b=0)
}
#####
# cusum per le medie
dati--->medie dei campioni
# sigma-->s.e.(mu)
# delta-->salto in termini di sigma per cui si vuole ottimizzare la carta
# mu0--->media in controllo
#####
cusumS<-function(dati,sigma,delta,mu0){
n<-length(dati)
spiu<-c(rep(0,n))
smeno<-c(rep(0,n))
mu1<-0
#calcolo mu1
mu1<-mu0+delta*sigma
#calcolo kpiu
kpiu<-(mu0+ ( abs(mu1-mu0)/2 ) )
#calcolo kmeno
kmeno<-(mu0- ( abs(mu1-mu0)/2 ) )
#calcolo statistiche spiu
for (t in 1:length(dati)){
if (t==1){
spiu[1]<-max(0,(dati[1]-kpiu))}
else{
spiu[t]<-max(0,(spiu[t-1]+dati[t]-kpiu))} }
#calcolo statistiche smeno
for (r in 1:length(dati)){
if (r==1){
smeno[1]<-max(0,(-dati[1]-kmeno))}
else{
smeno[r]<-max(0,(smeno[r-1]-dati[r]-kmeno))} }
r<-r+1
t<-t+1
list(mu1=mu1,kpiu=kpiu,kmeno=kmeno,spiu=spiu,smeno=smeno)
}

```

```
#####
# GRAFICO CUSUM PER LE MEDIE
#dati-->medie dei campioni
# sigma-->s.e.(mu)
# delta-->salto in termini di sigma per cui si vuole ottimizzare la carta
# mu0---->media in controllo
# h----->calcolare h
#
#      library(spc)
#      xcusum.crit(k=0.5,200)
#      per un k=delta/2=0,5 e un arl0=200, ad esempio
#      ottenuto h,la soglia sarà=h*sigma
#
# tit---->titolo grafico
#####

graficoCUSUM<-function(dati,delta,sigma,mu0,h,tit="CUSUM per le medie bilaterale")
{
S<-cusumS(dati,sigma,delta,mu0)
s_piu<-S$s_piu
s_meno<-S$s_meno
k_piu<-S$k_piu
k_meno<-S$k_meno
mu_1<-S$mu1

LCL<-h*sigma
UCL<-h*sigma

matplot(cbind(s_piu,s_meno),type="l",ylim=c(min(c(min(s_meno),LCL)),max(c(max(s_piu),UCL))),lty=1:2,main=tit)
points(s_piu)
points(s_meno)
abline(a=LCL,b=0,col="red")
abline(a=UCL,b=0,col="red")
abline(a=0,b=0,col="green")
w<-cbind(s_piu,s_meno)
print(w)
list(mu1=mu_1,kpiu=k_piu,kmeno=k_meno)
}

```

```
#####
#CUSUM PER SIGMA2E
# dati-->sigma2E dei campioni
# sigma-->s.e.(sigma2E)
# delta-->salto in termini di sigma per cui si vuole ottimizzare la carta
# sigma0-->sigmaE in controllo
#####
cusumS_sigma2E<-function(dati,sigma=32.197,delta=1,sigma0=7.135){
n<-length(dati)
s<-c(rep(0,n))
sigma1<-0
#calcolo sigma1
sigma1<-sqrt( (sigma0^2)+delta*sigma )
#calcolo k
k<- ( 2*log(sigma1/sigma0) / ((sigma0^(-2))-(sigma1^(-2))) )
#calcolo statistiche s
for (t in 1:length(dati)){
  if (t==1){
    s[1]<-max(0,(dati[1]-k))}
  else{
    s[t]<-max(0,(s[t-1]+dati[t]-k))}
}
t<-t+1
list(sigma1=sigma1,k=k,s=s)
}
#####
# GRAFICO CUSUM PER SIGMA2E
#dati---->sigma2E dei campioni
# sigma--->s.e.(sigma2E)
# delta--->salto in termini di sigma per cui si vuole ottimizzare la carta
# sigma0-->sigmaE in controllo
# h----->calcolare h
#
# tit---->titolo grafico
#####
graficoCUSUMsigma2E<-function(dati,delta=1,sigma=32.197,sigma0=7.135,h,tit="CUSUM per sigma2E
unilaterale")
{
S<-cusumS_sigma2E(dati,sigma,delta,sigma0)
s<-S$s
k<-S$k
sigma1<-S$sigma1

UCL<-h*sigma

matplot(cbind(s),type="l",ylim=c(-10,max(c(max(s),UCL))),lty=1:2,main=tit)
points(s)
abline(a=UCL,b=0,col="red")
abline(a=0,b=0,col="green")
list(s,k=k,sigma1=sigma1)
}

```



```
#####
# CUSUM PER SIGMA2B
#dati---->sigma2bmedio dei campioni
# sigma--->s.e.(sigma2b)
# delta--->salto in termini di sigma per cui si vuole ottimizzare la carta
# sigma0-->sigmab in controllo
# eta---->eta=sigma2E/n°osservazioni per campione
#####
cusumS_sigma2b<-function(dati,sigma=55.186,delta=1,sigma0=7.014,eta){

n<-length(dati)
s<-c(rep(0,n))
sigma1<-0

#calcolo sigma1
sigma1<-sqrt( (sigma0^2)+delta*sigma )
#calcolo k
k<- -eta+( ( log( ((sigma1^2)+eta)/((sigma0^2)+eta) ) ) / ( (((sigma0^2)+eta)^-1)-(((sigma1^2)+eta)^-1) ) )
#calcolo statistiche s
for (t in 1:length(dati)){
  if (t==1){
    s[1]<-max(0,(dati[1]-eta-k))
  }
  else{
    s[t]<-max(0,(s[t-1]+dati[t]-eta-k))
  }
}
t<-t+1
list(sigma1=sigma1,k=k,s=s)
}
#####
# dati---->sigma2bmedio dei campioni
# sigma--->s.e.(sigma2b)
# delta--->salto in termini di sigma per cui si vuole ottimizzare la carta
# sigma0-->sigmab in controllo
# h----->calcolare h
#
# tit---->titolo grafico
# eta---->eta=sigma2E/n°osservazioni per campione
#####
graficoCUSUMsigma2b<-function(dati,delta=1,sigma=55.186,sigma0=7.014,h,eta,tit="CUSUM per sigma2b
unilaterale"){
S<-cusumS_sigma2b(dati,sigma,delta,sigma0,eta)
s<-S$s
k<-S$k
sigma1<-S$sigma1
UCL<-h*sigma
matplot(cbind(s),type="l",ylim=c(-10,max(c(max(s),UCL))),lty=1:2,main=tit)
points(s)
abline(a=UCL,b=0,col="red")
abline(a=0,b=0,col="green")
list(s,k=k,sigma1=sigma1)
}

```

```
#####
#CALCOLA LA SOGLIA PER LA CARTA CUSUM SU SIGMA2B DATO UN ARLO FISSATO
#####

calcola_arl_siglab<-
function(arl0=200,sigmaE=7.135,siglab=7.014,sigma=55.186,delta=1,h=3.5,nosservazioni=2){
valoriRLperiterazione<-10000
max_iterazioni<-100
scartoARL<-2 #differenza massima tra arl0 e arl trovato
arl<-c(rep(0,max_iterazioni)) #vettore che contiene i valori di arl trovati nelle iterazioni
soglia<-c(rep(0,max_iterazioni)) #vettore che contiene i valori della soglia usati nelle
iterazioni
soglia[1]<-h*sigma #valore della soglia iniziale

valori_run_length<-c(rep(0,valoriRLperiterazione)) #numero valori run length per ogni iterazione (vengono
ritornati solo quelli dell'ultima iterazione)
s<-c(rep(0,10000)) #numero campioni

sigma1<-0
numero_iterazioni=0

eta<-(sigmaE^2)/nosservazioni
#calcolo sigma1
sigma1<-sqrt( (siglab^2)+delta*sigma )
#calcolo k
k<--eta+( ( log( ((sigma1^2)+eta)/((siglab^2)+eta) ) ) / ( (((siglab^2)+eta)^(-1))-(((sigma1^2)+eta)^(-1))
) )

q=0 #quando vale 1 blocca le iterazioni
t<-1 #contatore iterazioni
z1<-date() #ora di avvio iterazioni
while(q==0)
{

for(j in 1:valoriRLperiterazione) #genero le run length
{
i=1
s2bm<-simula()$sigma2bmedio
s[1]<-max(0,(s2bm-eta-k))

while(s[i]<soglia[t])
{
i=i+1
s2bm<-simula()$sigma2bmedio
s[i]<-max(0,(s[i-1]+s2bm-eta-k))
if(i>10000)s[i]<-99999
}

#il valore di i è la j-esima run length
valori_run_length[j]<-i
}
}
}

```

```

arl[t]<-mean(valori_run_length)      #'arl trovato è la media delle RL

#controllo la differenza tra arl0 e arl trovato, quindi abbasso o alzo la soglia se necessario
if(arl[t]<(arl0-scartoARL))
  {
  soglia[t+1]=soglia[t]+1
  q=0
  t<-t+1
  numero_iterazioni=numero_iterazioni+1
  if(numero_iterazioni==max_iterazioni)q=1
  }

else if(arl[t]>(arl0+scartoARL))
  {
  soglia[t+1]=soglia[t]-1
  q=0
  t<-t+1
  numero_iterazioni=numero_iterazioni+1
  if(numero_iterazioni==max_iterazioni)q=1
  }

else{  numero_iterazioni=numero_iterazioni+1
      q=1
      t<-t+1}
}

z2<-date()                          #ora di fine iterazioni

list(ultimi_valoriRL=valori_run_length,k=k,arl0=arl0,soglia=soglia,arl=arl,numero_iterazioni=numero_iterazioni,time_inizio=z1,time_fine=z2)
}

```

```
#####
#CALCOLA LA SOGLIA PER LA CARTA CUSUM SU SIGMA2E DATO UN ARLO FISSATO
#####
```

```
calcola_arl_sigmae<-function(arl0=200,sigmaE=7.135,sigmab=7.014,sigma=32.197,delta=1,h=5){
valoriRLperiterazione<-10000
max_iterazioni<-100
scartoARL<-2 #differenza massima tra arl0 e arl trovato
arl<-c(rep(0,max_iterazioni))
soglia<-c(rep(0,max_iterazioni)) #vettore che contiene i valori della soglia via via modificati
soglia[1]<-h*sigma #valore della soglia iniziale
```

```
valori_run_length<-c(rep(0,valoriRLperiterazione)) #numero valori run length per ogni iterazione
s<-c(rep(0,10000)) #numero campioni
sigma1<-0
```

```
numero_iterazioni=0
```

```
#calcolo sigma1
```

```
sigma1<-sqrt( (sigmaE^2)+delta*sigma )
```

```
#calcolo k
```

```
k<- ( 2*log(sigma1/sigmaE) ) / ((sigmaE^(-2))-(sigma1^(-2))) )
```

```
q=0 #quando vale 1 blocca le iterazioni
```

```
t<-1 #contatore iterazioni
```

```
z1<-date() #ora di avvio iterazioni
```

```
while(q==0)
```

```
{
```

```
for(j in 1:valoriRLperiterazione) #genero le run length
```

```
{
```

```
i=1
```

```
s2e<-simula()$sigma2e
```

```
s[1]<-max(0,(s2e-k))
```

```
while(s[i]<soglia[t])
```

```
{
```

```
i=i+1
```

```
s2e<-simula()$sigma2e
```

```
s[i]<-max(0,(s[i-1]+s2e-k))
```

```
if(i>10000)s[i]<-99999
```

```
}
```

```
#il valore di i è la j-esima run length
```

```
valori_run_length[j]<-i
```

```
}
```

```
arl[t]<-mean(valori_run_length) #l'arl trovato è la media delle RL
```

```
#controllo la differenza tra arl0 e arl trovato, quindi abbasso o alzo la soglia se necessario
```

```
if(arl[t]<(arl0-scartoARL))
```

```

    {
    soglia[t+1]=soglia[t]+1
    q=0
    t<-t+1
    numero_iterazioni=numero_iterazioni+1
    if(numero_iterazioni==max_iterazioni)q=1
    }

    else if(arl[t]>(arl0+scartoARL))
        {
        soglia[t+1]=soglia[t]-1
        q=0
        t<-t+1
        numero_iterazioni=numero_iterazioni+1
        if(numero_iterazioni==max_iterazioni)q=1
        }

    else{  numero_iterazioni=numero_iterazioni+1
          q=1
          t<-t+1}
    }

z2<-date()                                #ora di fine iterazioni

list(ultimi_valoriRL=valori_run_length,k=k,arl0=arl0,soglia=soglia,arl=arl,numero_iterazioni=numero_iterazioni,time_inizio=z1,time_fine=z2)
}

```

```
#####
#CALCOLO ARL CARTE
#####
profilo_arl0_200<-function(cambia_mu=1,cambia_se=1,cambia_sb=1){
z1<-date()          #ora di inizio
numero_valori_RL<-10000

#dati
media<-40
sigmab=7.014
sigmaE=7.135
std_err_media=3.864
std_err_sigma2b=55.186
std_err_sigma2E=32.197
nosservazioni<-2
ngruppi<-5
delta=1          #delta con cui è ottimizzata la cusum
mu0<-media

###CUSUM SULLE MEDIE###
mu1<-mu0+delta*std_err_media  #calcolo mu1
kpiu<-(mu0+ ( abs(mu1-mu0)/2 ) ) #calcolo kpiu
kmeno<-(mu0- ( abs(mu1-mu0)/2 ) ) #calcolo kmeno
library(spc)
h<-3.5          #xcusum.crit(k=0.5,200)
UCLcusum_mu<- h*std_err_media  #soglia superiore
LCLcusum_mu<- -h*std_err_media  #soglia inferiore

###CUSUM SU SIGMA2E###
sigma1<-sqrt( (sigmaE^2)+delta*std_err_sigma2E )          #calcolo sigma1
k_se<- ( 2*log(sigma1/sigmaE) / ((sigmaE^(-2))-(sigma1^(-2))) )          #calcolo k
soglia_se<-161  #soglia per la CUSUM per sigma2E trovata con calcola_arl_sigmae()

###CUSUM SU SIGMAB###
sigma1<-sqrt( (sigmab^2)+delta*std_err_sigma2b )          #calcolo sigma1
eta<-(sigmaE^2)/nosservazioni
#calcolo k
k_sb<--eta+( ( log( ((sigma1^2)+eta)/((sigmaE^2)+eta) ) ) / ( (((sigmaE^2)+eta)^(-1))-(((sigma1^2)+eta)^(-1)) ) )
soglia_sb<-268.2  #soglia per la CUSUM per sigma2b trovata con calcola_arl_siglab()
###SHEWHART PER LE MEDIE###
LCLmu<-29.2
UCLmu<-50.8
###SHEWHART PER SIGMA2E###
LCLse<-3.1
UCLse<-187.2
###SHEWHART PER SIGMA2B###
LCLsb<-0
UCLsb<-254.7
###CARTA EWMA###
limite_sup_t<-c(rep(0,10000))
```

```

limite_inf_t<-c(rep(0,10000))
lambda<-0.1
L<-xewma.crit(0.1,200,sided="two")
stop<-c(rep(0,7))
  #stop[1] se==1 ferma calcolo statistiche spiu e smeno per CUSUM sulle medie
  #stop[2] se==1 ferma calcolo statistiche s per CUSUM su sigma2E
  #stop[3] se==1 ferma calcolo statistiche s per CUSUM su sigma2b
  #stop[4] se==1 ferma calcolo statistiche Shewhart sulle medie
  #stop[5] se==1 ferma calcolo statistiche Shewhart su sigma2E
  #stop[6] se==1 ferma calcolo statistiche Shewhart su sigma2b
  #stop[7] se==1 ferma calcolo statistiche EWMA sulle medie
RL <-matrix(data=0, nrow = 7, ncol = 10000)          #matrice di 7 righe e 10000 colonne per le RL delle
cartearl_carte <-matrix(data=0, nrow = 7, ncol = 13) #matrice che contiene gli arl -->7 colonne, 13 righe
valori_salto <-matrix(data=999, nrow = 3, ncol = 13) #matrice che contiene i valori dei salti in termini di
                                                    sigma -->3 colonne, 13 righe
valori_parametri<-matrix(data=999, nrow = 3, ncol = 13)#matrice che contiene i valori dei parametri -->
                                                    3 colonne, 13 righe

  #per i valori:
  #[1,]-->media
  #[2,]-->s2Eb
  #[3,]-->s2E

#vettori che contengono le statistiche S delle carte CUSUM e statistiche EWMA
spiu_cusum_medie <-c(rep(0,10000))
smeno_cusum_medie <-c(rep(0,10000))
s_cusum_s2E <-c(rep(0,10000))
s_cusum_s2b <-c(rep(0,10000))
ewma <-c(rep(0,10000))
#delta per il disegno del profilo ARL delle carte
delta<-seq(from=0,to=3,length=13) #0-0.25-0.5-...-3 salti in termini di std_err(parametro)

for(i in 1:13){
if(cambia_mu==1){
  media1<-media+delta[i]*std_err_media
  valori_salto[1,i]<-delta[i]
  valori_parametri[1,i]<-media1
}
else
  {
  media1 <-media
  valori_salto[1,i]<-0
  valori_parametri[1,i]<-media
  }
if(cambia_sb==1){
  sigmab1<-sqrt((sigmab^2)+(delta[i]*std_err_sigma2b))
  valori_salto[2,i]<-delta[i]
  valori_parametri[2,i]<-sigmab1
}
else
  {
  sigmab1=sigmab

```

```

        valori_salto[2,i]<-0
        valori_parametri[2,i]<-sigmab
    }
if(cambia_se==1){
    sigmaE1<-sqrt((sigmaE^2)+(delta[i]*std_err_sigma2E))
    valori_salto[3,i]<-delta[i]
    valori_parametri[3,i]<-sigmaE1
}
else
{
    sigmaE1=sigmaE
    valori_salto[3,i]<-0
    valori_parametri[3,i]<-sigmaE
}

j<-0

for(s in 1:numero_valori_RL){          #numero valori RL
    stop<-c(rep(0,7))
    ferma<-0                          #va a 1 se tutti gli stop sono a 1 (ciclo while)
    j<-j+1
    c<-1
    #simulo il campione c=1
    x<-simula(media1,sigmab1,sigmaE1,ngruppi,nosservazioni)

    mediax<-x$mu
    s2Ex <-x$sigma2e
    s2bx <-x$ sigma2b
    s2bmx <-x$sigma2bmedio

    #calcolo le statistiche delle carte CUSUM per il primo campione
    spiu_cusum_medie[1] <-max(0,mediax-kpiu)
    smeno_cusum_medie[1]<-max(0,mediax-kmeno)
    s_cusum_s2E[1]   <-max(0,s2Ex-k_se)
    s_cusum_s2b[1]  <-max(0,s2bmx-eta-k_sb)

    #carte Shewhart (primo campione)
    if((mediax>UCLmu)|(mediax<LCLmu)){
        stop[4]<-1
        RL[4,j]<-1
    }
    if((s2Ex>UCLse)|(s2Ex<LCLse)){
        stop[5]<-1
        RL[5,j]<-1
    }
    if((s2bx>UCLsb)|(s2bx<LCLsb)){
        stop[6]<-1
        RL[6,j]<-1
    }
}

```



```
#carta ewma
```

```
limite_inf_t[c]<-media-(L*std_err_media*sqrt((lambda/(2-lambda))*(1-(1-lambda)^(2*c))))
limite_sup_t[c]<-media+(L*std_err_media*sqrt((lambda/(2-lambda))*(1-(1-lambda)^(2*c))))
ewma[1]<-lambda*mediax+(1-lambda)*media
if((ewma[1]>limite_sup_t[c])|(ewma[1]<limite_inf_t[c]))
  {
    stop[7]<-1
    RL[7,j]<-1
  }
```

```
#PER I CAMPIONI DOPO IL PRIMO...
```

```
while(ferma==0) #genero campioni finchè non ho trovato la j-esima RL per tutte le carte
{
  c <- c+1
  x <- simula(media1,sigtab1,sigmaE1) #simulo il campione c; c=2,...
  mediax<- x$mu
  s2Ex <- x$sigma2e
  s2bx <- x$sigma2b
  s2bmx <- x$sigma2bmedio
  ###CUSUM SULLE MEDIE###
  if(stop[1]==0)
    {
      spiu_cusum_medie[c]<-max(0,spiu_cusum_medie[c-1]+mediax-kpiu)
      smeno_cusum_medie[c]<-max(0,smeno_cusum_medie[c-1]-mediax-kmeno)
      if((spiu_cusum_medie[c]>UCLcusum_mu)|(smeno_cusum_medie[c]<LCLcusum_mu))
        {
          stop[1]<-1
          RL[1,j]<-c
        }
    }
  ###CUSUM SU SIGMA2E###
  if(stop[2]==0)
    {
      s_cusum_s2E[c]<-max(0,s_cusum_s2E[c-1]+s2Ex-k_se)
      if(s_cusum_s2E[c]>soglia_se)
        {
          stop[2]<-1
          RL[2,j]<-c
        }
    }
  ###CUSUM SU SIGMA2B###
  if(stop[3]==0)
    {
      s_cusum_s2b[c]<-max(0,s_cusum_s2b[c-1]+s2bmx-eta-k_sb)
      if(s_cusum_s2b[c]>soglia_sb)
        {
          stop[3]<-1
          RL[3,j]<-c
        }
    }
}
```

```

###SHEWHART PER LE MEDIE###
if(stop[4]==0){

  if((mediax>UCLmu)|(mediax<LCLmu)){
    stop[4]<-1
    RL[4,j]<-c
  }
}
###SHEWHART PER SIGMA2E###
if(stop[5]==0){

  if((s2Ex>UCLse)|(s2Ex<LCLse)){
    stop[5]<-1
    RL[5,j]<-c
  }
}
###SHEWHART PER SIGMA2B###
if(stop[6]==0){

  if((s2bx>UCLsb)|(s2bx<LCLsb)){
    stop[6]<-1
    RL[6,j]<-c
  }
}
###EWMA PER LE MEDIE###
if(stop[7]==0){
limite_inf_t[c]<-media-(L*std_err_media*sqrt((lambda/(2-lambda))*(1-(1-lambda)^(2*c))))
limite_sup_t[c]<-media+(L*std_err_media*sqrt((lambda/(2-lambda))*(1-(1-lambda)^(2*c))))
ewma[c]<-lambda*mediax+(1-lambda)*ewma[c-1]
if((ewma[c]>limite_sup_t[c])|(ewma[c]<limite_inf_t[c]))
  {
    stop[7]<-1
    RL[7,j]<-c
  }
}
if(c>10000)ferma=1
if(mean(stop)==1)ferma<-1      #tutte le RL sono state trovate

} #fine while

} #fine for (s arl)

arl_carte[,i]<-apply(RL,1,mean)      #matrice RL completa --> media su righe per ottenere gli ARL per le
7 carte con delta j-esimo

} #fine for salti delta

z2<-date()      #ora di fine
list(arl_carte=arl_carte,salti=valori_salto,valore_parametri=valori_parametri,time_inizio=z1,time_fine=z2)
}

```



## BIBLIOGRAFIA

- Brook D. e Evans D. A. (1972), "An approach to the probability distribution of cusum run length". *Biometrika* 59, 539-548.
- Chang T. C. e Gan F. F. (2004), "Shewhart charts for monitoring the variance components". *Journal of Quality Technology*, Vol. 36 No 3, 293-308
- Cox D. R. e P. J. Solomon, "Components of variance", *Monographs on Statistics and Applied Probability* 97. Chapman & Hall/CRC
- Crosier R. B. (1986), "A new two-sided cumulative quality control scheme". *Technometrics* 28, 187-194
- Crowder S. V. (1987), "A simple method for studying run-length distributions of exponentially weighted moving average charts". *Technometrics* 29, 401-407
- Goel A. L. e Wu S. M. (1971), "Determination of A.R.L. and a contour nomogram for CUSUM charts to control normal mean". *Technometrics* 13, 221-230
- Lucas J. M. e Crosier R. B. (1982), "Fast initial response for cusum quality-control schemes: Give your cusum a headstart". *Technometrics* 24, 199-205
- Lucas J. M. e Saccucci M. S. (1990), "Exponentially weighted moving average control schemes: Properties and enhancements". *Technometrics* 32, 1-12
- Montgomery, D. C. (2001). "Controllo statistico della qualità". McGraw-Hill
- Vance L. C. (1986), "Average run lengths of cumulative sum control charts for controlling normal means". *Journal of Quality Technology* 18, 189-193
- Waldmann K.-H. (1986), "Bounds for the distribution of the run length of one-sided and two-sided CUSUM quality control schemes". *Technometrics* 28, 61-67
- Yashchin, E. (1994), "Monitoring variance components". *Technometrics*, Vol. 36 No. 4, 379-393