

Università degli studi di Padova



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

TESI DI LAUREA

Sorveglianza della capacità. Confronto tra
gli schemi EWMA e AEWMA.

Capability monitoring. Comparison
between the EWMA and AEWMA
schemes.

31 ottobre 2008

Relatore: Ch.mo Prof. Guido Masarotto

Laureando: Alessandro Rancan – 534506-GEI

ANNO ACCADEMICO 2008-09

Indice

Prefazione	1
1 Controllo Statistico del Processo	5
1.1 Carte di Controllo.....	5
1.1.1 Carta di Controllo di Shewhart.....	10
1.1.2 Carta di Controllo EWMA.....	12
1.1.3 Carta di Controllo AEWMA	14
2 Sorveglianza della Capacità	17
2.1 Un'introduzione al significato di Capacità del Processo.....	17
2.2 Sorveglianza della Capacità.....	19
3 Confronto tra gli schemi EWMA e AEWMA	23
3.1 Introduzione.....	23
3.2 Metodo di calcolo.....	23
3.3 Risultati.....	27
3.4 Conclusioni	29
Appendice Codice sorgente	31
Bibliografia	37
Ringraziamenti	39

Prefazione

Negli ultimi anni, nella gran parte dei mercati¹, ogni impresa deve competere almeno con un'altra per aggiudicarsi la preferenza del cliente; inoltre il progresso della tecnologia e della scienza assieme alla necessità di abbassare i costi per poter aumentare il proprio profitto ha fatto sì che l'attenzione si focalizzasse sulla *Qualità*.

In tutte le realtà in cui si produce un bene o un servizio la mancanza di qualità comporta dei costi aggiuntivi sia per il richiedente che per l'offerente; un banale esempio è la perdita di tempo e di denaro, conseguenze del fermare la produzione, lo scarto, la rilavorazione, etc... dovute ad un prodotto difettoso o con caratteristiche diverse da quelle concordate con il cliente.

Principalmente per questo motivo la *Qualità* è diventata un fattore discriminante nel processo di decisione del consumatore a fronte dell'incertezza dovuta alla omogeneità degli attributi fisico - funzionali dei prodotti offerti dalle imprese tra loro concorrenti.

Per il produttore stesso, poi, attuare una strategia rivolta al miglioramento continuo della *Qualità* comporta vantaggi sostanziali in termini di efficienza e, di conseguenza, ad aumentare i guadagni.

Perciò l'impresa consapevole del valore della *Qualità* e che agisce in tal senso aumenta la sua quota di mercato e i profitti migliorando la posizione nel mercato.

¹ Intesi come l'insieme di imprese produttrici di prodotti quali beni o servizi in diretta concorrenza tra loro.

Per riuscire a meglio comprendere il concetto di *Qualità* possiamo prendere in considerazione l'elenco delle componenti della *Qualità* proposto da Garvin (1987):

- **Aspetti formali** (quali packaging, estetica e design)
- **Affidabilità** (cioè la frequenza al guasto)
- **Conformità alle normative** (il prodotto costruito è l'esatta esecuzione del progetto)
- **Durata del servizio** effettivo del prodotto
- **Funzionalità** (capacità del prodotto di compiere funzioni diverse)
- **Manutenibilità** ossia con quale grado di facilità il prodotto può essere riparato
- **Prestazione del servizio** (in che modo esso svolge le funzioni richieste del cliente)
- **Livello di qualità percepito** (in funzione dell'esperienza di consumo e della reputazione dell'azienda o del prodotto)

Come dicevamo, la *Qualità* è presente in svariate situazioni produttive quindi non ha un'unica definizione. Nell'ottica delle normative europee per la creazione di un Sistema di Gestione della Qualità essa è definita come "*la soddisfazione delle esigenze del cliente*" quindi come la condizione del bene o servizio in questione di "*essere appropriato all'uso*". A questa concezione viene preferita un'altra, più specifica: "*la qualità è inversamente proporzionale rispetto alla variabilità²*", perciò un prodotto le cui caratteristiche rilevanti presentano una bassa variabilità è considerato un prodotto di qualità.

² Variabilità non desiderata.

Nell'approccio strategico di miglioramento della qualità il management dell'impresa produttrice del bene o servizio si applica per la "*riduzione della variabilità sia nel processo che nel prodotto*" in una visione d'insieme di tutti i soggetti coinvolti da e nell'impresa³.

Ecco quindi entrare in gioco la statistica nell'applicazione di metodi, tecniche e strumenti per la riduzione di variabilità che nel loro insieme danno vita al Controllo Statistico della Qualità (risalente almeno ai primi decenni del XX secolo con la pubblicazione del testo di W.A.Shewhart, *Economic Control of Quality of Manufactured Product*, 1931).

I suoi aspetti fondamentali sono :

- il campionamento in fase di accettazione
- il controllo statistico del processo di produzione
- la programmazione degli esperimenti⁴.

In questa tesi tratteremo di un particolare strumento per il controllo del processo, SPC⁵.

³ Mi riferisco al Total Quality Management, TQM.

⁴ Per individuare quali fattori, e in che misura, hanno influenzato il concretizzarsi del livello di qualità osservato.

⁵ Statistical Process Control, appunto.

Capitolo 1 : Controllo Statistico del Processo

1.1 Le Carte di Controllo

È più facile rendere stabile, ripetibile e migliore un processo produttivo che verificare la qualità di ciascun prodotto finito, per questo l'attenzione è rivolta sull'individuazione e la riduzione della variabilità dei fattori coinvolti nella produzione. Il controllo statistico di processo ha come obiettivo primario individuare il più rapidamente possibile la presenza di fattori specifici così da poter avviare tempestivamente azioni correttive.

Per fare ciò vengono in nostro aiuto i "magnifici sette", ovvero gli strumenti statistici di base del controllo statistico del processo applicabili ad ogni tipo di processo. Essi sono:

- Istogrammi e grafici "steam-leaf"⁶
- Fogli di controllo
- Grafici di Pareto
- Diagrammi causa ed effetto
- Diagrammi sulla concentrazione dei difetti
- Grafici a dispersione
- Carte di controllo.

⁶ Grafici a rami e foglie.

Lo strumento di nostro interesse è la Carta di controllo (ideato negli anni Venti da Walter A. Shewhart), che è il più semplice per definire una procedura di SPC. I dati riguardanti le caratteristiche della qualità vengono classificati in attributi, ovvero caratteristiche non misurabili su scala continua né su scala quantitativa, e variabili, rappresentabili su una scala continua di valori e descritti da una misura di centralità e una di variabilità. Esistono perciò carte di controllo per attributi e per variabili. I tre scopi fondamentali delle Carte di Controllo sono:

- Riduzione della variabilità del processo
- Monitoraggio e controllo del processo
- Stima dei parametri del prodotto e del processo⁷.

La larga diffusione di questo strumento è dovuta a cinque principali ragioni:

- migliorano la produttività poiché attraverso la riduzione degli sprechi aumentano l'efficienza e, a sua volta, la capacità produttiva⁸
- prevengono efficacemente la produzione di pezzi difettosi, eliminando quindi gli scarti e il ritiro della merce
- evitano inutili e dannosi aggiustamenti al processo in caso di variazioni anomale
- forniscono, in genere, valide informazioni diagnostiche per interventi preventivi sul processo
- procurano indicazioni sulla capacità e sulla stabilità nel tempo del processo utili in fase di progettazione.

⁷ Come media, deviazione standard, percentuale di pezzi non conformi, etc...

⁸ Intesa come numero di pezzi accettabili per unità di tempo.

Indipendentemente da come sia stato progettato, un processo produttivo è comunque soggetto ad una certa variabilità intrinseca, risultato dell'effetto cumulato di svariati fattori (casuali o costanti) piccoli ma ineliminabili.

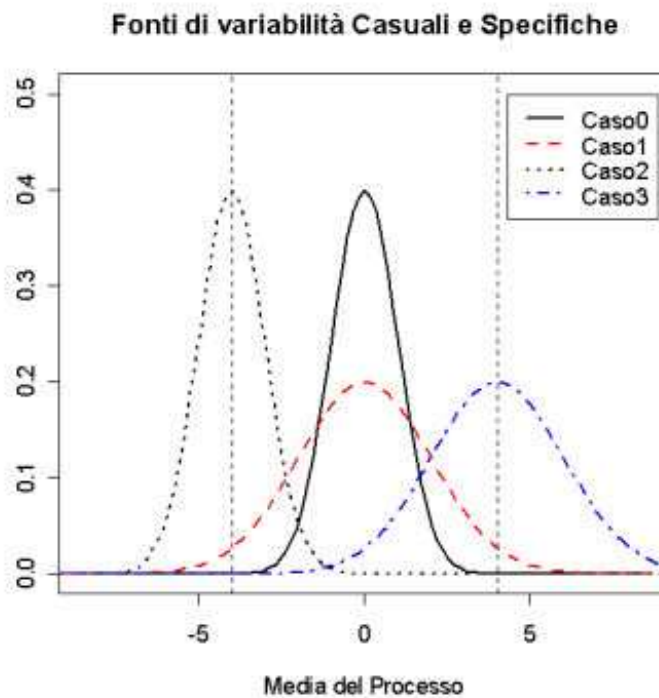


Figura 1

Un processo la cui variabilità è causata solamente da fattori casuali è detto "*sotto controllo*" e la maggior parte dei valori della grandezza in esame sono compresi all'interno dei limiti di specifica USL e LSL (Upper Specific Limit e Lower Specific Limit). Esistono poi dei fattori specifici quali macchinari malfunzionanti, errori umani o materiali difettosi che influiscono sul processo in modo evidente generando prodotti non conformi. Tale processo è detto "*fuori controllo*". In Figura 1, nel Caso0 il processo è sotto controllo, negli altri tre casi invece è fuori controllo a causa di una variazione rispettivamente della varianza, della media e di entrambe.

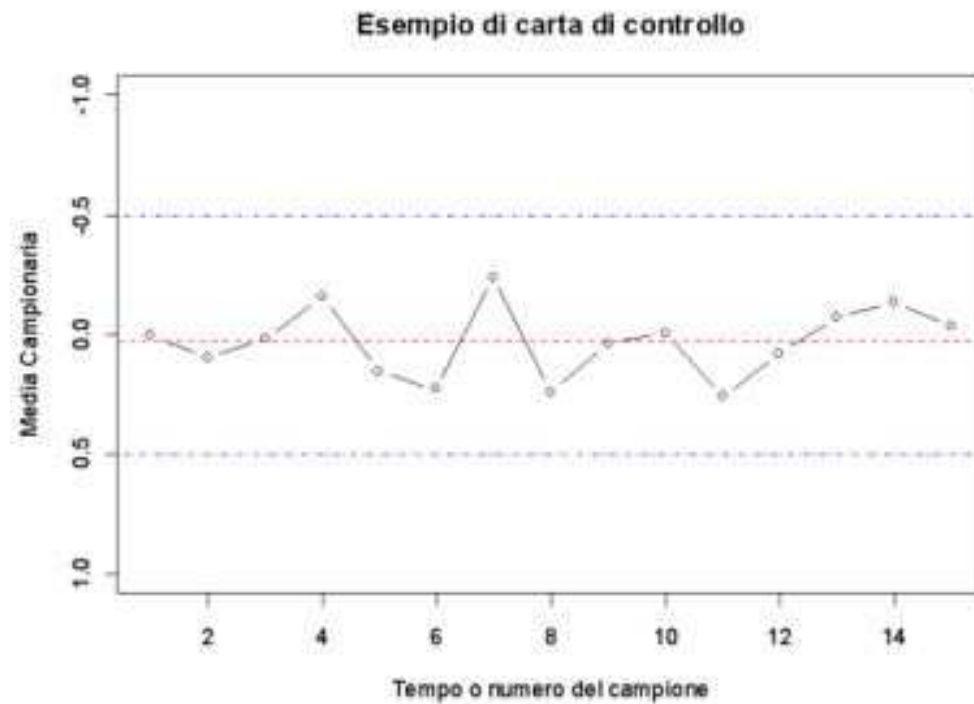


Figura 2

Un esempio di carta di controllo è riportato in Figura 2. La linea centrale, CL, rappresenta il valor medio della qualità, che è solitamente il valore desiderato. Le altre due linee orizzontali sono chiamate UCL e LCL, rispettivamente Upper e Lower Control Limit. Operando sulla posizione dei limiti nella carta di controllo si può variare il grado di incertezza delle affermazioni ricavabili dalle osservazioni. Ad esempio, ponendo i limiti a $\pm 1.645\sigma$, il 90.0% delle osservazioni sarà interno all'intervallo e la probabilità di rilevare valori esterni anche se il processo è sotto controllo sarà dunque del 10%⁹.

⁹ Assumendo che la caratteristica di qualità si distribuisca normalmente.

Possono presentarsi quattro diversi casi di processo i quali, a parte quello "ideale"¹⁰, vanno migliorati:

- processo stabile che produce sempre output entro i limiti di specifica
- processo stabile che produce una quantità significativa di output oltre i limiti di specifica
- processo instabile che produce fortunatamente output entro i limiti di specifica
- processo instabile che produce spesso output fuori dai limiti di specifica.

Le osservazioni campionarie vengono raccolte periodicamente e rappresentate nella carta di controllo; se il valore in oggetto cade tra i limiti di controllo e mostra un andamento casuale allora il processo è sotto controllo, se invece si osserva un valore al di fuori di questi due limiti il processo verrà considerato fuori controllo e verranno intraprese azioni correttive per individuare ed eliminare le cause. La dimensione campionaria e la frequenza di campionamento vengono definiti in fase di progettazione della carta di controllo: come è facile immaginare, quanto più grande è la numerosità del campione tanto più sarà facile individuare un piccolo scostamento del processo. Analizzare grandi campioni con elevata frequenza non è conveniente dal punto di vista economico; è più ragionevole invece esaminare frequentemente piccoli campioni oppure grandi campioni ad intervalli più larghi in termini di tempo o di lotti di produzione. Per calcolare correttamente la dimensione campionaria e la frequenza di campionamento è possibile usare due strumenti:

- la lunghezza media delle sequenze (**ARL** *Average Run Length*)
- il tempo medio al segnale (**ATS** *Average Time to Signal*).

¹⁰ Il primo dell'elenco.

1.1.1 Carte di Controllo di Shewhart

Come dicevamo precedentemente, Shewhart è stato un precursore del Controllo Statistico del Processo e le sue Carte di Controllo costituiscono il fondamento della teoria applicata a tutte le carte di controllo. Le carte di controllo di Shewhart sono carte "senza memoria" dato che considerano solo le informazioni correnti e non quelle passate, ossia tratta i campioni in modo indipendente l'uno dall'altro. Il processo viene sorvegliato utilizzando una trasformazione della caratteristica da monitorare e i limiti sono calcolati in base alla stima della varianza delle osservazioni. Oltre a questi limiti ne vengono calcolati altri posti più vicino alla media del processo e considerati limiti d'allerta. Infatti in alcune versioni della carta il processo viene considerato fuori controllo se un certo numero di osservazioni successive si trova oltre la stessa soglia di allerta¹¹.

Le carte di controllo di Shewhart possono essere per:

- variabili
- il valore medio
- lo scarto medio
- il range medio
- misure singole
- attributi
- frazione di non conformi
- frazione di non conformità per unità.

L'ARL è il numero medio di valori che devono essere osservati prima di rilevare un valore esterno ai limiti di controllo¹² e segnalare quindi un allarme.

¹¹ Indicativo di uno scostamento del processo.

¹² Indicativo di un andamento del processo fuori controllo.

Moltiplicando l'ARL per l'intervallo medio di tempo tra due campioni si ottiene l'ATS, ovvero il tempo medio che intercorre tra due segnali di fuori controllo.

La situazione ideale presuppone un ARL più grande possibile quando il processo è sotto controllo e più piccolo possibile in caso di fuori controllo. Per un'impresa, infatti, è fondamentale avere il minor numero possibile di falsi allarmi e rilevare tempestivamente lo scostamento del processo. Dipende però dalle singole realtà produttive perché per un'impresa potrebbe risultare più conveniente avere molti falsi allarmi ma avere pochissimi lotti difettosi mentre, per un'altra, potrebbe essere molto più costoso verificare la conformità dei prodotti. Si tratta di un trade-off che affronteremo più volte successivamente. Se nessuna osservazione cade al di fuori dei limiti di controllo ma sono avvenuti dei cambiamenti nel processo la probabilità di **non** rilevare questi cambiamenti altro non è che la probabilità di commettere un errore del secondo tipo:

$$\beta = P\{LCL \leq \bar{x} \leq UCL | \mu = \mu_1 = \mu_0 + k\sigma\}$$

Perciò l'ARL risulta essere:

$$ARL = \sum_{r=1}^{\infty} r \beta^{r-1} (1 - \beta) = \frac{1}{(1 - \beta)}$$

Ignorando tutte le osservazioni precedenti, la carta di controllo di Shewhart è poco sensibile a piccoli scostamenti del processo, di ampiezza solitamente non superiori a 1.5σ . È quindi un buon strumento per grandi scostamenti ma non per piccoli.

1.1.2 Carte di Controllo Exponentially Weighted Moving Average

Ovvero carta di controllo a medie mobili pesate esponenzialmente (EWMA). È stata introdotta da Roberts (1959) ed è definita come:

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

dove y_i è una trasformazione della caratteristica di interesse; i è l'indice temporale della caratteristica, il parametro λ è un valore compreso tra 0 e 1; il valore iniziale richiesto dal primo campionamento per $i=1$, z_0 , è generalmente la media del processo.

Sostituendo ricorsivamente nella precedente i valori di z_{i-1} si ottiene

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

in cui i pesi $\lambda (1 - \lambda)^j$ decrescono in progressione geometrica e la loro somma vale uno. Dato che è una media ponderata di tutte le osservazioni dipende meno dall'assunzione di normalità ed è quindi lo strumento ideale nel caso di osservazioni individuali. Utilizzando la carta EWMA per un periodo sufficientemente elevato, la varianza di z_i che è pari a

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

ed i limiti si possono approssimare a

$$CL^{13} = \mu_0 \pm k\sigma \sqrt{\frac{\lambda}{2 - \lambda}}$$

dove k è un parametro d'ampiezza che dipende dalla situazione.

¹³ Limite di controllo, non linea centrale.

Un pregio della carta di controllo EWMA è la sua versatilità, infatti può essere utilizzata in diversi processi a seconda delle necessità poiché variando λ si controllano i pesi per le osservazioni passate. Infatti, ricordando la statistica :

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

all'aumentare del valore di λ viene data meno importanza alle osservazioni passate fino al caso in cui $\lambda = 1$ per cui la statistica viene usata per osservazioni singole.

A differenza delle carte di controllo di Shewhart, trattando appunto anche i campioni precedenti, la carta EWMA permette di segnalare uno scostamento del processo individuando un cambiamento nel trend senza che il processo stesso sia oltre i limiti.

Nel caso in cui si effettua una correzione del processo la migliore stima della media del nuovo processo non è il valore corrente della statistica EWMA ma è la media delle osservazioni successive alla correzione. Inoltre non necessita di campioni numerosi ed è quindi uno strumento più conveniente in termini di tempo e denaro.

Esiste però un problema di applicazione abbastanza rilevante. Per ottenere una veloce individuazione di un piccolo scostamento del processo il valore ottimale di λ deve essere scelto piccolo, mentre per individuare uno scostamento più grande il valore ottimale di λ deve essere grande. Questo è il problema cosiddetto di *inerzia*; un piccolo valore di λ fa convergere lentamente la varianza di z_i perciò i valori anomali nella fase iniziale vengono difficilmente segnalati. Si tratta quindi di un ostico trade-off per l'impresa che dovrà scegliere quale situazione le conviene.

Tuttavia un certo grado di inerzia rimane poiché quando la statistica EWMA si trova nelle vicinanze di un limite essa reagirà lentamente ad uno scostamento del processo verso il limite opposto. Questa situazione è definita *worst case*.

1.1.3 Carte di Controllo Adaptive Exponentially Weighted Moving Average

Le carte di controllo a medie mobili pesate esponenzialmente (AEWMA) uniscono la velocità di reazione delle carte di Shewhart per grandi scostamenti alla robustezza delle carte EWMA, di cui sono un'estensione. La carta AEWMA è basata sulla statistica:

$$z_t = z_{t-1} + \phi(e_t) \quad (1)$$

L' "errore" e_t è lo scarto tra il valore corrente y_t e la statistica z_{t-1} e $\Phi(e_t)$ è un'opportuna funzione. Fissata una soglia γ adatta ad assicurare il tempo medio desiderato tra falsi allarmi, la carta segnala un allarme quando

$$|z_t - \eta_0| > \gamma$$

dove η_0 è il valore obiettivo a cui ci riferivamo precedentemente. Nel caso in cui $y_t \neq z_{t-1}$ la (1) si può riscrivere come:

$$z_t = (1 - w(e_t)) z_{t-1} + w(e_t) y_t \quad \text{dove} \quad w(e) = \phi(e)/e$$

che non è altro che un EWMA con pesi variabili. Nei casi in cui $\Phi(e) = \lambda e$ e $\Phi(e) = e$ si ottengono le carte EWMA e Shewhart rispettivamente. Per riuscire ad ottenere i pregi di ambedue le carte la funzione $\Phi()$ viene scelta tale che sia:

- monotona crescente in e
- simmetrica
- $\Phi(e)/e \approx 1$ quando e è grande in valore assoluto
- $\Phi(e) \approx \lambda e$ quando e è piccolo in valore assoluto.

Quindi quando l'errore è grande la statistica si comporta come una Shewhart e può risolvere il problema dell'inerzia iniziale (Capizzi e Masarotto, 2003)¹⁴.

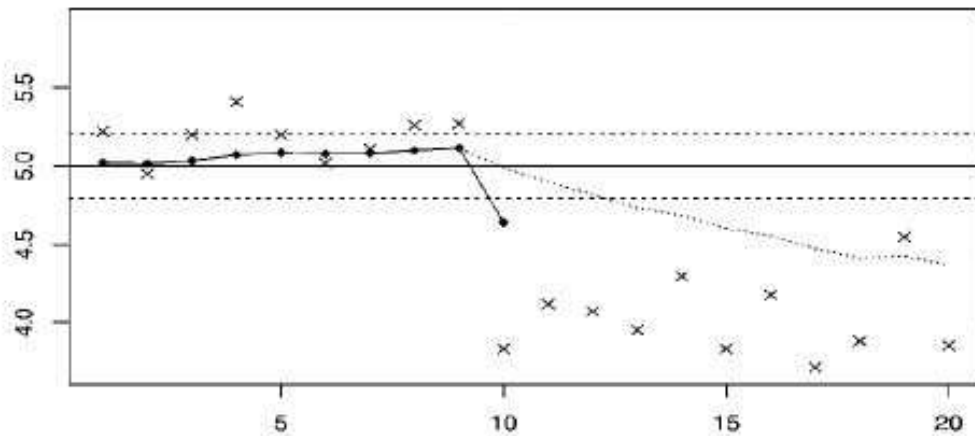


Figura 3

Come possiamo notare dall'esempio in Figura 3, la carta di controllo AEWMA, in linea continua, rileva subito lo scostamento del processo nel *worst case* a differenza della carta EWMA con $\lambda = 0.1$, in linea tratteggiata.

I "difetti" della carta EWMA, come abbiamo visto in precedenza, sono dovuti all'impossibilità di costruire un'unica carta che riesca a segnalare con tempestività sia il piccolo che il grande scostamento del processo e che riesca a stimare il vero livello corrente del processo nel *worst case*.

Il criterio dell'AEWMA, proposto da Capizzi e Masarotto, riesce nell'intento di costruire un'unica carta capace di segnalare con rapidità una grande varietà di scostamenti utilizzando diverse funzioni per $\Phi()$. A confronto con altre carte, perciò, i suoi pregi sono la semplicità e l'efficienza.

¹⁴ Vedi bibliografia.

Una esempio di funzione $w_t = \Phi(e)/e$ può essere :

$$w_t = 1 - (1 - \lambda) \frac{1}{\max(1, |y_t - z_{t-1}|/\gamma)} \quad \text{dove} \quad z_t = (1 - w_t) z_{t-1} + w_t y_t$$

ed è quella che verrà usata per le prove in questa trattazione.

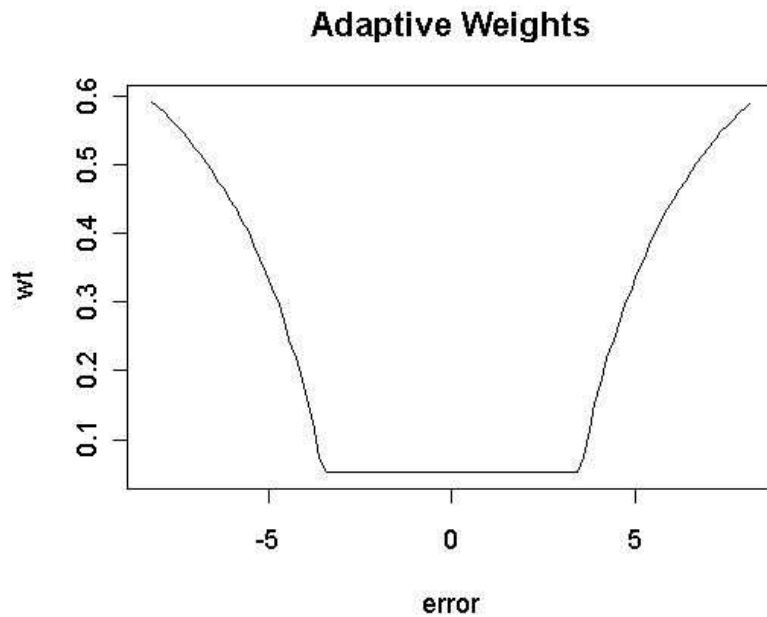


Figura 4

Dal grafico della funzione w_t in Figura 4 si può notare che, se il valore assoluto dell'errore $|y_t - z_{t-1}|$ è minore del valore soglia γ , allora i pesi w_t sono pari a λ , poiché $\Phi(e) = \lambda e$. Se invece è maggiore del valore soglia γ , allora i pesi w_t aumentano di valore all'aumentare dell'errore, così da riuscire ad "aggiornare" più velocemente il livello della statistica z_t .

Capitolo 2 : Sorveglianza della Capacità.

2.1 Un'introduzione al significato di Capacità del Processo

Un altro utile strumento usato nelle fasi di progettazione, analisi e valutazione della prestazione del processo per poter eliminare o quanto meno ridurre la variabilità in relazione ai livelli nominali di specifica è la capacità. La valutazione della capacità si basa sull'analisi della distribuzione della caratteristica di qualità in relazione ai limiti di specifica, alla media e alla variabilità. In particolare, la variabilità della caratteristica d'interesse può essere interpretata come istantanea (o naturale) oppure rispetto al tempo. Ricordando che il 99.73% dei valori di una distribuzione Normale sono compresi nell'intervallo $\mu-3\sigma$ e $\mu+3\sigma$, il rapporto tra processo conforme ai limiti di specifica (USL e LSL) e capacità del processo viene detto *process capability ratio*:

$$C_p = \frac{U - L}{6\sigma}$$

Generalmente la deviazione standard del processo, σ , e il valor medio, μ , devono essere stimati dato che sono incogniti e non decisi in fase di progettazione. È anche possibile calcolare specifiche unilaterali:

- specifica superiore $CU = (USL - \mu)/(3\sigma)$
- specifica inferiore $CL = (\mu - LSL)/(3\sigma)$.

Esistono quattro diverse misure della Capacità a seconda della sensibilità agli scostamenti della media del processo e prendono il nome di indici:

$$C_p = (U - L)/6\sigma$$

Misura la dispersione delle specifiche in rapporto alla dispersione del processo ma non tiene conto della posizione della media; non può essere negativo ed è chiamato anche indice di capacità potenziale.

$$C_{pk} = \min(U - \mu, L - \mu)/3\sigma .$$

Considera la centratura del processo e può essere negativo. Se $C_{pk} = C_p$ allora il processo è centrato; tanto più C_{pk} è minore di C_p tanto più il processo non è centrato. È chiamato anche indice di capacità effettiva.

$$C_{pm} = (U - L)/6\sqrt{\sigma^2 + (\mu - T)^2} \quad \text{dove} \quad T = (U + L)/2 .$$

Migliore del precedente perché, oltre a tener conto della centratura, non aumenta per piccoli valori di sigma .

$$C_{pmk} = \min(U - \mu, L - \mu)/3\sqrt{\sigma^2 + (\mu - T)^2}$$

Costruito a partire dai precedenti è il più sensibile. Vale zero quando il processo è centrato su un limite di specifica ed è limitato inferiormente in -0.333.

Senza entrare nei dettagli, Vännman ha definito una famiglia di indici di capacità, dipendenti da due parametri u e v , che riunisce i quattro precedenti in unica scrittura:

$$C_p(u, v) = \frac{d - u|\mu - T|}{3\sqrt{\sigma^2 + (\mu - T)^2}}$$

dove

$$C_p(0, 0) = C_p \quad C_p(1, 0) = C_{pk} \quad C_p(0, 1) = C_{pm} \quad C_p(1, 1) = C_{pmk} .$$

Un processo con un indice di capacità $C_{p..}$ minore di 1 è un processo non capace, con un indice di capacità $C_{p..}$ compreso tra 1 e 1.5 è marginalmente capace e con un indice di capacità $C_{p..}$ maggiore di 1.5 è capace. È possibile quindi effettuare verifiche d'ipotesi sul valore della capacità del processo. L'analisi di capacità del processo è fondamentale in un programma di miglioramento della *qualità* complessivo. L'obiettivo è quello di utilizzare i dati al fine di:

- prevedere come il processo rispetterà le tolleranze
- assistere gli addetti alla Ricerca & Sviluppo
- stabilire l'intervallo di campionamento per le procedure di sorveglianza
- definire i requisiti di prestazione di nuove attrezzature
- selezionare i fornitori
- supporto alla pianificazione della produzione
- ridurre la variabilità del processo.

Le principali tecniche usate sono istogrammi o carte di probabilità, carte di controllo e programmazione degli esperimenti.

2.2 Sorveglianza della Capacità

Presupposto dell'analisi della capacità del processo è che questo sia stabile ossia non siano presenti cause di variabilità specifiche. In alcune situazioni però può essere d'interesse monitorare la capacità sebbene il processo sia instabile.

Ad esempio aziende con più fornitori per lo stesso input, le quali, a causa delle diversità di questi ultimi, nelle tecnologie usate e nel grado di qualità del prodotto, incorrono nel caso di processi non capaci ma sotto controllo perché le variazioni della distribuzione della caratteristica di qualità sono piccole rispetto alle tolleranze.

Un altro esempio può essere il caso di particolari processi produttivi in cui le specifiche variano per ogni pezzo e quindi non è possibile considerare l'andamento di una caratteristica di qualità se non la capacità del processo.

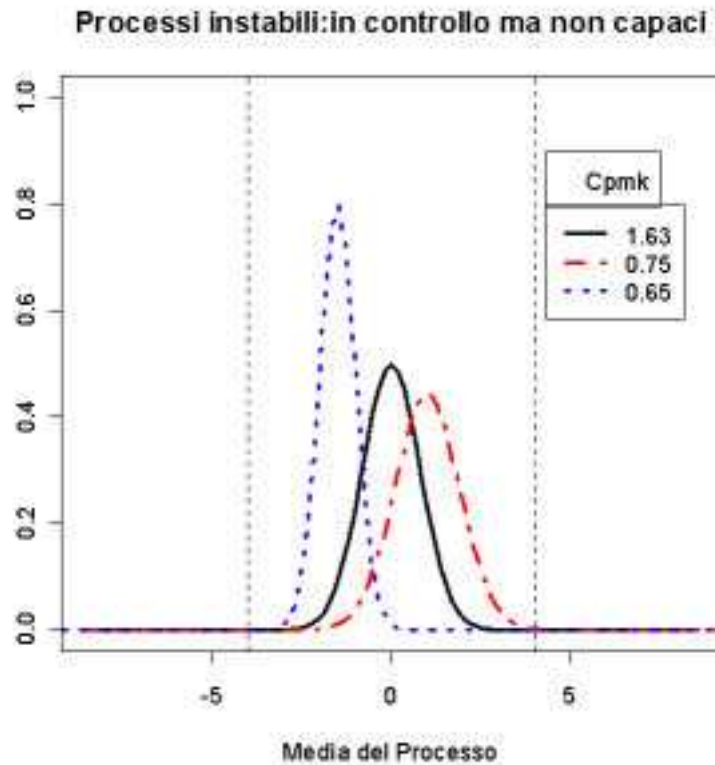


Figura 6

In figura 6 è mostrato un esempio di processo instabile; le tre diverse distribuzioni “simulano” il processo in tre situazioni differenti. Non ci sono situazioni di fuori controllo ma i processi riportati in linea tratteggiata non sono capaci, al contrario di quello in linea continua; essi infatti hanno un indice di capacità minore di 0.8 mentre quello in linea continua è maggiore di 1.6 .

L'instabilità dovuta alle caratteristiche del materiale acquistato dall'impresa non può essere eliminata direttamente dalla stessa; d'altronde, sorvegliare il processo con le carte classiche aumentandone la sensibilità produrrebbe troppi falsi allarmi.

Per valutare la capacità di un processo malgrado la sua instabilità¹⁵ **Castagliola** e **Vännman**¹⁶ hanno recentemente proposto una carta di controllo EWMA per sorvegliare la capacità attraverso gli indici di Vännman. Inoltre hanno dimostrato che la loro procedura individua con efficienza sia incrementi che decrementi del livello di capacità del processo.

La progettazione della carta ha come presupposto la normalità della distribuzione della caratteristica di interesse e che la capacità del processo sia inizialmente costante. Dopo aver stimato la media e la deviazione standard del processo vengono calcolati gli indici di capacità $\hat{C}_p(u,v)_i$ che sono variabili casuali indipendenti. Dato che la distribuzione di $\hat{C}_p(u,v)_i$ è asimmetrica si applica una trasformazione logaritmica $Y_i = a + b \ln(\hat{C}_p(u,v)_i)$ dove a e b sono dei parametri che dipendono dalla distribuzione di $\hat{C}_p(u,v)_i$ in controllo:

$$b = \left(\ln \left(\frac{V(\hat{C}_p(u,v))}{E^2(\hat{C}_p(u,v))} + 1 \right) \right)^{-1/2}$$

$$a = -b \ln \left(E(\hat{C}_p(u,v)) \left(\frac{V(\hat{C}_p(u,v))}{E^2(\hat{C}_p(u,v))} + 1 \right)^{-1/2} \right)$$

Si monitora Y_i con la statistica EWMA:

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

e i limiti per monitorare la $\hat{C}_p(u,v)_i$ sono:

$$CL = E(Y_i) \pm k \left(\frac{\lambda}{2 - \lambda} \right)^{1/2} \sigma(Y_i)$$

dove k è una soglia appropriata determinata per avere un numero accettabile di falsi allarmi.

¹⁵ Non necessariamente, può essere anche stabile.

¹⁶ Vedi bibliografia.

Capitolo 3 : Confronto tra gli schemi EWMA e AEWMA

3.1 Introduzione

L'idea di questa trattazione è di verificare se, come nelle altre situazioni, anche nel caso della sorveglianza della capacità lo schema AEWMA sia più efficiente dello schema EWMA per la sua capacità di migliorarne i difetti, ovvero, di segnalare più rapidamente un allarme sia nel caso del piccolo che del grande scostamento con un solo λ ottimale.

Fissati¹⁷ quindi, i parametri λ e γ si cerca, tramite simulazione per via software, di studiare la sensitività della carta, in termini di ARL di fuori controllo, per la combinazione di valori scelta del livello di significatività α , e degli scostamenti di μ e σ .

3.2 Metodo di calcolo

Seguendo la procedura proposta da Castagliola e Vännman dopo aver simulato un processo sotto controllo, la cui caratteristica di qualità si distribuisce normalmente, ed aver ottenuto N campioni di n osservazioni ciascuno si è calcolato la stima di massima verosimiglianza di μ_i e σ_i e la capacità del processo per ogni campione. Dopodiché si è ricavato il valore di a e b e applicato la trasformazione logaritmica al vettore di indici \hat{C}_p ottenendo $Y_t = a + b \ln(\hat{C}_p)$. Con i valori di riferimento del processo sotto controllo a , b , μ_Y e σ_Y si è proceduto a definire il quantile k in modo tale che:

$$P \left(\max_{1 \leq t \leq N_0} \left| \frac{z_t - \mu_Y}{\sigma_Y \sqrt{\lambda/(2-\lambda)}} \right| \leq k \right) = 1 - p_0$$

¹⁷ Di fatto ne proveremo alcuni.

dove:
$$z_t = (1 - w_t) z_{t-1} + w_t y_t$$

e
$$k = \Phi^{-1}(1 - p_0)$$

per cui le carte segnalano un allarme quando:

$$|z_t| \leq \mu_Y + k \sigma_Y \sqrt{(\lambda / (2 - \lambda))}$$

quindi è
$$P(\text{RunLength} \leq N_0) = 1 - p_0 \quad .$$

In altre parole la probabilità che il massimo valore della statistica z_t sia minore della quantità k è la probabilità che il numero di campioni necessario per segnalare un allarme sia minore di N_0 , ossia il numero di campioni effettuati.

I valori di riferimento del processo sotto controllo μ_Y e σ_Y sono usati per calcolare il quantile, inizializzare le statistiche EWMA e AEWMA e ottenere i limiti della carta di controllo.

Come accennato in precedenza la funzione scelta per i pesi è:

$$w_t = 1 - (1 - \lambda) \frac{1}{\max(1, |Y_t - z_{t-1}| / \gamma)}$$

la quale, come è facile notare, nel caso in cui $\gamma = \infty$ rende z_t una statistica EWMA, mentre, quando a γ viene assegnato un valore appropriato, z_t è una statistica AEWMA. Il valore scelto per la soglia di allarme γ , con il quale si definisce il grado di reattività della statistica AEWMA e se ne regola la prestazione, è 3.5 . Considerando che la deviazione standard σ_Y fissata pari ad 1 nell'utilizzo della carta AEWMA, e che z_t ha una deviazione standard trascurabile (meno di 0.5) con la regola del 3σ ¹⁸ si dovrebbe riuscire a “pareggiare” il valore assoluto dell'errore così da risolvere in parte il problema dell'inerzia della carta EWMA. Infatti fino a che l'errore e non supera, in valore assoluto, la soglia $\gamma = 3.5$ la carta AEWMA si comporta come una carta Shewhart.

¹⁸ Il 99.73% dei valori di una distribuzione Normale sono compresi nell'intervallo $\mu - 3\sigma$ e $\mu + 3\sigma$.

Presupposto della carta di controllo AEWMA è che la Y_t abbia forma simmetrica. Per cui abbiamo deciso di utilizzare

$$Y_t = \Phi^{-1} \left(F \left(\hat{C}_p \right) \right) \sim N(0,1)$$

dove F è la funzione di ripartizione empirica.

Nella pagina seguente in Figura 7 è riportato un esempio della statistica Y_t e della statistica AEWMA z_t con $\gamma = 3.5$ e $\lambda = 0.05$. Nel primo caso il processo è in controllo e la trasformata Y_t e la statistica z_t hanno una forma abbastanza simmetrica; nel secondo caso, quello di piccolo scostamento, le due statistiche hanno ancora una forma abbastanza simmetrica mentre nel terzo caso, quello di grande scostamento, le due statistiche sono molto asimmetriche.

Nell'utilizzo della carta AEWMA per la statistica Y_t i valori di μ_Y e σ_Y vengono fissati rispettivamente a 0 e a 1.

Con il valore del quantile si calcolano poi i limiti della carta e si procede ad una seconda simulazione, questa volta del processo fuori controllo, e si ricavano i valori della statistica ARL.

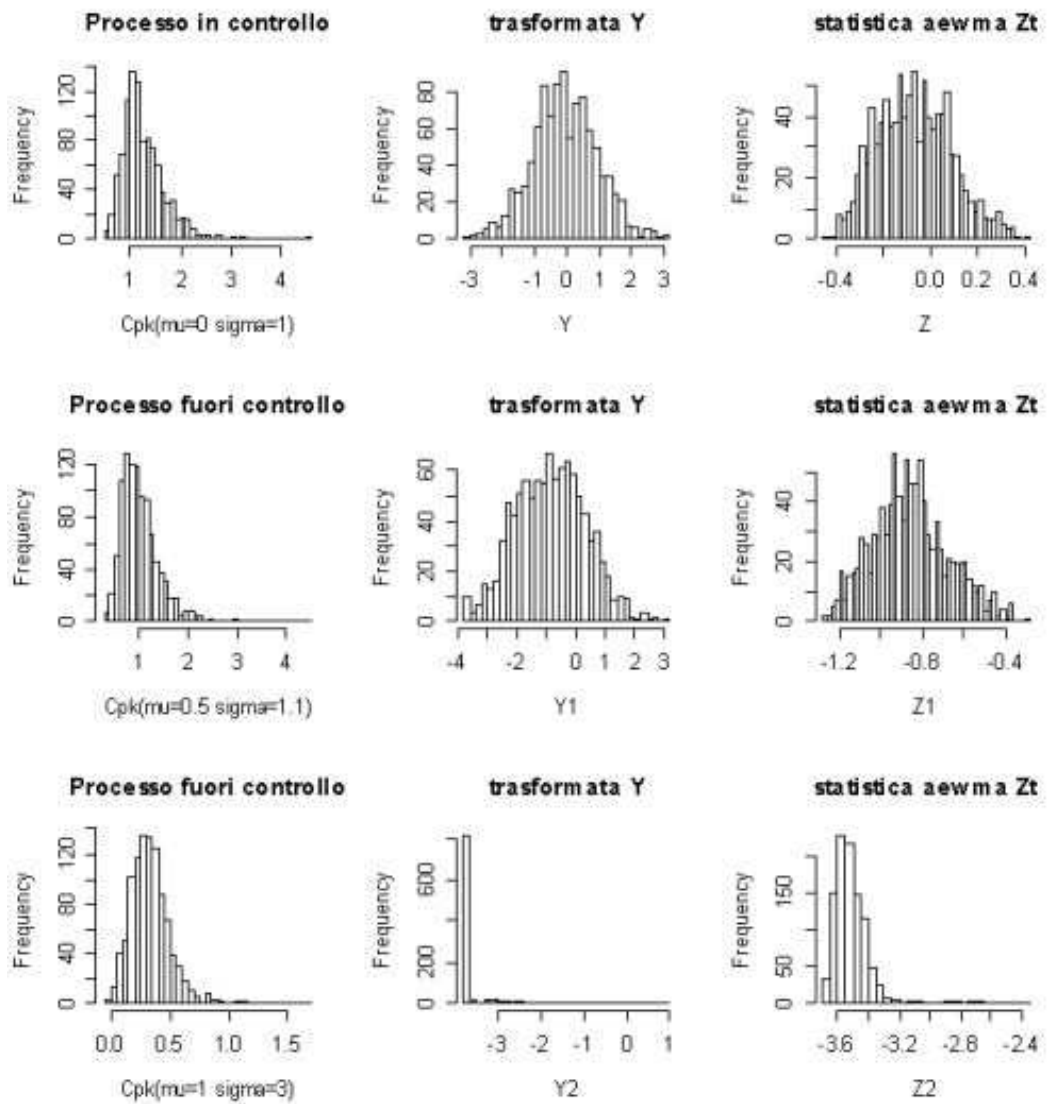


Figura 7

Nelle prove effettuate il processo considerato in controllo è la distribuzione normale standard, il numero dei campioni è $N=1000$ e il valore di $p_0 = \alpha = 0,05$. Il numero di ripetizioni della simulazione è $B=2000$ e l'indice scelto è il $\hat{C}_p(1,1)$ che corrisponde a:

$$C_{pmk} = \min(U - \mu, L - \mu) / 3\sqrt{\sigma^2 + (\mu - T)^2} .$$

Nella Tabella 1 sono riportati i limiti di controllo e nelle successive si possono osservare i risultati della simulazione dell'ARL di fuori controllo per le carte EWMA e AEWMA per diverse combinazioni della numerosità dei campioni n (5, 7, 10) dei parametri λ (0.05, 0.3) e γ (3, 3.5, 4) e di casi di scostamento che sono:

- piccolo $\mu = 0.5$ e $\sigma = 1.1$
- medio $\mu = 1$ e $\sigma = 1.5$
- grande $\mu = 1$ e $\sigma = 3$.

3.3 Risultati

Tabella 1

Control Limit		for $\alpha = 0.05$			
n	EWMA		AEWMA		
	$\lambda = 0.05$	$\lambda = 0.3$	$\gamma = 3$	$\gamma = 3.5$ $\lambda = 0.05$	$\gamma = 4$
5	0.438	1.390	0.581	0.492	0.493
7	0.457	1.447	0.658	0.488	0.490
10	0.475	1.452	0.690	0.490	0.494

Tabella 2

Average Run Length						for $\mu = 0.5; \sigma = 1.1$ $Cp(1,1) = 1$					
EWMA			AEWMA								
n	-	-	$\gamma = 3$	$\gamma = 3.5$	$\gamma = 4$						
	$\lambda = 0.05$	$\lambda = 0.3$				$\lambda = 0.05$					
5	23.900	119.630	25.918	21.729	22.129						
7	18.185	49.026	21.759	16.080	16.318						
10	13.709	18.758	14.924	12.188	12.618						

Tabella 3

Average Run Length						for $\mu = 1; \sigma = 1.5$ $Cp(1,1) = 0.59$					
EWMA			AEWMA								
n	-	-	$\gamma = 3$	$\gamma = 3.5$	$\gamma = 4$						
	$\lambda = 0.05$	$\lambda = 0.3$				$\lambda = 0.05$					
5	6.042	4.733	3.388	5.193	5.503						
7	4.829	3.373	2.899	3.959	4.625						
10	3.862	2.432	1.812	3.227	4.134						

Tabella 4

Average Run Length						for $\mu = 1; \sigma = 3$ $Cp(1,1) = 0.32$					
EWMA			AEWMA								
n	-	-	$\gamma = 3$	$\gamma = 3.5$	$\gamma = 4$						
	$\lambda = 0.05$	$\lambda = 0.3$				$\lambda = 0.05$					
5	3.223	2.108	1.360	2.889	3.640						
7	2.733	1.722	1.185	2.474	3.314						
10	2.320	1.400	1.076	2.127	3.065						

3.4 Conclusioni

Premettendo che le prove di questa breve trattazione non pretendono di coprire la vasta gamma di situazioni che possono presentarsi, dalle poche prove effettuate, per quanto riguarda la statistica EWMA, i risultati confermano l'impossibilità di ottenere un piccolo valore della statistica ARL con un unico λ ottimale. Guardando la Tabella 1 si può osservare che i limiti per la carta EWMA con $\lambda = 0.05$ sono simili a quelli per la carta AEWMA per $\gamma = 3.5$ e $\gamma = 4$ e i limiti per la carta EWMA con $\lambda = 0.3$ sono simili a quelli per la carta AEWMA per $\gamma = 3$.

Come era prevedibile il piccolo scostamento ($\mu = 0.5$ e $\sigma = 1$) viene individuato velocemente dalla statistica EWMA con λ piccolo ($\lambda=0.05$) mentre per λ grande ($\lambda=0.3$) i valori della statistica ARL sono molto più alti. Viceversa per gli altri due scostamenti i valori della statistica ARL per λ grande sono minori di quelli per λ piccolo. Nel complesso si può notare come, per entrambe le statistiche, più è numeroso il campione, più piccolo è il valore della statistica ARL. Per quanto riguarda la statistica AEWMA sembrerebbe che essa "soffra" del trade-off della statistica EWMA nei confronti però del parametro soglia γ .

Infatti per $\gamma = 3$ i valori della statistica ARL sono inferiori rispetto agli stessi per γ pari a 3.5 e 4 tranne che per lo scostamento piccolo dove invece sono molto più alti. Difatti i risultati in Tabella 1 mostravano una somiglianza nei valori dei limiti per la carta EWMA con $\lambda = 0.3$ e la carta AEWMA con $\gamma = 3$. I valori della statistica ARL per γ pari a 4 sembrano essere sempre leggermente superiori a quelli corrispondenti a $\gamma = 3.5$. Questi ultimi sono sempre leggermente minori dei rispettivi valori della statistica EWMA con $\lambda = 0.05$ e sono leggermente superiori a quelli con $\lambda = 0.3$ tranne che nel caso del piccolo scostamento in cui sono molto inferiori.

Per quanto si evince da questi risultati sembrerebbe che lo schema AEWMA sia più efficiente, anche se di poco, dello schema EWMA.

Appendice

Codice Sorgente

Di seguito vado a mostrare il codice del software statistico R usato per la simulazione del processo, il calcolo degli indici di capacità, del quantile, e della statistica ARL.

```
> simulaCp<-function(n,mu1,sigma1,N=1000,tau=(N+1),mu2=2,
+ sigma2=4,u=1,v=1,USL=4,LSL=-4){
+ d=(USL-LSL)/2
+ T=(USL+LSL)/2
+ Cp<-double(N)
+ X<-cbind(matrix(rnorm(n*(tau-1),mu1,sigma1),n,(tau-1)),
+ matrix(rnorm(n*(N-tau+1),mu2,sigma2),n,(N-tau+1)))
+ S<-apply(X,2,function(s) c(mean(s),sd(s)))
+ for (i in 1:N){
+ Cp[i]=(d-u*(abs(S[1,i]-T)))/(3*sqrt(S[2,i]^2+v*(S[1,i]-T)^2))}
+ Cp}
```

Questa funzione simula N campioni di n osservazioni normali da un processo che può essere in controllo, fuori controllo o che abbia subito uno scostamento. Calcola poi media e deviazione standard per ogni campione e riporta come output uno dei quattro indici di capacità della famiglia di indici di Vännman. Sono poi fissati i parametri che variano a seconda dei casi di interesse.

```
> n=7
> B=2000
> lambda=0.3
> alpha=0.05
> ad=0
```

```

> gamma=Inf
> mu1=0
> sigma1=1
> mu1_2=0.5
> sigma1_2=1.1

```

La costante `ad` è usata come variabile indicatrice per applicare o no lo schema AEWMA; `ad` è un'abbreviazione di “adaptive”.

```

> Cp_ic<-simulaCp(n,mu1,sigma1,N=10000)
> summary(Cp_ic)

```

Viene quindi simulato dapprima il processo sotto controllo per poter definire i parametri della carta di controllo. Con la funzione `summary(Cp_ic)` si possono osservare i valori di minimo, massimo, media, mediana, 1° e 3° quartile.

```

> aewma<-function(Y,lambda,gamma,z0){
+ Z<-double(length(Y))
+ z0=mY
+ l1=(1-lambda)
+ g1=1/gamma
+ for (g in 1:length(Y)){
+     e<-Y[g]-z0
+     w<-1-l1/max(1,abs(g1*e))
+     Z[g]<-z0<-z0+w*e}
+ Z}

```

È la funzione usata per applicare la statistica EWMA o AEWMA a seconda che i parametri `ad` e `gamma` valgano rispettivamente `ad = 0`, `gamma = Inf` oppure `ad = 1` e `gamma = 3` o `3.5` o `4`.

```

> pCp<-ecdf(c(-1000,1000,Cp_ic))

```

È la funzione di probabilità della ripartizione empirica del vettore di indici di capacità del processo in controllo.

```

> Ecp=mean(Cp_ic)
> Vcp=var(Cp_ic)
> g=1+Vcp/(Ecp^2)
> b=(log(g))^-0.5
> a=-b*log(Ecp*g^-0.5)
> Y=a+b*log(Cp_ic)
> m_Y=mean(Y)
> s_Y=sd(Y)
> cat("a =", a, "b =", b, "m_Y =", m_Y, "s_Y =", s_Y, "\n")
a = -0.6015635 b = 2.945029 m_Y = 0.01817798 s_Y = 0.9221586
> Zmax<-double(B)
> for (i in 1:B){
+ Cp<-simulaCp(n,mu1,sigma1,N=100)19
+ {if (ad==0){
+ Y=a+b*log(Cp)
+ mY=m_Y
+ sY=s_Y}
+ else{
+ Y<-qnorm(pCp(Cp))
+ mY=0
+ sY=1}}
+ Z<-aewma(Y,lambda,gamma,mY)
+ Zmax[i]<-max(abs(Z-mY))}
> zmax<-Zmax/(sY*(lambda/(2-lambda))^0.5)
> k<-quantile(zmax,1-alpha)
> cat("mY =",mY,"sY =",sY,"lambda=",lambda,"k=",k,"\n")

```

Sono i comandi utilizzati per poter ripetere B volte la simulazione di un processo fuori controllo, il calcolo degli indici di capacità scelti, la loro

¹⁹ N=100, fissato per ricavare la statistica ARL con probabilità 1-p₀.

trasformazione, la statistica EWMA o AEWMA, la sua standardizzazione e il calcolo del quantile.

```
> mu1=mu1_2
> sigma1=sigma1_2
```

Vengono così aggiornati i valori della media e della deviazione standard del processo, dalla situazione in controllo a quella fuori controllo.

```
> limite<- k * sqrt(lambda/(2 - lambda)) * sY
> limite2<- 2 * limite
```

Viene calcolato uno dei due limiti simmetrici della carta.

```
> rl=double(B)
> for(o in 1:B){
+ Cp<-simulaCp(n,mu1,sigma1)
+ {if (ad==0){
+ Y=a+b*log(Cp[Cp>0])20
+ mY=m_Y
+ sY=s_Y}
+ else{
+ Y<-qnorm(pCp(Cp))
+ mY=0
+ sY=1}}
+ Z<-aewma(Y,lambda,gamma,mY)
+ rl[o]<-min(which(c(abs(Z-mY),limite221)>limite))}
> ARL<-mean(rl)
> cat("mY =",mY,"sY =",sY,"lambda =",lambda,"\n",
+     "gamma      =",gamma,"limite      =",limite,"ARL      =",ARL,"\n")
```

²⁰ Simulando N=1000 processi fuori controllo alcuni di essi sono negativi e comportano un errore nel calcolo del logaritmo, quindi elimino questi casi anomali.

²¹ Limite2 serve per impedire che rl non sia mai soddisfatta ed avere quindi un valore <1000.

Con questi comandi viene ripetuta B volte la simulazione di N=1000 indici di capacità di un processo fuori controllo, ne viene calcolata la trasformata²² e la statistica EWMA o AEWMA. Successivamente viene fatta la media tra tutte le prime osservazioni della statistica che superano il limite ottenendo così la statistica ARL.

²² Come in precedenza con gli indici del processo in controllo.

Bibliografia

Douglas C. Montgomery (2005): "Introduction to Statistical Quality Control", Fifth Edition. Edizione Italiana "Controllo Statistico della Qualità" (2006) McGraw-Hill.

Castagliola, P.; Vannman, K. (2007): "*The efficiency of the EWMA capability chart*" IEEE International Conference on Industrial Engineering and Engineering Management, Singapore 2-4 Dec. 2007 Page(s): 1389 – 1393.

Capizzi, G. e Masarotto, G. (2003): "*An Adaptive Exponentially Weighted Moving Average Control Chart*", Technometrics, 45, 199 – 207.

Ringraziamenti

Raggiunto questo traguardo devo innanzitutto ringraziare Dio per tutte quelle volte che ero in difficoltà e mi ha aiutato (e sono state tante!). Se penso che dopo il diploma lavavo piatti nove ore al giorno e mi ero rassegnato all'idea di poter proseguire gli studi...

Un grandissimo grazie va alla mia famiglia che mi ha sopportato e supportato in questi anni di "*studio*".

A mio padre Roberto, mia madre Ornella e mia sorella Alice un enorme grazie per quello che hanno fatto per me in questi anni. Lo so, sono un rompiscatole e ho un caratteraccio però vi voglio bene!

Grazie allo zio Davide, la mia spalla per quel che riguarda la carriera universitaria. Lo devo soprattutto a lui se mi sono iscritto... due volte, per giunta! Grazie alla nonna Cecilia, che anche se ormai mi vede solo in fotografia, come dice lei, mi vuole sempre bene, e allo zio Roberto che per me è sempre disponibile. A i miei zii Gianni e Ornella e alle cugine Marilisa e Claudia. Alla fine sono il primo della famiglia a laurearsi, spero ne siano orgogliosi Tiziano, Ester, Bruno, Luca, Maddalena, Antonio, Rosina e tutti gli altri da lassù. Un grazie a Tiziano, spero di riuscire a ripagarti per quello che hai fatto per me e la mia famiglia.

Un ringraziamento particolare va al mio relatore, professor Masarotto. È stato davvero disponibile, comprensivo e d'aiuto.

Grazie alla mia cuginetta Eliana, per le ore passate insieme a parlare di ciò che più ci piace ("Muoviti a laurearti!").

Un grazie anche a Davide e Debora, due persone che fanno parte della mia vita e che per me ci sono sempre.

Ringrazio Valeria che da sei anni mi è sempre vicino, malgrado la distanza...

Un'amica più unica che rara, come dice lei.

Un ringraziamento anche a Don Antonio, suor "Flò", gli animatori e tutti i bambini che mi fanno correre come un matto il sabato pomeriggio.

Passo ora a ringraziare i colleghi statistici, iniziando dagli amici che davvero mi sopportano e in particolare da quelli che mi hanno dato una grossa mano in questi tre anni, ovvero Enrico e Pietro.

Il *signor Pierre*, con cui ho *condiviso* praticamente tutto di questi tre anni, è un ottimo collega e amico nonché impareggiabile comico. Posso dire di non aver mai riso tanto come in questi anni. Di sicuro ho recuperato il tempo perso come giocatore di carte... prima di entrare nell'*aula studio*²³ sapevo appena cosa fossero, le carte.

Un grazie particolare a *Tonio*. Mi ricordo ancora la prima volta che ci incontrammo in treno, il primo giorno di precorso... Non ho più smesso di ridere di e con lui. Sono contento di averlo conosciuto, oltre che un personaggio è un amico generoso e disponibile. Parte di questa mia laurea è merito suo, anche se negherà. Abbiamo talmente tante cose da ricordare che ci siamo riproposti di scrivere una sorta di "diario", quindi non inizio nemmeno...

Ringrazio poi Laura, a cui ho chiesto, e ottenuto, mille favori, e che spero tornerà da noi, Marta B., Valentina (le primissime amiche!), Andrea (Schez), Andrea (Piga), Andrea (Kappa85), Andrea (Vacca... e dopo qualcuno si lamenta dell'esistenza dei soprannomi... per forza!), Alessio, Emanuele (Léle), Matteo(Stocco), Tommy, Aliona, Riccardo (Banotta), Silvia, Enrico Moro (per cui, come dice il buon *Pierre*, vale il prezzo delle tasse universitarie), Alberto (Grosso), Michele, Daria, Chiara, Elisa, Marta V., Elena, Federica, tutti i sosia (tantissimi) e quanti ho dimenticato.

²³ È il nome dell'osteria del complesso di Santa Caterina. Ricorda un po' il paese dei Balocchi.

Ringrazio Trenitalia per avermene fatte passare di tutti i colori (anche oggi! Il treno è arrivato, pienissimo, ed è partito subito, prima che riuscissi a farmi largo tra la gente!).

Ringrazio infine la mia fedele compagna R. Quando non sapevo *che fare* mi ha sempre aiutato a decidere!

Mi dispiace di aver dimenticato qualcuno nella fretta... non me ne vogliano.

Un ringraziamento al contrario alla mia mente contorta... maledetta quella volta che mi venne in mente SKYLVR!

Spero di non dimenticare mai i bei momenti vissuti a Santa Caterina, di viverne altrettanti, e magari... ritrovarmi di nuovo fra due anni a scrivere i ringraziamenti della tesi!