



Università degli Studi di Padova
Facoltà di Scienze Statistiche
Corso di Laurea in Statistica e Tecnologie Informatiche

METODI SELF-STARTING

Relatore: Prof. GIOVANNA CAPIZZI

Laureando: CICHI IULIA
Matricola: 569265

Anno Accademico 2008/2009

Indice

Introduzione	3
Capitolo 1	
Procedure di controllo Self-Starting univariate	7
1.1 Le statistiche Q univariate	7
1.2 Carte Shewart sulle statistiche Q.....	8
1.3 Carte con memoria sulle statistiche Q	14
1.4 Confronti tra Q-Shewart, Q-CUSUM e Q-EWMA.....	18
Capitolo 2	
Procedure di controllo Self-Starting multivariate	21
2.1 Le statistiche Q multivariate.....	21
2.2 Performance delle carte Self-Starting multivariate	28
2.2.1 Misure di performance delle carte Self-Starting multivariate	28
2.2.2 Disegno delle carte Self-Starting multivariate.....	29
2.3 Prestazioni delle carte basate su “accumulo di vettori” e “accumulo di scalari”	34
2.4 Confronto della carta SSMEWMA e carta “accumulo di vettori” ...	36
2.5 Discussione	39
Capitolo 3	
Confronto tra due carte Self.-Starting multivariate.....	41
3.1 Le carte di controllo per “accumulo di vettori”	41
3.1.1 Esempio di carte U e T.....	41
3.1.2 Confronto delle prestazioni al variare del periodo IC	43
3.1.3 Simulazioni	43
3.2 Carte di controllo SSMEWMA	44
3.2.1 Simulazioni per carta SSMEWMA.....	45
3.2.2 Confronto dei metodi Sullivan e Hawkins.....	48
Conclusioni	51
Appendice	53
Bibliografia	69

Introduzione

Le tecnologie di produzione avanzata, spinte dalla necessità di soddisfare le mutevole esigenze dei clienti, con una maggiore efficienza e dei costi di produzione relativamente bassi, si basano oggi su nuove strategie di produzione.

In molti settori, si assiste ad un drastico cambiamento nella filosofia aziendale caratterizzata sempre più spesso di volumi bassi di produzione eseguiti in tempi sempre più rapidi, quali ad esempio gli ambienti job-shop. Come risultato si ha la necessità di sviluppare tecniche di *Statistical Process Control* (SPC) adeguate in quei contesti dove il volume ridotto di produzione, le produzioni temporanee o le rapide innovazioni, qui generalmente denominati *short-runs*, producono un numero di osservazioni insufficiente per una stima accurata dei parametri del processo e per il disegno dei limiti di controllo con cui sorvegliare la stabilità del processo.

Dalla prima carta di controllo statistico sviluppata da Shewhart (1931), le carte di controllo sono state uno strumento statistico fondamentale nel settore industriale. Un problema comune con quale si confrontano gli operatori è che la progettazione delle carte di controllo prevede l'uso dei parametri del processo in controllo che si presume siano conosciuti esattamente.

In pratica tuttavia parametri del processo sono raramente conosciuti e sono spesso sostituiti con delle stime calcolate su un campione in controllo, detto di Fase I. L'idea principale è che con grandi campioni le stime sono vicine ai veri parametri; ma, anche se quest'approccio sembra ragionevole, è stato dimostrato che i parametri stimati diminuiscono sostanzialmente le prestazioni previste per le carte di controllo (Jones et. al., 2001, Woodall e Montgomery, 1999). In letteratura si trova che, per stabilire in modo

accurato i limiti di controllo per una carta Shewhart, per esempio, è necessario raccogliere circa 25 campioni, da 4 a 6 osservazioni (Montgomery 1996). Quesenberry (1993) raccomanda un campione di dimensione 300 per stabilire i limiti per carte di controllo per osservazioni singole. Tuttavia, campioni di questa grandezza possono non essere disponibili nelle produzioni di basso volume, e il numero dati necessari per soddisfare tali raccomandazioni potrebbe non essere disponibile prima dell'inizio della produzione o addirittura superare il numero totale delle parti prodotte in un ciclo di produzione breve.

Quando poi più caratteristiche non indipendenti sono monitorate simultaneamente, dovrebbero essere utilizzate delle carte di controllo multivariate in grado di tener conto della correlazione delle variabili. Purtroppo, le prestazioni delle carte multivariate diminuiscono quando i parametri del processo sono sostituiti con delle stime. Nel caso multivariato sono dunque necessari campioni di Fase I di dimensioni molto grandi per ottenere piccole variabilità dei parametri stimati, specialmente per la matrice di covarianze. Costoso, e quindi indesiderabile, questo rimedio può essere perfino impossibile in alcune situazioni, come ad esempio nei problemi *short-runs*.

Una possibilità è di usare alcuni metodi *self-starting* secondo i quali letture successive del processo sono utilizzate per aggiornare parametri stimati e simultaneamente verificare le condizioni di fuori controllo del processo. L'idea principale è quella di eliminare il problema dei parametri sconosciuti, producendo dai dati con parametri ignoti, un corrispondente flusso di osservazioni indipendenti ed identicamente distribuite come una normale standard. Tre situazioni favoriscono l'utilizzo delle procedure *self starting*:

- quando i primi fuori controllo prodotti sono costosi;
- quando vi è un notevole ritardo tra le unità prodotte;

- quando non sono disponibili campioni sufficientemente grandi per la stima accurata dei parametri.

Per monitorare un processo quando la media e la varianza sono ignote e per far fronte al problema della stima della media e varianza del processo nel caso di produzioni di breve tempo, Hawkins (1987), Quesenberry (1991,1995,1997) Hawkins e Olwell(1997) hanno sviluppato carte delle carte self-starting univariate, chiamate Carte Q.

In ambito multivariato, le carte *self-starting* sono state proposte da Schaffer (1998), Quesenberry (1997), Sullivan e Jones (2002) e Hawkins (2007). In particolare, il lavoro di Hawkins (2007) discute l'utilizzo di una carta EWMA multivariata (MEWMA), introdotta da Lowry et. al. (1992) come estensione multivariata della carta EWMA (Roberts, 1959) nel caso in cui il vettore delle medie e la matrice di covarianze in controllo non siano note e non sia possibile disporre di campioni di Fase I sufficientemente grandi.

In questa tesi ci si propone in particolare di presentare alcune delle carte *self-starting* enumerate sopra. Il capitolo 1 parla delle carte univariate di Hawkins (1987) e di Quesenberry(1991,1995). Il capitolo 2 presenta le carte multivariate di Sullivan e Jones (2002) e di Hawkins(2007). Queste ultime due si analizzano e si tenta un confronto delle prestazioni nel capitolo 3, utilizzando i risultati ottenuti con i programmi in R disponibili nel Appendice.

Capitolo 1

Procedure di controllo self-starting univariate

1.1 Le statistiche Q univariate

Si assuma che le misure X del processo in controllo siano variabili casuali indipendenti e identicamente distribuite di una normale $N(\mu, \sigma^2)$, dove μ e σ^2 non sono note. L'allontanamento dalla situazione di controllo può assumere una forma di shift in μ o in σ , a partire da qualche istante e persistente per qualche tempo. Indicando con:

- \bar{x}_r la media delle prime r osservazioni
- S_r la somma dei quadrati delle deviazioni delle prime r osservazioni

dalla media \bar{x}_r ,

Le statistiche \bar{x}_r e S_r si possono ottenere facilmente utilizzando le seguenti formule ricorsive, per $r \geq 2$,

$$\bar{x}_r = \bar{x}_{r-1} + (x_r - \bar{x}_{r-1}) / r$$

$$S_r = S_{r-1} + (r-1)(x_r - \bar{x}_{r-1})^2 / r$$

Si definisce la statistica

$$T_r = \frac{a_r (x_r - \bar{x}_{r-1})}{\sqrt{S_{r-1} / (r-2)}} \quad (1)$$

dove $a_r = \sqrt{(r-1)/r}$.

La statistica (1) è il test che compara la compatibilità dell'osservazione X_r con i suoi predecessori ed è il test t di Student che verifica la compatibilità dei due campioni X_r e $\{X_1, X_2, \dots, X_{r-1}\}$. I residui standardizzati T_r , hanno anche importanti proprietà teoriche, poiché i

successivi T_r sono statisticamente indipendenti e seguono la distribuzione t con $r-2$ gradi di libertà.

Mentre l'indipendenza è molto preziosa per rendere possibile il controllo della run lengths (RL) delle somme cumulate, la distribuzione t_{r-2} (in particolare perchè il numero di gradi di libertà varia con il numero di osservazioni) non è così conveniente come sarebbe una distribuzione normale standard.

Per questa ragione si introducono le statistiche Q , trasformazioni della sequenza originale di osservazioni in una nuova sequenza di osservazioni indipendenti ed identicamente distribuite (i.i.d.) come una normale standard $N(0,1)$.

Sotto le assunzioni di normalità e stabilità (media e la deviazione standard sono costante) del processo tutte queste statistiche hanno dunque distribuzione normale standard e sono indipendenti e possono dunque essere sorvegliate per mezzo delle carte tradizionali.

Assumendo che i dati seguano una distribuzione normale con uno o entrambi parametri ignoti, e che si abbia un cambiamento in μ o in σ , si presenta nel seguito le varie statistiche Q e si discutono le proprietà di alcune carte tradizionali, come la Shewhart e il CUSUM, basate su Q .

1.2 Carte Shewhart sulle statistiche Q

Quesenberry (1991, 1995) suggerisce di rappresentare le statistiche Q su una carta Shewhart con limiti ± 3 . L'abilità di riconoscere molti 'patterns' anomali dei valori rappresentati, segnalare una causa assegnabile di cambiamento nel processo, è un'estremamente utile proprietà delle carte Shewhart in generale, e delle carte Q-Shewhart in particolare. Si nota infatti che la sensibilità delle carte Q-Shewhart nel determinare cambiamenti

temporali dei parametri ad un passo, quando questo è successo, è massima possibile.

Supponiamo che, in particolare, avvenga uno shift in μ dal valore nominale

$\mu = \mu_0$ al valore $\mu = \mu_0 + \delta\sigma$. Si indichi con Φ , Φ^{-1} e G_v , la funzione di distribuzione normale standard, la sua inversa e la funzione di distribuzione t con v gradi di libertà

Quesenberry (1991) suggerisce le seguenti statistiche Q per individuare un cambiamento nella media del processo.

- a. media μ_0 e deviazione standard σ_0 del processo note (caso KK)

$$Q_r(X_r) = \frac{X_r - \mu_0}{\sigma_0}, r = 1, 2, \dots \quad (2)$$

- b. media μ incognita e deviazione standard σ_0 del processo nota (caso UK)

$$Q_r(X_r) = \left(\frac{r-1}{r}\right)^{\frac{1}{2}} \frac{(X_r - \bar{X}_{r-1})}{\sigma_0}, r = 2, 3, \dots \quad (3)$$

- c. media μ_0 nota e deviazione standard σ del processo incognita (caso KU)

$$Q_r(X_r) = \Phi^{-1} \left\{ G_{r-1} \left(\frac{X_r - \mu_0}{S_{0,r-1}} \right) \right\}, r = 2, 3, \dots \quad (4)$$

dove $S_{0,r}^2 = \frac{1}{r} \sum_{j=1}^r (X_j - \mu_0)^2$.

- d. media μ e deviazione standard σ del processo incognite (caso UU)

$$Q_r(X_r) = \Phi^{-1} \left\{ G_{r-2} \left(\frac{r-1}{r} \right)^{\frac{1}{2}} \left(\frac{X_r - \bar{X}_{r-1}}{S_{r-1}} \right) \right\}, r = 3, 4, \dots \quad (5)$$

H_v – funzione di distribuzione χ^2 con v gradi di libertà

F_{v_1, v_2} - funzione di distribuzione F di Snedocor con (v_1, v_2) gradi di libertà

Sia $R_r = X_r - X_{r-1}$, e rispettivamente H_v e F_{v_1, v_2} la funzione di distribuzione χ^2 con v gradi di libertà e quella di una F di Snedocor con (v_1, v_2) gradi di libertà. Qualora si abbia uno shift della deviazione standard, Quesenberry (1991) suggerisce le seguenti statistiche Q per la sorveglianza della varianza del processo

e media μ incognita e deviazione standard σ_0 del processo nota (Caso K)

$$Q(R_r) = \Phi^{-1} \left\{ H_1 \left(\frac{R_r^2}{2\sigma_0^2} \right) \right\}, r = 2, 4, 6, \dots \quad (6)$$

f. media μ e deviazione standard σ del processo incognite (caso U)

$$Q(R_r) = \Phi^{-1} \left\{ F_{1, v} \left(\frac{vR_r^2}{R_2^2 + R_4^2 + \dots + R_{r-2}^2} \right) \right\}, r = 4, 6, \dots \quad (7)$$

$$v = \frac{r}{2} - 1$$

Si osservi in particolare che

- caso KK: è il test per verificare che un campione di dimensione 1 da una distribuzione normale di varianza σ_0^2 abbia media pari a μ_0
- caso UK: è il test per verificare che un campione di dimensione 1 e un campione di dimensione $(r-1)$ abbiano la stessa media (assumendo che abbiano la stessa varianza nota σ_0^2)
- caso KU: è il test per verificare che un campione di dimensione 1 e un campione di dimensione $(r-1)$ abbiano la stessa media μ_0 , assumendo che hanno entrambi distribuzioni normali con stessa varianza sconosciuta.
- caso UU: è il test per verificare che un campione di dimensione 1 e un campione di dimensione $(r-1)$ hanno la stessa media (ignota), assumendo che provengono da distribuzioni normali con la stessa varianza ignota

Quesenberry (1991) dimostra che le statistiche Q delle equazioni 2–7 sono sequenze di variabili casuali i.i.d. $\sim N(0,1)$.

Così, se queste statistiche sono tracciate su una carta di controllo Shewhart con limiti di controllo ± 3 sappiamo quale tipo di pattern aspettarci quando il modello è stabile e i parametri μ e σ sono costanti.

In particolare, data una sequenza di statistiche $Q : Q_1, Q_2, Q_3, \dots, Q_r, \dots$, Quesenberry (1995) analizza in particolare quattro carte tipo Shewhart per individuare un cambiamento in μ o σ .

1. *The 1 of 1 Test*: data Q_r , il test segnala un incremento in μ se $Q_r > 3$ e il test segnala una diminuzione in μ se $Q_r < -3$;
2. *The 9 of 9 Test*: data Q_r , il test segnala un incremento in μ se $Q_r, Q_{r-1}, Q_{r-2}, \dots, Q_{r-8}$ sono tutte > 0 , e il test segnala una diminuzione in μ se $Q_r, Q_{r-1}, Q_{r-2}, \dots, Q_{r-8}$ sono tutte < 0 . Per effettuare questo test devono dunque essere disponibili 9 statistiche Q consecutive.
3. *The 3 of 3 Test*: data Q_r , il test segnala un incremento in μ se Q_r, Q_{r-1}, Q_{r-2} sono tutte > 1 e il test segnala una diminuzione in μ se Q_r, Q_{r-1}, Q_{r-2} sono tutte < -1 . Per effettuare questo test devono essere disponibili 3 statistiche Q consecutive.
4. *The 4 of 5 Test*: dati Q_r , il test segnala un incremento in μ se 4 delle 5 valori $Q_r, Q_{r-1}, \dots, Q_{r-4}$ sono > 1 , e una diminuzione in μ se 4 delle 5 sono < -1 . Devono essere disponibili 5 statistiche consecutive.

Si è dimostrato che questi sono i più potenti e imparziali test per confronto di

due campioni (Lehmann-1959). E da qui segue che i *test 1 of 1* sulle relative carte Q sono le più sensibili possibili per individuare un cambiamento della media subito dopo lo shift.

Esaminando le formule, è facile vedere che l'effetto del cambiamento della media sulle statistiche Q varia al variare del tipo di conoscenza sui parametri del processo.

Ad esempio, per i casi KK, KU, UK l'effetto di un cambiamento nella media sarà di incrementare la media delle statistiche Q quando $\delta > 0$, e di

diminuire la loro media quando $\delta < 0$. Per il caso UU il cambiamento della media non colpisce solo il numeratore, ma anche il denominatore S_{r-1} . Se succede un cambiamento all'osservazione $r=c$ i valori S_r , per $r=c+2, r=c+3, \dots$, tendono ad essere elevati e aumentano con r . Questa tendenza ad aumentare del denominatore produce una diminuzione delle statistiche Q e di conseguenza una riduzione della probabilità di individuare un cambiamento subito dopo che si è verificato con il *test 1 of 1*.

Nel caso in cui ci si disponga in ogni istante temporale di un campione di osservazioni è possibile calcolare le carte Q per le medie dei campioni \bar{X}_r e per le varianze dei campioni S_r^2 , per dati raggruppati in campioni di dimensioni n_r .

$$\bar{\bar{X}}_r = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2 + \dots + n_r \bar{X}_r}{n_1 + n_2 + \dots + n_r}$$

$$S_{p,r}^2 = \frac{(n_1 - 1)S_1^2 + \dots + (n_r - 1)S_r^2}{n_1 + n_2 + \dots + n_r - r}$$

Le statistiche Q per sorvegliare la media del processo per dati raggruppati in campioni sono:

- a. media μ_0 e deviazione standard σ_0 del processo note (caso KK)

$$Q_r(\bar{X}_r) = \sqrt{n_r} \frac{(\bar{X}_r - \mu_0)}{\sigma_0}, r = 1, 2, \dots \quad (8)$$

- b. media μ incognita e deviazione standard σ_0 del processo nota (caso UK)

$$Q_r(\bar{X}_r) = \sqrt{\frac{n_r(n_1 + \dots + n_{r-1})}{n_1 + \dots + n_r}} \frac{(\bar{X}_r - \bar{\bar{X}}_{r-1})}{\sigma_0}, r = 2, 3, \dots \quad (9)$$

- c. media μ_0 nota e deviazione standard σ del processo incognita (caso KU)

$$S_{0,r}^2 = \frac{\sum_{\alpha=1}^r \sum_{j=1}^{n_{\alpha}} (X_{\alpha j} - \mu_0)^2}{n_1 + \dots + n_r} \quad (10)$$

$$Q_r(\overline{X}_r) = \Phi^{-1} \left\{ G_{n_1+\dots+n_r} \left(\frac{\sqrt{n_r} (\overline{X}_r - \mu_0)}{S_{0,r}} \right) \right\}, r = 2,3,\dots$$

d. media μ e deviazione standard σ del processo incognite (caso UU)

$$Q_r(\overline{X}_r) = \Phi^{-1} \left\{ G_{n_1+\dots+n_r-r} \left[\sqrt{\frac{n_r(n_1+\dots+n_{r-1})}{n_1+\dots+n_r}} \left(\frac{\overline{X}_r - \overline{X}_{r-1}}{S_{p,r}} \right) \right] \right\}, r = 2,3,4,\dots$$

(11)

Le statistiche (8) - (11) per le \overline{X}_r sono variabili casuali i.i.d. $N(0,1)$, sotto condizioni di stabilità del processo, e come detto in precedenza per misurazioni singole sono i migliori test per la media.

Statistiche Q per la varianza del processo per dati raggruppati in campioni:

e. media μ incognita e deviazione standard σ_0 del processo nota (Caso K)

$$Q(S_r^2) = \Phi^{-1} \left\{ H_{n_r-1} \left(\frac{(n_r-1)S_r^2}{\sigma_0^2} \right) \right\}, r = 1,2,\dots \quad (12)$$

f. media μ e deviazione standard σ del processo incognite (caso U)

$$w_r = \frac{(n_1+\dots+n_{r-1}-r+1)S_r^2}{(n_1-1)S_1^2+\dots+(n_{r-1}-1)S_{r-1}^2} \quad (13)$$

$$Q(S_r^2) = \Phi^{-1} \left\{ F_{n_r-1, n_1+\dots+n_{r-1}-r+1}(w_r) \right\}, r = 2,3,\dots$$

Sotto assunzioni che i valori del campione r-esimo sono osservazioni indipendenti da una $N(\mu_r, \sigma^2)$, le statistiche Q delle equazioni (12), (13) sono sequenze di v.c. indipendenti $N(0,1)$

1.3 Carte con memoria sulle statistiche Q

Hawkins (1987) introduce una carta cusum per sorvegliare il livello e la dispersione del processo nel caso in cui non si conoscano i parametri del processo e si vuole sapere se ci sono state derive dalle condizioni iniziali di start-up. Utilizza la sequenza dei residui standardizzati T_r dati nella formula (1) che sono indipendenti ma i gradi di libertà della loro distribuzione t_{r-2} cambia con numero di osservazioni.

Per questo motivo, Hawkins fa uso di un risultato teorico di Wallace(1959) per ottenere le variabili u_r

$$u_r = \left(\frac{8v+1}{8v+3} \right) \left[v \log_e \left(1 + \frac{T_r^2}{v} \right) \right] \text{ dove } v = n-2$$

Le quantità u_r sono funzioni delle deviazioni standardizzate $\frac{x_r - \overline{x_{r-1}}}{S_{r-1}}$, e se $j \rightarrow \infty$ allora $u_r \rightarrow (x - \mu)/\sigma$. Queste sono approssimazioni di variabili casuali indipendenti di una normale standard $N(0,1)$, ciascuna delle quali misura la deviazione del corrispondente x dai suoi predecessori, usando efficientemente la media e deviazione standard di tutte le precedenti sequenze. Hawkins propone, appena le osservazioni x_r sono disponibili, di disegnare un CUSUM delle u_r per controllare gli scostamenti della media di u_r da zero. Ogni spostamento significa implica che la media delle x_r è plausibilmente cambiata dal suo livello atteso.

Allo scopo di controllare la variabilità del processo fa uso della statistica definita in Hawkins(1981):

$$v_r = \sqrt{|u_r| - 0.822} / 0.349$$

Le v_r sono molto vicine a $N(0,1)$ quando processo è in controllo e non sono molto influenzate dai cambiamenti della media delle x_r , invece un cambiamento nella deviazione standard delle x_r porta a un cambiamento

nella loro media. Si nota che le v_r sono funzione delle u_r , e così anche le v_r sono indipendenti.

Per le statistiche u_r e v_r si istituiscono 4 CUSUM per cambiamenti al di sopra e al di sotto della media e della deviazione standard. Dati i 4 seguenti valori iniziali:

$$L_0^+ = L_0^- = S_0^+ = S_0^- = 0$$

e le x_r si aggiornano le seguenti statistiche cusum:

$$L_r^+ = \max(0, L_{r-1}^+ + u_r - k)$$

$$L_r^- = \min(0, L_{r-1}^- + u_r - k)$$

$$S_r^+ = \max(0, S_{r-1}^+ + v_r - k)$$

$$S_r^- = \min(0, S_{r-1}^- + v_r - k)$$

Si ottiene un segnale quando il valore assoluto di una delle quantità sopra eccede h – l'intervallo di decisione. Il valore di riferimento k corrisponde alla semi ampiezza del salto. I parametri h e k sono scelti decidendo prima per quale cambiamento della media si desidera migliorare le performance .

Un segnale dato da:

- L^+ e L^- si interpreta come un aumento e una diminuzione rispettivamente della media delle x_r
- S^+ e S^- si interpreta come un aumento e una diminuzione rispettivamente della deviazione standard delle x_r

Quando il processo è in controllo, ognuna di queste cusum sarà caratterizzata da una certa oscillazione intorno allo zero. Nella diagnosi di un segnale fuori controllo si studiano gli spostamenti oltre lo zero che si concludono con un segnale. Il primo valore diverso da zero di questi allontanamenti (o escursioni) segnalati è l'istante in cui si dichiara che nel processo è avvenuto un cambiamento. Tutte le quattro statistiche cusum si possono rappresentare sulla stessa carta utilizzando diversi simboli grafici.

Mentre le L^+ e S^+ occupano la parte positiva, le L^- e S^- la meta negativa, ciò implica una carta di facile lettura.

Se la media si sposta verso l'alto di un valore δ : $\mu \rightarrow \mu + \delta$, la media del numeratore di T_r : $a_r(x_r - \overline{x_{r-1}})$, si sposta da $0 \rightarrow \delta a_r$ e T_r diventerà una t non centrata con parametro di noncentralità δa_r . Quindi u_r si sposta da una media nulla a una media positiva e il cusum per la media avrà uno spostamento verso l'alto. Se all'inizio la statistica di controllo non è sufficientemente grande per dare un segnale, dopo l'aggiornamento la media $\overline{x_r}$ si sposterà da $\mu \rightarrow \mu + \delta/r$. Quando si calcola la differenza $x_{r+1} - \overline{x_r}$, le successive cusum saranno sostanzialmente positive con un ulteriore aumento della cusum. Allora indifferentemente da quanti termini si aggiungono al cusum, questo avrà una deriva strettamente positiva con un conseguente un segnale.

Se la "storia" è lunga (periodo prima dello shift), le informazioni precedenti saranno predominanti, la media si sposterà molto lentamente verso un nuovo livello con una deriva del cusum più rapida nell'individuare uno shift nella media. Se la storia è breve, la media si sposta più rapidamente verso un nuovo livello, la deriva del cusum è molto meno pronunciata e l'individuazione del cambiamento molto lenta e meno probabile.

Un primo effetto dello shift nella media di x_r è sul cusum L della media, ma oltre al cambiamento del numeratore di T_r anche la varianza di u_r sarà modificata seppure in piccola misura. Così l'effetto principale del cambiamento sulla media è sul cusum L e un più piccolo, generalmente trascurabile, effetto sul cusum S . In contrasto, un cambiamento nella deviazione standard σ implica un cambiamento del livello delle v_r , con una principale deriva del cusum S , ma avrà anche un effetto secondario sul cusum L . Così, se σ aumenta, h e k aumenteranno essendo date in unità di σ con un conseguente un effetto sul cusum S .

In sintesi:

- se il cusum per la media segnala → significa un cambiamento della media

- se il cusum per la deviazione standard segnala → significa un cambiamento su σ

- se entrambe segnalano può significare uno shift su entrambe media e varianza oppure uno shift solo nella varianza .

Se un cusum inizia subito la deriva e finisce con un segnale è opportuno far ripartire il cusum dopo aver omesso le prime osservazioni.

Studiando le prestazioni del cusum in termini di ARL si osserva che, l'effetto della lunghezza della storia del processo in-controllo (IC) è più marcato sostanzialmente quando si passa da 25 a 50 osservazioni IC e non molto marcato per le lunghezze più grandi. Una serie di 25 osservazioni iniziali IC sembra essere troppo breve, specialmente se guardiamo la σ . Se si utilizzano 50 osservazioni invece la risposta è due volte più veloce per la deviazione standard. Così in un approccio di disegno del tipo Fase I/FaseII sarebbe difficile da giustificare il raddoppiamento di uno studio di start-up solo per verificare il miglioramento delle prestazioni, mentre utilizzando il metodo Self –Starting si utilizzano le osservazione correnti del processo senza spese supplementari o ritardi. Il vantaggio delle Self Starting è che la stima della media e della deviazione standard può essere migliorata via che le osservazioni diventano disponibili, ciò che implica che le prestazioni in termini di ARL migliorano senza costi ulteriori. Le situazioni in cui le prestazioni sono inferiori sono quando le osservazioni ottenute subito dopo lo start-up sono già fuori controllo ma di un'entità troppo piccola per essere individuate da un cusum.

1.4 Confronti tra Q-Shewhart e Q-CUSUM e Q-EWMA

Nel caso si vogliano individuare efficientemente sia piccoli sia grandi cambiamenti nei parametri, si consiglia l'utilizzo di una combinazione di carte cusum/shewhart. La scala delle Shewhart e Cusum sono scelte in modo che le stesse linee orizzontali servono come limite di decisione per ognuna. Un segnale Shewhart indica un probabile errore di un solo campione (un outlier), e questo campione è quindi analizzato, mentre un segnale cusum richiede una revisione generale della metodologia.

Anche Quesenberry (1995) aveva proposto di monitorare le statistiche Q, utilizzando le carte CUSUM o EWMA, osservando che le loro prestazioni sono comparabili e sono in complesso migliori delle carte Shewhart. Così, in aggiunta ai 4 test sopra applicati alle carte QShewhart, si considerano anche le carte EWMA e CUSUM calcolate per le statistiche Q (Quesenberry, 1995).

5) La statistica EWMA, Z_t è data da :

$$Z_r = \lambda Q_r + (1-\lambda) Z_{r-1}, r=1, 2, \dots;$$

$$Z_0 = 0$$

dove λ è il peso della osservazione corrente

Questi valori sono rappresentati sulla carta con limiti di controllo $\pm K\sqrt{\lambda/(2-\lambda)}$,

Se $Z_r > UCL(Z_r)$ significa un incremento in μ

Se $Z_r < LCL(Z_r)$ significa una diminuzione in μ

6) Le statistiche CUSUM, S_r^+ , S_r^-

$$S_r^+ = \max(0, S_{r-1}^+ + Q_r - k_s)$$

$$S_r^- = \min(0, S_{r-1}^- + Q_r + k_s)$$

$$S_0^+ = S_0^- = 0$$

Se $S_r^+ > h_s$ significa un incremento in μ

Se $S_r^- < -h_s$ significa una diminuzione in μ

Per $\delta=0$ (caso stabile) tutte le statistiche sono i.i.d.. Dalla stima della probabilità di falsi allarmi si ottiene che:

- l'EWMA, CUSUM, 1 of 1 Test danno risultati simili
- 9 of 9 Test $\approx 0.73 \cdot 1$ of 1 Test
- 4 of 5 Test $\approx 1.2 \cdot 1$ of 1 Test
- 3 of 3 Test $\approx 2.3 \cdot 1$ of 1 Test

Per $\delta > 0$ si osserva che:

-la rapidità con la quale segnalano i test Shewhart dipende dal numero dei dati necessari per effettuare il test. Il test 9 of 9 presenta una performance insoddisfacente.

-il test 1 of 1 non è competitivo come gli altri test per individuare un cambiamento, ma è più sensibile nell'individuare un solo outlier (valori anomali). Gli outliers sono rimossi e non sono utilizzati nei successivi calcoli per le statistiche Q.

- non c'è un test migliore per tutti i casi.

- le performance delle carte Shewhart 3 of 3, 4 of 5, 9 of 9 sono difficili da comparare perchè non hanno lo stesso tasso di falsi allarmi. Il rapporto tra i loro tassi di falsi allarmi è

$$9 \text{ of } 9 : 4 \text{ of } 5 : 3 \text{ of } 3 \approx 3 : 5 : 10$$

-il test 4 of 5 è considerato migliore per le carte Shewhart, ed è raccomandato per entrambe μ e σ .

-il test 3 of 3 ha un considerevole vantaggio rispetto 4 of 5, ma ha un numero elevato di falsi allarmi.

-le carte EWMA e CUSUM segnalano subito, sono migliori delle Shewhart e sono comparabile tra loro.

Se l'aumento di falsi allarmi è considerato tollerabile al fine di migliorare la sensibilità, il test 3 of 3 è un compromesso accettabile per i processi start-up, prima che siano sufficienti dati disponibili per migliorarne la sensibilità.

L'effetto del ridotto numero di osservazioni nel periodo in controllo e prima del cambiamento è, come per le singole osservazioni, quello di abbassare le prestazioni dei test.

Capitolo 2

Procedure di controllo self-starting multivariate

2.1 Le statistiche Q multivariate

Supponiamo di avere una sequenza di osservazioni indipendenti normali p-variate per quale il vettore delle medie cambia da μ_a a μ_b all'osservazione T+1:

$$X_r \sim N_p(\mu_a, \Sigma) \text{ per } r=1,2,\dots,T$$

$$X_r \sim N_p(\mu_b, \Sigma) \text{ per } r=T+1,T+2,\dots,$$

con vettore delle medie, matrice di covarianze e istante del cambiamento ignoti. Se si verifica un cambiamento è importante che sia segnalato il più presto possibile senza eccessivi falsi allarmi. La deviazione dell'osservazione corrente dalla stima del vettore delle medie rappresenta l'innovazione :

$$\bar{x}_{r-1} = \frac{1}{r-1} \sum_{i=1}^{r-1} x_i \quad (1)$$

$$d_r = (x_r - \bar{x}_{r-1})$$

$$u_r = d_r \cdot \sqrt{\frac{r-1}{r}} \approx N(0, \Sigma)$$

Il vettore delle medie campionarie e la matrice di covarianze possono essere calcolate usando le seguenti formule di aggiornamento proposte da Quesenberry (1997)

$$\bar{x}_r = \frac{1}{r} ((r-1)\bar{x}_{r-1} + x_r)$$

$$Q_r = Q_{r-1} + \frac{r-1}{r} d_r d_r^T$$

$$S_r = \frac{1}{r-1} Q_r \quad (2)$$

$$\bar{x}_0 = \mathbf{0}_p$$

$$Q_0 = \mathbf{0}_{p \times p}$$

Quesenberry (1997) dimostra anche che le statistiche u_r sono indipendenti e identicamente distribuite (i.i.d.) quando il processo è in controllo (IC) e propone le seguenti trasformazioni per ottenere le statistiche da rappresentare sulla carta di controllo:

$$e(x_r) = \Phi^{-1}[F_{p,r-p-1}[A_r]] \quad (3)$$

$$A_r = \left(\frac{(r-1)(r-p-1)}{r * p * (r-2)} \right) (x_r - \overline{x_{r-1}})^T S_{r-1}^{-1} (x_r - \overline{x_{r-1}})$$

per $r=p+2, \dots$

Dove Φ^{-1} è l'inversa della normale standard, e $F_{p,r-p-1}$ è la funzione di distribuzione cumulativa (cdf) della distribuzione F con p e $r-p-1$ gradi di libertà. In controllo le statistiche seguono una distribuzione normale.

Schaffer (1998) introduce le carte self-starting multivariate basate su una forma quadratica del vettore di osservazioni e suggerisce l'accumulo di statistiche scalari usando una carta EWMA univariata (carta S):

$$e_{S,r} = \lambda e(x_r) + (1 - \lambda) e_{S,r} \quad (4)$$

dove λ è il peso assegnato all'osservazione corrente e S indica l'approccio di accumulo degli scalari.

Perché le innovazioni sono vettori, l'approccio alternativo è l'"accumulo di vettori" (carta U) usando un modello EWMA o CUSUM multivariato prima della conversione in una statistica scalare (Sullivan e Jones, 2002):

$$\begin{aligned} z_r &= \lambda u_r + (1 - \lambda) z_{r-1} \\ \Sigma_{z_r} &= \frac{\lambda [1 - (1 - \lambda)^{2r}]}{2 - \lambda} S_{r-1} \\ e_{V,r} &= z_r^T \Sigma_{z_{ij}}^{-1} z_r \end{aligned} \quad (5)$$

Se $\lambda=1$ corrisponde al caso in cui non quando succede accumuli e quindi a una convenzionale carta Hotelling T^2 per le innovazioni.

Sebbene le innovazioni sono identicamente distribuite quando il processo è IC, le statistiche non lo sono. Una conseguenza importante è che la loro varianza diminuisce rapidamente ad aumentare del numero delle osservazioni incluse nella stima dei parametri; come risultato questo approccio presenta lo svantaggio di segnale presto dei falsi allarmi (*early false alarms*), se il limite UCL è costante. Se il limite di controllo è sufficientemente grande per controllare i falsi allarmi si ottiene che la carta non è più tanto sensibile nell'individuare ragionevolmente piccoli e medi cambiamenti.

Questa carta di “accumulazioni di vettori” può essere migliorata con una trasformazione in modo che le statistiche abbiano la stessa distribuzione. Se negli approcci self-starting univariati le statistiche sono frequentemente trasformate in distribuzioni normali standard, nel multivariato le statistiche hanno forme quadratiche che potrebbero essere trasformate in valori negativi difficili da interpretare. Un alternativa è di trasformare le statistiche in distribuzioni χ^2_p perchè le statistiche $e_{v,r}$ delle (5) seguono questa distribuzione asintoticamente. Questa scelta accentua le variazioni dei limiti di controllo rispetto a p . Per ridurre la variabilità dei limiti di controllo si possono trasformare le statistiche in variabile casuali (v.c.) χ^2_1 . Questa trasformazione permette di interpretare le statistiche rappresentate sulla carta come distanza quadratica dal vettore delle medie stimato in controllo. Perchè il grafico delle distanze possa essere più utile si trasformano le statistiche in radice quadrata delle v.c. χ^2_1 .

Le equazioni delle statistiche trasformate per le carte delle “accumulazioni di vettori” sono date da:

$$\begin{aligned}
e_{V,r} &= Z_r^T \sum_{Z_r}^{-1} Z_r \\
\frac{r-p-1}{p(r-2)} e_{V,r} &\approx F_{p,r-p-1} \\
e_{T,r} &= \sqrt{(\chi_1^2)^{-1} \left[F_{p,r-p-1} \left[\frac{r-p-1}{p(r-2)} e_{V,r} \right] \right]}
\end{aligned} \tag{6}$$

Il metodo self starting multivariato di Hawkins (2007) è il primo che trasforma il flusso di vettori originali di parametri ignoti in un flusso di vettori di parametri noti con stessa dimensionalità. Questo flusso di vettori ha molte possibilità di utilizzo, può essere utilizzato per costruire qualsiasi carta di controllo, come la T^2 di Hotelling o qualsiasi metodo multivariato basato sulla statistica CUSUM. Hawkins mostra l'utilizzo dei flussi trasformati per il disegno di una carta MEWMA. Nel caso self starting "font-end" questa carta avrà le stesse proprietà IC come se la vera media e la matrice di covarianze del processo fossero esattamente note. La metodologia di Hawkins prevede due fasi concettuali:

- Prima fase: si trasforma il flusso di vettori $\mathbf{X} \sim \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ in una sequenza di vettori reciprocamente indipendenti \mathbf{R} , con vettore delle medie $\mathbf{0}$ e matrice di covarianze diagonale, con varianze sulla diagonale ignote.
- Seconda fase: si trasforma il flusso di vettori \mathbf{R} in una sequenza di vettori reciprocamente indipendenti \mathbf{U} che seguono una distribuzione normale multivariata, $\mathbf{N}(\mathbf{0}, \mathbf{I})$

Possiamo quindi applicare qualsiasi carta multivariata alla sequenza \mathbf{U} per individuare cambiamenti nella sequenza originale \mathbf{X} . Il primo passo è una formulazione standardizzata del problema. La standardizzazione multivariata implica la trasformazione del vettore \mathbf{X} in un vettore normale standard :

$$\mathbf{Z} = \mathbf{A} (\mathbf{X} - \boldsymbol{\mu})$$

dove la matrice \mathbf{A} è tale che: $\mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^T = \mathbf{I}$. Con questa trasformazione $\mathbf{Z} \sim \mathbf{N}(\mathbf{0}, \mathbf{I})$.

Qualsiasi matrice A che soddisfi questa condizione sarà sufficiente per trasformare la \mathbf{X} in una normale multivariata. Con questa scelta se la x_r e z_r sono le r -esime componenti di \mathbf{X} e \mathbf{Z} rispettivamente, allora z_r è il residuo di regressione, quando x_r è ottenuto dai suoi predecessori x_1, x_2, \dots, x_{r-1} , standardizzato per unità di varianza. I risultati della trasformazione \mathbf{Z} sono conosciuti come “*regression – adjusted variables*” raccomandati da Hawkins (1993) per processi cascata.

I residui ricorsivi si basano sulla ripetizione del modello di regressione ad ogni nuovo arrivo dei dati. Essi sono una trasformazione lineare dei residui ordinari ma a differenza di loro sono indipendenti e normalmente distribuiti.

Consideriamo la regressione di x_r da x_1, x_2, \dots, x_{r-1} , e scriviamo la matrice $\mathbf{X}_{n,p}$ degli elementi $x_{r,j}$, $r=1,2,\dots,n$; $j=1,2,\dots,p$ e $x_{r0}=1$.

Indicando con $\mathbf{Y}_{r,j}$ la j -esima colonna della matrice e $x_{r,j}$ la r -esima riga, la regressione multipla è data da:

$$\mathbf{Y}_{r,j} = \mathbf{x}_{r,j-1} * \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_{r,j} \quad (7)$$

Scrivendo come $\hat{\mathbf{x}}_{r,j}$ la stima ai minimi quadrati di $\boldsymbol{\beta}_j$ basata sui primi r vettori. La previsione di $x_{r,j}$ basata sulle osservazioni correnti sui predittori e sul parametro di regressione stimato nell'istante immediatamente precedente, è:

$$\hat{\mathbf{x}}_{r,j} = \mathbf{x}_{r,j-1} * \hat{\boldsymbol{\beta}}_{r-1,j} \quad (8)$$

L' r -esimo residuo ricorsivo della variabile x_j è definito da :

$$r_{r,j} = (x_{r,j} - \hat{\mathbf{x}}_{r,j}) / \sqrt{(1 + h_{r,j})} \quad (9)$$

che si distribuisce come una $N(0, \sigma^2(j))$, dove $\sigma^2(j)$ è la varianza condizionata della regressione di x_j su x_1, x_2, \dots, x_{j-1} e $h_{r,j}$ è:

$$h_{r,j} = \mathbf{x}_{r,j-1} [\mathbf{X}_{r-1,j-1}^T \mathbf{X}_{r-1,j-1}]^{-1} \mathbf{x}_{r,j-1}^T \quad (10)$$

Se supponiamo di aver osservato r vettori del processo e indichiamo con $r_{r,j}$ il residuo ricorsivo ottenuto quando di $\hat{\mathbf{x}}_{r,j}$ è calcolato da $x_{r1}, x_{r2}, \dots,$

$x_{r,j-1}$ a partire dalle regressioni adattate sui vettori casuali 1, 2, ..., r-1, si osserva che $r_{r,j}$ è definito solo per $r \geq j+1$.

E' noto che la sequenza $r_{j+1,j}$, $r_{j+2,j}$, $r_{j+3,j}$,, $r_{r,j}$,, comprende variabili indipendenti $N(0, \sigma^2(j))$. Per questi residui c'è indipendenza longitudinale ma anche laterale. Così i residui ricorsivi sono tra loro indipendenti.

Scrivere i residui in una matrice ci dà il primo passo della trasformazione dal flusso originale di osservazioni con vettore delle medie e matrice di covarianze ignote in un flusso con vettore delle medie $\mathbf{0}$ e matrice diagonale di covarianze. E' chiaro che i residui ricorsivi non possono essere calcolati per i primi casi. Avendo n vettori di osservazioni del processo ci saranno n-p vettori, i primi p non sono completamente definiti, che daranno una matrice \mathbf{R} di dimensione $(n-p) \times p$.

Secondo passo è quello di trasformare la matrice \mathbf{R} in una matrice \mathbf{U} che segue una distribuzione normale multivariata standard $\mathbf{N}(\mathbf{0}, \mathbf{I})$. Si nota che la somma $:(r_{j+1,j})^2 + (r_{j+2,j})^2 + \dots + (r_{r-1,j})^2$ segue una distribuzione $\sigma^2(j)\chi^2_{r-j-1}$ ed è indipendente da $r_{r,j}$.

I residui ricorsivi standardizzati $t_{r,j}$ si possono definire come:

$$t_{r,j} = r_{r,j} / \sqrt{((r_{j+1,j})^2 + (r_{j+2,j})^2 + \dots + (r_{r-1,j})^2) / (r-j-1)} \quad (11)$$

che seguono una distribuzione t con $(r-j-1)$ gradi di libertà.

I $t_{r,j}$ sono trasformati nella sequenza normale standard :

$$u_{r,j} = \Phi^{-1} [F_{r-j-1}(t_{r,j})] \quad (12)$$

per $r \geq j+2$. Dove $F_{r-j-1}(\cdot)$ è la funzione di distribuzione cumulativa della distribuzione t con $(r-j-1)$ gradi di libertà; $\Phi^{-1}(\cdot)$ è l'inversa della distribuzione normale. Le trasformazioni (12) seguono una distribuzione $N(0,1)$ e sono tra loro indipendenti. Il loro calcolo rappresenta il secondo momento dell'approccio di Hawkins durante il quale si definisce la matrice \mathbf{U} di vettori indipendenti $N(0, \mathbf{I})$. In sintesi le trasformazioni possono essere così sintetizzate:

Vettore originale **X di parametri ignoti** → vettori dei residui ricorsivi

R → **T** → **U** vettori normali standard **di parametri noti**.

La matrice **U** è completamente definita solo quando il secondo vettore **R** è completato, cioè al (p+2)-esimo vettore **X**. Avendo n vettori del processo ci saranno (n-p-1) vettori **U**. Si osservi che esiste una stretta connessione tra i vettori **U** e i vettori **Z** che potevano essere ottenuti nel caso in cui i parametri fossero stati noti.

Sia U_r il vettore colonna composto dalla r-esima riga di **U**, al variare di r questi vettori p-dimensionali possono essere monitorati utilizzando una carta di controllo multivariata tradizionale. Hawkins in particolare suggerisce la schema MEWMA di Lowry et. al. (1992):

$$\mathbf{M}_r = \lambda \mathbf{U}_r + (1-\lambda) \mathbf{M}_{r-1}, \quad \text{dove } 0 < \lambda \leq 1 \text{ e } \mathbf{M}_0 = \mathbf{0} \quad (13)$$

Il primo vettore non banale U_r si ottiene per l'osservazione $r=p+2$, nel caso di r osservazioni otteniamo (n-p-1) vettori \mathbf{M}_r non banali, con $\mathbf{M}_r \sim N(\mathbf{0}, \Sigma_{M_r})$. La matrice di covarianze Σ_{M_r} si ottiene della formula:

$$\Sigma_{M_r} = [\lambda/(2-\lambda)] [1-(1-\lambda)^{2(r-p-1)}] \mathbf{I}_p \quad (14)$$

che dà le quantità marginali chi-quadrato:

$$\mathbf{T}_r^2 = \mathbf{M}_r^T \Sigma_{M_r} \mathbf{M}_r = [\lambda/(2-\lambda)] [1-(1-\lambda)^{2(r-p-1)}] \|\mathbf{M}_r\|^2 \quad (15)$$

La carta dà un segnale di fuori controllo non appena $\mathbf{T}_r^2 > h$, dove h è scelto per raggiungere uno specificato ARL in controllo. Tale valore può essere preso dalle tabelle messe a disposizione da Hawkins (2007).

Hawkins trova più conveniente rappresentare graficamente le quantità $\|\mathbf{M}_r\|^2$. Un segnale di fuori controllo è dato non appena queste quantità superano la soglia

$$[\lambda/(2-\lambda)] [1-(1-\lambda)^{2(r-p-1)}] h \quad (4)$$

All'aumentare di r i limiti di controllo tendono a $h\lambda/(2-\lambda)$. Spesso si utilizzano questi limiti di controllo asintotici per aumentare la robustezza

della EWMA . Hawkins non segue questa pratica per una maggiore possibilità di una risposta più rapida ad uno shift precoce.

Questa procedura è chiamata Self Starting MEWMA oppure SSMEWMA. Fatto eccezione per le prime $(p+1)$ osservazioni iniziali, questa SSMEWMA è una MEWMA di quantità indipendenti esattamente $N(\mathbf{0}, \mathbf{I}_p)$, e si avranno esattamente le proprietà RL in controllo dei dati normali multivariati con vettore delle medie e matrice di covarianze note.

2.2 Performance delle carte self-starting multivariate

2.2.1 Misure di performance delle carte self-starting multivariate

Per studiare le prestazioni delle carte di controllo si utilizzano alcune misure di performance quali ad esempio la Run length (RL) ovvero il numero d'osservazioni (o campioni) da estrarre per avere un segnale d'allarme. L'ARL è la media della RL ovvero il tempo medio d'attesa per segnalare un allarme. Si definisce per le due situazioni: un falso allarme o un vero allarme. Così, ARL-IC misura quanto spesso in media ci saranno dei falsi allarmi, mentre ARL-OC quanto velocemente s'individua in media un vero cambiamento.

In molti contesti applicativi, la distribuzione di riferimento per la RL è la distribuzione geometrica, che è completamente caratterizzata dalla sua media. Così anche la RL sarà caratterizzata dall'ARL, e da tutte le altre quantità come i momenti e i percentili.

Un'altra possibile misura di performance è la probabilità di un falso allarme ovvero la probabilità che la carta identifichi una causa assegnabile ma in realtà il processo è stabile.

2.2.2 Disegno delle carte self-starting multivariate

Il disegno della carte introdotte nelle sezioni precedenti consiste nella specificazione di λ e del limite di controllo. La costante λ dipende dalla dimensione dello shift e dal numero di variabili. Si osserva che le migliori scelte di λ per la carta T concordano con i valori precedentemente pubblicati per la progettazione della carta MEWMA (Lowry e.al. 1992, Ranger e Prabhu ,1997).

Per rendere valido il confronto delle prestazioni delle diverse carte di controllo i corrispondenti limiti devono essere scelti in modo che tutte le carte abbiano la stessa ARL in controllo. Una carta di controllo con un basso ARL-OC (fuori controllo) per uno shift di una determinata grandezza dovrebbe essere la migliore. Sullivan (2002) trova errato questo approccio per le carte self starting, mostrando che se i limiti di controllo per le carte U, S e T sono scelti in modo che tutte danno la stessa ARL-IC, la carta T ha prestazioni apparentemente “eccezionali” al diminuire di λ , a prescindere da δ . Ad esempio per un valore di $\lambda = 0.001$ si osserva la migliore prestazione fuori controllo ma con una relativa incidenza di falsi allarmi. Questa mostra che le prestazioni eccezionali della carta sono dovute a una distribuzione della RL con molti *early false alarms* ed equilibrata successivamente. La carta con $\lambda = 0.1$ funziona meglio perchè la sua probabilità di falsi allarmi è inferiore ma subito dopo lo shift la probabilità di segnalare è superiore a quella della carta $\lambda = 0.001$. Utilizzando ARL-IC come criterio per stabilire i limiti di controllo, si trovano limiti di controllo troppo piccoli e di conseguenza le carte tendono presto a segnalare dei falsi allarmi e solo per questo, in fase di confronto, le loro prestazioni sembrano migliori. Le piccole ARL-OC sono realmente legate a numero grande di falsi allarmi e non ad un'individuazione di uno shift. In pratica le carte con molti falsi allarmi precosi possono causare problemi che vanno dal fastidio fino ai costi elevati. Per queste ragioni

s'introduce un nuovo criterio per determinare i limiti di controllo per le carte di controllo self starting multivariate.

Una carta di controllo ideale ha la probabilità di segnalare **zero** prima che accada uno shift e **uno** per prima osservazione shiftata. Poichè in pratica una carta ha necessariamente alcuni falsi allarmi, un approccio ragionevole è di determinare i limiti di controllo basandoci su una probabilità piccola di falsi allarmi (f.a.p.) . La f.a.p. cumulativa per un certo numero fisso d'osservazioni può essere definita come probabilità di almeno un falso allarme fino a un momento specifico. Anche se le statistiche per le carte self starting multivariate sono indipendenti solo per $\lambda=1$ è opportuno scegliere il f.a.p. cumulativo corrispondente a una carta Shewhart: $1 - (1 - 0.0027)^{100} = 0.23689$ all'osservazione 100. La f.a.p. cumulativa non risente come l'ARL della grave asimmetria della distribuzione RL e quindi fornisce limiti di controllo che sono più utili per il confronto delle prestazioni.

La valutazione delle prestazioni della SSMEWMA richiede il valore della costante h del limite di controllo per un ARL desiderato, valori che sono dati nelle tabelle di Hawkins per diverse dimensioni p e costanti di liscio λ . Se avviene uno shift dopo che il processo rimasto in controllo per un periodo sufficientemente lungo le performance della carta SSMEWMA saranno le stesse del caso parametri noti. Ma, se accadono cambiamenti relativamente presto, ci saranno due aspetti da analizzare :

Primo: la dimensione del periodo prima dello shift, quando il processo è in controllo, denotato qui con IC. L'utilizzo delle stime dei parametri al posto dei veri valori riduce la potenza dei test e così, il tempo di risposta ad un cambiamento deve essere più breve del periodo iniziale in controllo.

Secondo: Asintoticamente la carta è invariante affine, cioè la capacità della carta di individuare uno shift di ampiezza δ nel vettore delle medie è

funzione dello scalare di non-centralità $\delta^T \Sigma^{-1} \delta$ e non dipende dall'esatta struttura di δ . In ogni caso, se le componenti di \mathbf{X} sono permutate, si ottengono diversi vettori \mathbf{U} e il tempo di segnalare non sarà solo funzione del parametro di non-centralità ma dipenderà da quali componenti sono shiftate.

Hawkins (2007) mostra i risultati di uno studio di simulazione per dati \mathbf{X} ($p=4$) - per dimensioni piccole e \mathbf{X} ($p=10$) – dimensioni grandi. Dopo un periodo iniziale in controllo di 30-100 osservazioni nel quale il processo funziona bene (le osservazioni sono entro i limiti di controllo), il vettore delle medie è cambiato. Si esaminano 4 tipi di cambiamenti:

- uno shift verso l'alto sulla prima componente
- uno shift verso l'alto sull'ultima componente
- uno shift verso l'alto sulla prima e sull'ultima componente
- uno shift in direzione opposta sulla prima e sull'ultima componente

L'ARL calcolato su 10000 sequenze indipendenti porta alle seguenti osservazioni:

- Nella figura 1(a) si osservano risposte più veloci per le dimensioni più piccole ($p=4$) che per quelle grandi ($p=10$). Questo è anche intuitivo perché la ragione per la reazione più lenta sta nelle imprecisioni delle stime dei parametri. Nel caso di dimensionalità più elevata si richiedono infatti più dati per le stime.
- Ancora dalla figura 1(a), come anche dalla 1(b), per $p=10$, si osservano risposte più veloci quando lo shift è sulla prima componente che sull'ultima. Questo è in contrasto con l'intuizione che ci fa pensare che lo shift sull'ultima componente tragga maggior vantaggio dall'aggiustamento per la regressione della procedura self-starting. La differenza è abbastanza grande nelle fasi iniziali – quando anche i valori stimati con la regressione soffrono di gravi distorsioni per la mancanza dei dati sufficienti, ma poi le curve si avvicinano rapidamente.

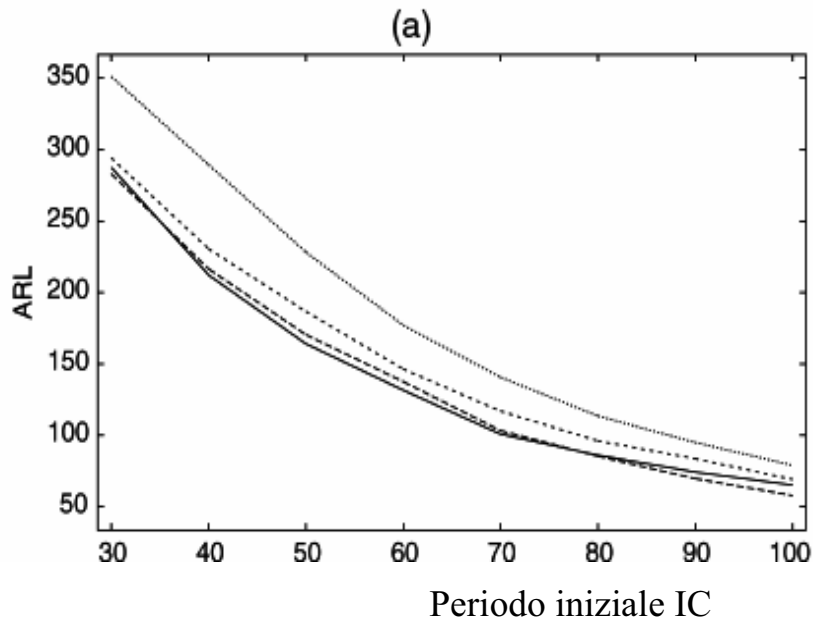


Figura 1: ARL della carta SSEWMA per diversi tipi di perturbazione [(a) $p=4$, prima; $p=4$, ultima; $p=10$, prima; $p=10$, ultima] da Hawkins (2007)

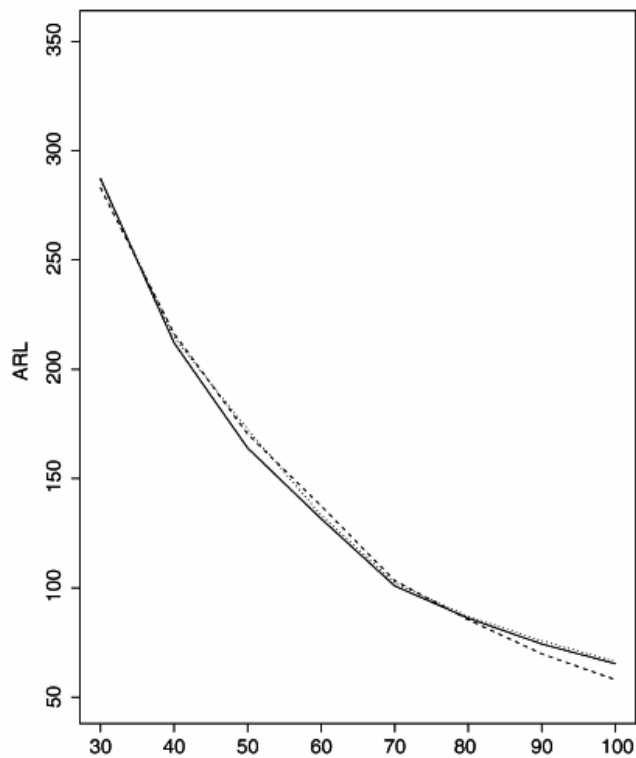


Figura 1: ARL della carta SSEWMA per diversi tipi di perturbazione [(b) $p=4$, prima ; $p=4$, ultima; $p=4$, entrambe] da Hawkins (2007)

- Per $p=4$ l'effetto dell'ordine è appena percepibile, le curve per lo shift sulla prima componente e sull'ultima componente praticamente coincidenti.

- Per periodo in controllo di 100 osservazioni le ARL hanno dimensioni tra 60 – 80, ancora tre volte più grandi del livello asintotico (per il caso di parametri noti Lowry e Woodall (1992) trovano $ARL=20$ per $p=10$ e $ARL=23$ per $p=4$).

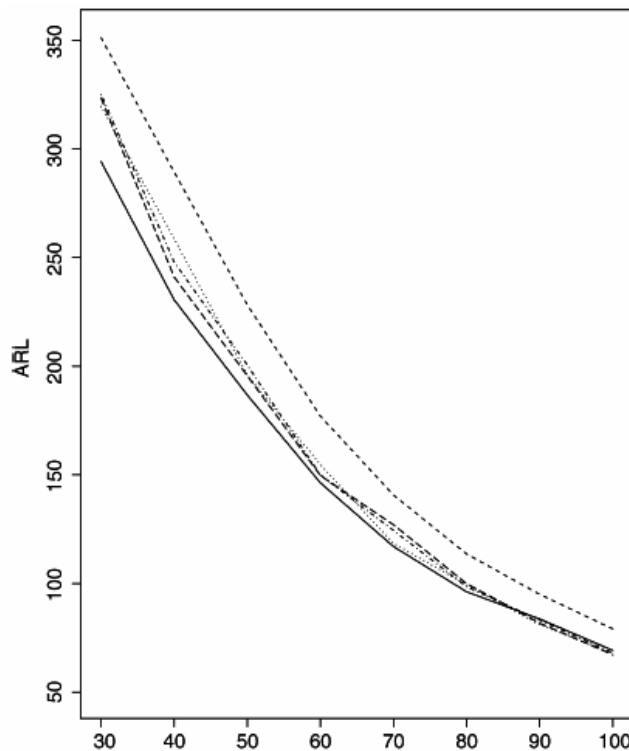


Figura 1: ARL della carta SSEWMA per diversi tipi di perturbazione [(c) $p=10$, prima; --- $p=10$, ultima; $p=10$, entrambe; - . - $p=10$, prima↑, ultima↓; _ _ $p=10$, prima↓, ultima↑] da Hawkins (2007)

Si dovrebbe prendere una sequenza iniziale IC di circa 500 osservazioni per ottenere 10% del valore asintotico. Per quanto questo possa essere un'indicazione dell'impatto sul processo di uno shift molto precoce, esso può essere visto come un commento indiretto delle performance delle procedure standard basate Fase I: raccolta dei dati e stima dei parametri del processo. Questi valori sembrano sostenere infatti che le dimensioni del campione dovrebbero essere di molte centinaia d'unità prima che le stime diventino accurate.

Se $p=10$ ci sono differenze tra le curve per shift sulla prima o sull'ultima componente o per gli shift in direzione opposta. Sembra dunque esserci, nel

caso di dimensioni p più elevate e brevi sequenze IC, un percettibile effetto della permutazione. Questo può avere alcune potenziali implicazioni: se cambiamenti in una particolare componente hanno una maggiore importanza, usare quella componente come prima nei vettori di dati può migliorare la risposta a cambiamenti che si verificano molto presto.

Nel complesso, la SSMEWMA presenta dei vantaggi permettendo sia di sfuggire alla necessità di conoscere i parametri sia di raggiungere prestazioni accettabili una volta accumulata una quantità di dati molto più piccola di quanto sarebbe necessario per un approccio tradizionale di Fase I.

2.3. Prestazioni delle carte basate su “accumulo di vettori” e “accumulo di scalari”

Rappresentando le distribuzioni RL-IC delle tre carte, per diversi valori di λ e dati bivariati, si osserva che la carta U presenta qualche svantaggio. La probabilità di segnalare aumenta bruscamente dopo le prime osservazioni, ciò indica la tendenza ad un numero eccessivo di falsi allarmi precoci. Poiché la distribuzione delle statistiche dipende dal numero di osservazioni accumulate, trasformando le statistiche in modo che siano identicamente distribuite, indipendentemente dal numero di osservazioni, si stabilizza la varianza ottenendo migliori prestazioni IC nelle prime fasi di controllo.

Per confrontare le prestazioni delle carte si fa un confronto delle distribuzioni RL-OC di ogni carta per uno shift che accade all'osservazione 100 (dati bivariati), per diversi valori di λ . Generalmente la carta T ha una maggiore probabilità di segnalare rispetto alle carte S e U. Per $\lambda = 1$ risulta che U ha una pendenza quasi orizzontale indicando una scarsa performance della carta per valori grandi della costante di lisciamiento. Le carte S e T

mostrano una probabilità di segnalare migliore della U per valori grandi di λ , ma hanno generalmente prestazioni migliori per valori piccoli di λ . La mancata reattività delle procedure self starting, nel caso si utilizzino parametri di liscio grandi, fa sì che osservazioni dal processo fuori controllo possano essere utilizzate per l'aggiornamento delle stime correnti con un conseguente ulteriore deterioramento della capacità di segnalare un vero allarme. Questo aspetto non sembra essere discusso nel disegno delle carte self starting.

Con un confronto delle tre carte utilizzando il percentile della RL 0.90, per dati bivariati, diversi valori di λ e lo shift di dimensioni $\{0, 0.5, 1, 1.5\}$, si ottengono i seguenti risultati:

- $\lambda < 0.22$ le carte T e U sono migliori della S;
- La carta T ha migliori prestazioni per tutti valori di λ ;
- $\lambda = 0.046$ la carta T ha migliori prestazioni per il più piccolo cambiamento considerato;
- $\lambda = 0.1$ dà migliori prestazioni per uno shift di 1;
- $\lambda = \{0.22, 0.1\}$ migliori prestazioni per uno shift di $\{1, 1.5\}$

In conclusione la carta T ha le migliori prestazioni globali.

Quando $\lambda=1$ le statistiche sono i.i.d. indifferentemente dalla dimensionalità (p), così $UCL=3$. Per valori più piccoli di λ , le statistiche rimangono identicamente distribuite ma non sono più indipendenti e la distribuzione congiunta cambia leggermente con p . Con aumento della dimensione, la variabilità della distribuzione RL-IC decresce ovvero brevi RL si verificheranno con maggiore probabilità. Si osserva che la probabilità di individuare uno shift, per un valore dato di λ , è maggiore per il caso bidimensionale che per il caso $p=10$, in pratica all'aumentare di p la ARL-OC aumenta. Comunque, può essere inappropriato utilizzare sempre lo stesso valore di λ dal momento che il suo valore ottimale dipende da p .

Le prestazioni della carta T diminuiscono con il periodo di riferimento IC. A causa dell'auto-aggiornamento delle stime dei parametri, i cambiamenti che si verificano molto presto possono influenzare pesantemente le stime mascherando lo shift. Poichè i cambiamenti precoci possono non essere individuati prontamente si raccomanda che questa carta sia completata da un'analisi retrospettiva fatta a una distanza dall'inizio della sorveglianza che dipende dalla dimensione p e dal più piccolo shift la cui identificazione è importante .

2. 4 Confronto della carta SSMEWMA e carta “accumulo di vettori”

E' stato discusso in precedenti studi (Sullivan et al., 2002) che, essendo poco conosciuta la distribuzione RL della MEWMA, l'ARL possa non essere sufficiente a riassumere il comportamento della RL ma sia piuttosto necessaria un'analisi più dettagliata ad esempio dei suoi percentili.

La distribuzione di riferimento per la RL è normalmente la distribuzione geometrica. Questa distribuzione descrive la RL delle carte Shewhart. E' completamente caratterizzata dalla sua media, e così l'ARL, e tutte le altre quantità come i percentili si possono calcolare. Se $f(r) = \text{probabilità}(RL=r)$ e $Q(r)=\text{probabilità}(RL \geq r)$, allora la funzione di rischio è definita da $H(r)=f(r) / Q(r)$. La distribuzione geometrica, si sa, ha una funzione di rischio costante la cui inversa è l'ARL.

Per il metodo di Sullivan (2002) si osserva una funzione di rischio non costante, con code più leggere della geometrica, e in queste circostanze ARL sarebbe una misura appropriata (Figura 2.a)

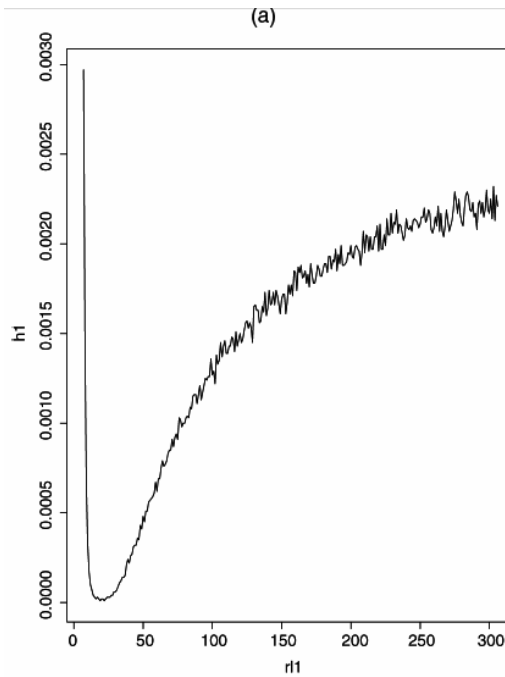


Fig2(a) Funzione di rischio per Metodo Sullivan (Hawkins, 2007)

Per il metodo Hawkins (2007), caso IC, la funzione di rischio scende rapidamente dopo di che resta costante, ciò mostra che, eccetto per periodi iniziali molto brevi, la distribuzione geometrica si adatta perfettamente alla RL-IC della SSMEWMA (Figura 2b). Nel caso fuori controllo (ad esempio uno shift di 1.5 su una componente dopo l'osservazione 100) il rischio non è costante, ma è basso per entrambe sequenze IC lunghe e corte, indicando nuovamente che l'ARL è una misura affidabile.

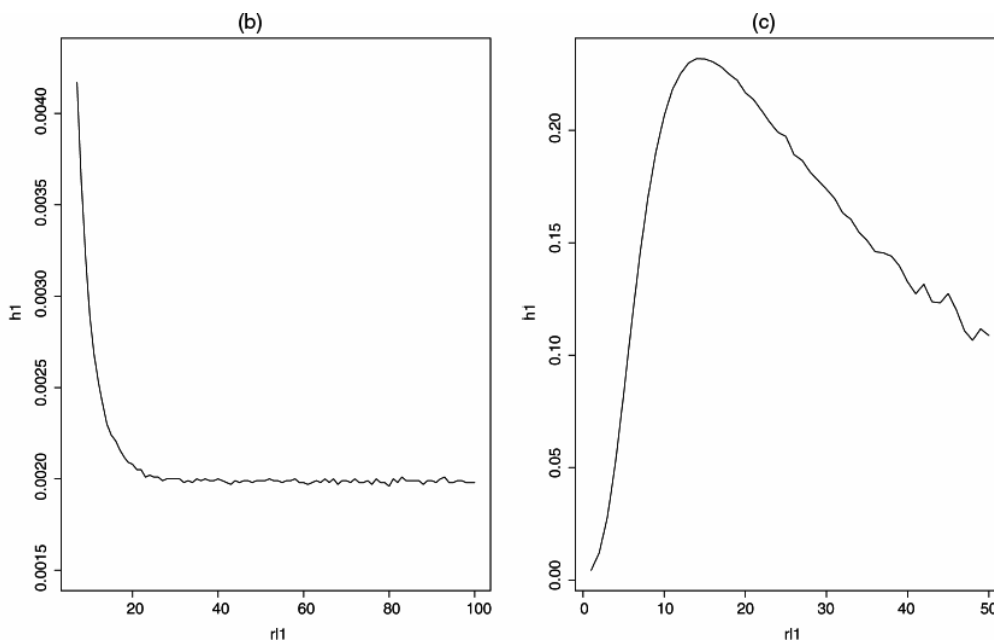
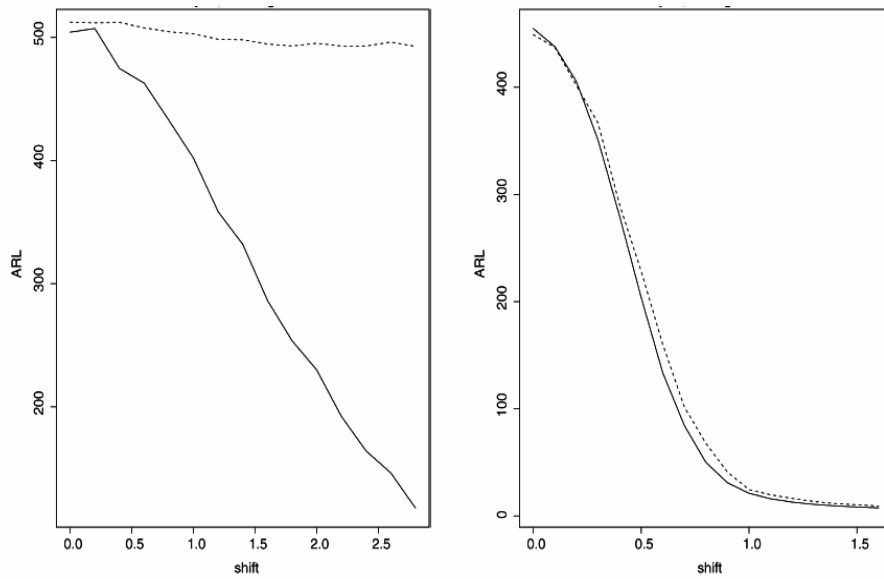


Fig2(b)Funzione di rischio SSMEWMA IC, (c) SSMEWMA OC; Hawkins, 2007)

Entrambi i metodi di Sullivan (2002) e di Hawkins (2007) convergono a una MEWMA di parametri noti, le loro potenziali differenze sono limitate al comportamento short-run (breve termine) e al possibile impatto dei metodi diversi per fissare i limiti di controllo. In questo senso Hawkins fa un confronto tra questi metodi per periodi brevi di 10 osservazioni e di 100 osservazioni per shift di vari dimensioni sulla prima componente. Per il confronto si deve garantire che entrambi metodi abbiano lo stesso ARL-IC ed ancora che il rischio non costante del metodo di Sullivan implica che ARL-IC dipende dal numero d'osservazioni incorporate senza segnale. Questo implica l'utilizzo di diversi limiti di controllo per raggiungere lo stesso ARL-IC via via che aumenta il numero d'osservazioni disponibile. Ad esempio, il valore $h=2.742$ per 10 osservazioni iniziali IC dà un'ARL-IC=500, mentre per 100 osservazioni iniziali IC da un ARL_IC=449. In altre parole se il periodo IC è molto breve, per esempio di 10 osservazioni, si osserva che la SSMEWMA risponde più rapidamente all'aumento dello shift nella media mentre il metodo Sullivan si rivela insensibile. D'altra parte, ogni volta che il periodo in controllo aumenta il vantaggio delle prestazioni della SSMEWMA svanisce.



(b)

Figura 3 Confronto delle prestazioni tra metodo Sullivan(.....) e SSMEWMA(_____) (a)IC=10, (b)IC=100 , (Hawkins,2007)

2.5 Discussione

La metodologia delle carte tradizionali comporta un'attenta raccolta e un' analisi del campione di Fase I. Una volta che quest'analisi è completata, la fase II di controllo inizia e, almeno teoricamente, non vi è alcun'ulteriore analisi sui veri parametri del processo. I metodi Self Starting rimuovono la necessità del campione di Fase I, il monitoraggio del processo inizia quasi subito, e non smettono mai di aggiornare i parametri del processo e di aumentare la precisione delle stime con l'aumento dei dati osservati.

Però, ci sono un certo numero di aspetti che richiedono attenzione. Una di queste è la reazione ad un segnale. Molto di più che nei casi tradizionali, per la carta Self Starting è essenziale diagnosticare i cambiamenti, quando si sono verificati e analizzare i dati raccolti durante un regime modificato, tenendoli fuori dal processo d'aggiornamento. Se ciò non avverrà allora le osservazioni shiftate saranno incorporate nelle stime dei parametri,

provocando una distorsione delle medie e gonfiando le varianze, risultando in controllo dei dati che in effetti non lo sono.

Un altro problema delle carte self starting è che le loro prestazioni diminuiscono per cambiamenti molto precoci. Se il processo è stato in controllo solo per un breve periodo e poi è successo un cambiamento, allora le stime saranno distorte, in particolare le varianze tendono ad essere gonfiate. Poiché questi shift precoci possono sfuggire all'individualizzazione, la carta non è in grado di recuperare, e non sarà neanche in grado di individuare grandi shift successivi. Questo ci suggerisce che la raccolta e il monitoraggio delle prime osservazioni dovrebbero essere fatti con cura abituale per gli studi di fase I, con l'unica differenza che questa fase non deve durare molto a lungo.

Capitolo 3

Confronto tra due carte self-starting multivariate

3.1 Le carte di controllo per “accumulo di vettori “

Si presentano alcuni risultati sulle prestazioni delle carte self starting multivariate, ottenute utilizzando programmi R presentati nell'Appendice.

3.1.1 Esempio per confronto tra le due carte U e T

Sullivan (2002), sfruttando un esempio da Holmes e Mergen (1993), vuole mostrare che carta T (accumulo di vettori trasformati) segnala più presto della carta U (accumulo di vettori trasformati). Si considerano 56 osservazioni individuali bivariate, da un impianto che produce sabbia, che rappresentano le percentuali di particelle di misura grande e media (peso). Da un'analisi grafica preliminare si trova che le prime 25 osservazioni sono diverse dalle ultime 31 e si presume un cambiamento del vettore delle medie dopo l'osservazione 25.

Osservazione: anche se i dati sono lontani da una distribuzione normale multivariata, la MEWMA è robusta alla non-normalità quando si usa una costante di liscio piccola ($r=0.1$)

Le prime 3 statistiche non possono essere calcolate, la MEWMA inizia dall'osservazione 4. La prima osservazione che sorpassa il limite di controllo della carta U è l'osservazione 52 mentre la carta T segnala il cambiamento all'osservazione 29. (Figura4(a): Carta U;Figura4(b):Carta T).

Da quest'esempio si osserva che la carta per “accumulo di vettori trasformati “ segnala più veloce della carta per “accumulo di vettori non trasformati”.

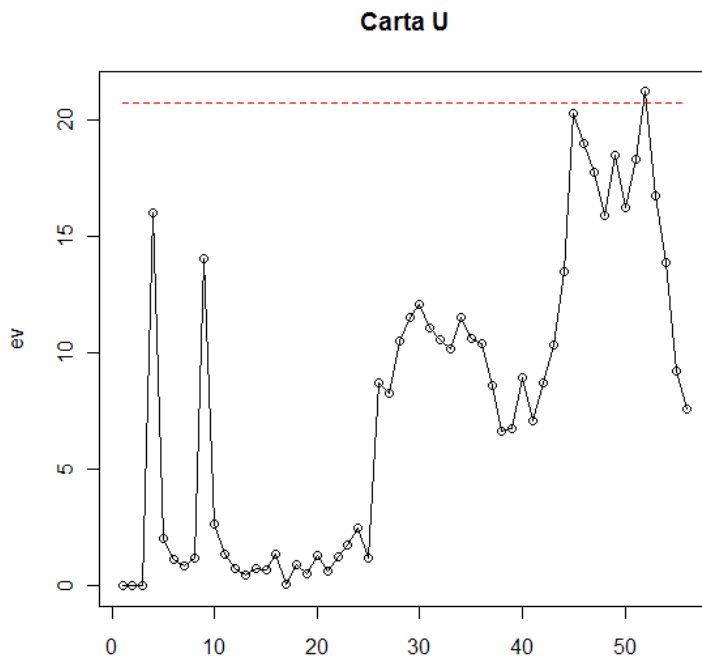


Figura4(a): Carta U

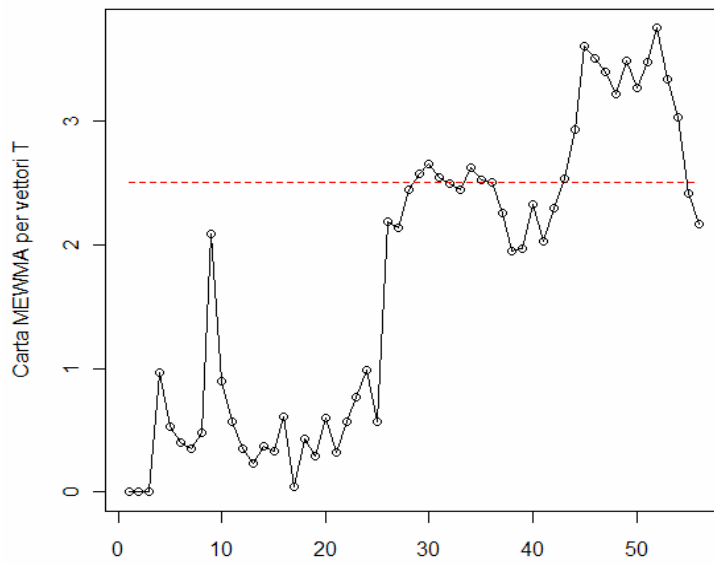


Figura4(b):Carta T

3.1.2. Confronto delle prestazioni al variare del periodo IC

Generiamo T dati da una distribuzione normale bivariata $N(\mathbf{0}, \mathbf{I}_2)$. All'osservazione T+1 la media subisce un cambiamento, si generano in questo caso osservazioni da una normale $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Si desidera vedere come la dimensione del periodo IC influisce la RL. La Tabella 1 riporta il numero di osservazioni fino alla prima segnalazione compreso e al netto del periodo in controllo (RLT e RL rispettivamente). Dai dati ottenuti si osserva che le prestazioni della carta T migliorano con aumento del numero di dati in controllo.

Tabella 1

<i>Periodo IC</i>	<i>RLT</i>	<i>RL</i>
10	33	23
30	43	13
50	62	12
100	<i>Falsi allarmi</i>	

3.1.3. Simulazioni

Si stimano le prestazioni della carta T, tramite simulazioni di 10000 osservazioni per diverse dimensioni del periodo $IC = \{10, 20, 30, 50, 75, 100\}$. Anche in questo caso si osserva che con un aumento del periodo IC le prestazioni della carta migliorano (Figura 5). Nella Tabella 2 si riportano, in termini di ARL, i risultati ottenuti per i periodi di controllo enunciati sopra.

Tabella 2

<i>IC</i>	10	20	30	50	75	100
<i>ARL</i>	23.895	18.850	16.603	14.257	12.702	12.088

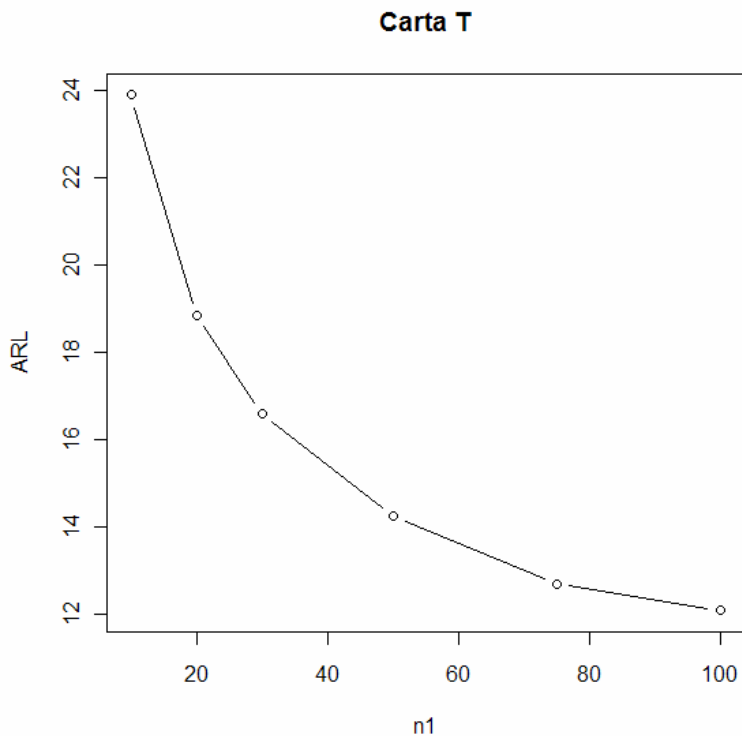


Figura 5: Le prestazioni della carta T al variare del periodo IC

3.2 Carta di controllo SSMEWMA

Si generano 30 dati da una normale multivariata di dimensione $p=5$, $N(\mu_1, I_5)$, e altri 30 da $N(\mu_2, I_5)$, dove $\mu_1=(0,0,0,0,0)$ e $\mu_2=(1,0,0,0,0)$, vale a dire si ipotizza uno shift sulla prima componente. La carta MEWMA per questi dati è rappresentata nella Figura 6

Si ottiene $RL_1 = 37$, e tenendo conto che le prime $m+1$ osservazioni non vengono riportate sulla carta MEWMA, la prima osservazione fuori controllo sarà la $43(=37+6)$ -esima

La $RL = 13$ è il numero d'osservazioni necessarie per individuare il cambiamento.

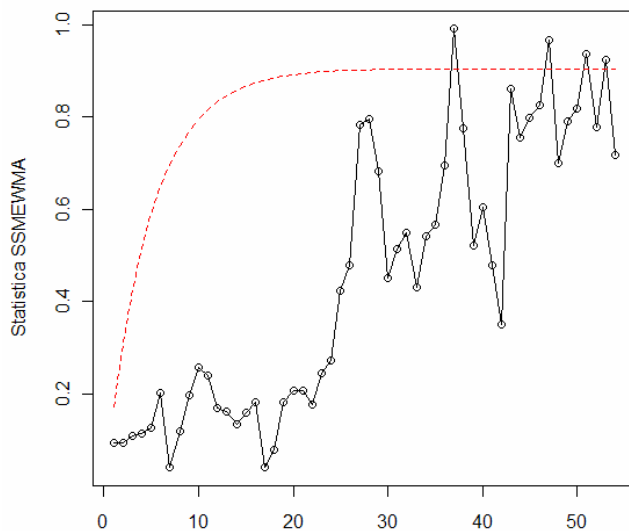


Figura 6: Carta SSMEWMA

3.2.1 Simulazioni per la carta SSMEWMA

Il programma di simulazioni per stimare l'ARL presentato qui non è efficiente dal punto di vista computazionale, richiedendo troppo tempo a causa dei modelli da stimare per calcolare la matrice dei residui. Così le vere prestazioni della carta non possono essere calcolate. Comunque è possibile dare un'idea su come le prestazioni della carta SSMEWMA cambiano al variare:

- della dimensione dello shift
- della struttura dello shift
- del numero di variabile(caratteristiche), qui m .
- al variare del periodo IC

Per 100 simulazioni, si calcolano le seguenti caratteristiche della RL:

- ||| l'ARL, numero medio d'osservazioni necessarie per individuare un cambiamento
- ||| NA il numero di volte che la carta non segnala entro un numero d'osservazioni dallo shift
- ||| $RL_{sim} < 0$ il numero di volte che la carta da falsi allarmi

III $RLsim > 0$ il numero di volte che la carta segnala.

Per quanto riguarda la dimensione dello shift (vedi Tabella XX) si osserva che ad aumentare dello shift le prestazioni della carta aumentano.

Tabella 3

<i>Shift</i>	<i>ARL</i>	<i>N</i> <i>A</i>	<i>RLsim</i> <i>< 0</i>	<i>RLsim</i> <i>> 0</i>
1	15.29	46	3	51
1.56	13.22	27	5	68
2	12.64	21	10	69

Tabella 4

Shift	NA	Falsi allarmi	ARL
↑prima comp.	27	5	13.22
↑ultima comp.	35	10	13.50
↑tutte comp	21	10	12.64
↓prima comp	27	9	12.23
↓ultima comp	31	8	15.45

Con riferimento alla struttura dello shift (vedi Tabella 4) Si osserva che lo shift sull'ultima variabile è segnalato più tardi . Ovviamente uno shift su tutte le componente è segnalato prima poiché aumenta anche il parametro di non-centralità. Come atteso inoltre:

- uno shift negativo sulla prima componente è segnalato prima di uno shift positivo (ARL=12.23 contro 13.22)

- uno shift negativo sull'ultima componente è segnalato piu' tardi uno di uno shift positivo sull'ultima (ARL=15.45 contro 13.50). Bisogna aumentare numero di simulazione per concludere.

Si osserva ancora che con aumento del numero di variabili (vedi Tabella5) la segnalazione dello shift in media avviene più tardi.

Tabella5

Numero variabile	2	3	4	5
ARL	13.72	14.35	15.31	15.29

Infine, Al variare del periodo in controllo $IC=\{10,20,30,50,75,100\}$, con 100 simulazioni per ogni caso si ottengono seguenti risultati(Tabella 6):

IC	10	20	30	50	75	100
ARL	3.235	11.685	10.692	10.015	10.650	9.222

che mostrano che col aumento del periodo IC dovrebbero aumentare le prestazioni della carta.(Figura 7)

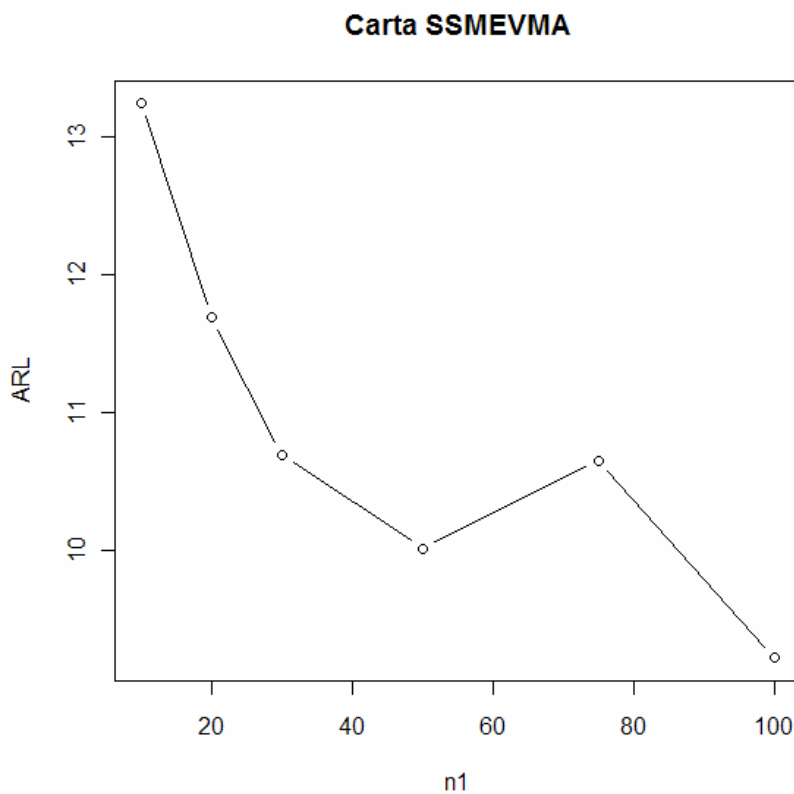


Figura 7:Carta SSMEVMA

3.2.2 Confronto dei metodi Sullivan e Hawkins

Per il confronto delle due carte di controllo Self Starting: Hawkins(SSMEWMA) e Sullivan(Carta T), si devono scegliere i limiti di controllo in modo che le due carte abbiano lo stesso ARL-IC. Qui si è scelto un $ARL-IC=200$. Per entrambe carte si generano dati normali bivariati con media $\mathbf{0}$ e matrice di covarianze identità(\mathbf{I}_2). Dopo un periodo IC, le medie subiscono uno shift sulla prima componente di 1 (con un parametro di non-centralità pari a 1). Con questi dati, dalle tabelle fornite dai due autori si trova $h=8.786$ per la SSMEWMA di Hawkins e $h=2.5082$ per la carta T.

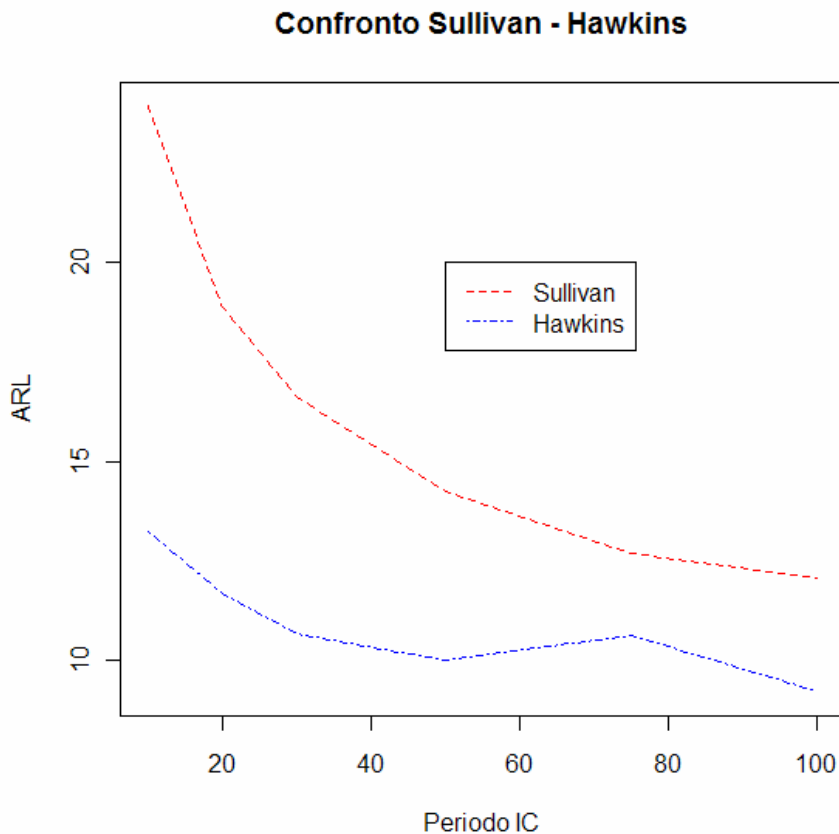


Figura8:Confronto Sullivan-Hawkins

I risultati per metodo Sullivan sono ottenuti da 10000 simulazioni per ogni periodo IC, così possono essere considerati una vera stima delle prestazioni della carta. I risultati per il metodo di Hawkins invece sono riportati per 100 simulazioni, così i risultati non rappresentano accuratamente le vere prestazioni. In questa situazione si osserva che la SSMEWMA risponde più rapidamente ad uno shift molto precoce.

Conclusioni

In questa tesi si è cercato di presentare dei metodi Self –Starting per monitorare in modo efficiente processi di produzione di basso volume. In questi ambienti non ci sono sufficienti dati disponibili per stimare parametri del processo necessari per le procedure tradizionali.

Nel caso univariato, gli approcci di Quessenberry(1991,1995) e Hawkins(1987) si basano principalmente sulle trasformazioni dei dati originali, di parametri ignoti in statistiche normale standard chiamate statistiche Q. Eliminando così il problema dei parametri ignoti, le statistiche Q possono essere monitorate con le carte Shewhart, CUSUM o EWMA. Si è osservato che ognuna di queste carte per le statistiche Q ha delle prestazioni migliori con aumento del periodo in controllo e la dimensione dello shift.

Se dovessimo monitorare più caratteristiche, piuttosto di osservarle separatamente, si potrebbero ottenere delle performance migliori utilizzando carte multivariate. Sono state presentate le proposte di Quesenberry(1997), Schaffer(1998) e Sullivan (2002). Mentre i primi tre trasformano i dati multivariati di parametri ignoti in statistiche scalare normale standard, Hawkins utilizza i residui di regressione standardizzati. È stato mostrato che la carta per “accumulo di vettori” di Sullivan e la carta SSMEWMA di Hawkins hanno delle prestazioni migliori, per ciò si è cercato nel Capitolo 3, utilizzando i programmi R del appendice di stimare le prestazioni di queste carte. È stato osservato, che come nel caso univariato, che l’aumento del periodo in controllo o della dimensione dello shift possono migliorare le performance delle carte. Se nel caso di “accumulo di vettori” la carta non dipende dalla struttura dello shift, per la carta MEWMA questa situazione sembra a cambiare nel caso di dimensioni

grandi. Anche la dimensione dei dati da monitorare può influire sulle prestazioni delle carte, così con dati di dimensioni piccole si ottengono delle prestazioni migliore.

Si è cercato anche di fare un confronto tra le carte di Sullivan e la carta SSMEWMA di Hawkins, Si osserva che le prestazioni della carta di Hawkins potrebbero essere migliori. Purtroppo il programma R per questo ultimo non è efficiente dal punto di vista computazionale non permettendo di fare tanti simulazioni come richiesto per delle vere stime. In quanto il problema deriva dalla stime dei tanti modelli per calcolare i residui di regressione un miglioramento si potrebbe forse ottenere utilizzando un algoritmo di regressione basato sulla decomposizione QR che permetterebbe di aggiornare la matrice dei residui senza stimare effettivamente i modelli di regressione.

Appendice

```
##### Modello Sullivan #####

n<-nrow(dati)
m<-ncol(dati)

#####          vettori          delle          medie
#####
#la funzione "medie" da i vettori delle medie, organizzati
in #una matrice utilizzando le formule di aggiornamento di
#Quesenberry
#j-esima riga della matrice e il vettore delle medie
stimato #per oss (j+1)-esima ed è la media delle prime j
osservazioni

medie<-function(dati)
{
n<-nrow(dati)
m<-ncol(dati)
media<-matrix(,n,m)
for (i in 1:n){
if (i==1) media[1,]<-dati[1,]
else
media[i,]<-1/i*((i-1)*media[i-1,]+dati[i,])
}
media
}
media<-medie(dati)

#####Le matrici D, U, delle innovazioni #####

#d(j)=x(j)-media(x1,...x(j-1)), sono le deviazioni delle
oss #correnti dalla media di tutte le precedenti oss
#u(j)=d(j)*(j-1)/j
#la funzione "innov" mi dà la matrice D delle d(j)
#la funzione "inovU" mi dà la matrice U delle u(j)

innov<-function(dati)
{
n<-nrow(dati)
m<-ncol(dati)
inov<-matrix(,n,m)
#U<-matrix(,n,m)
inov[1,]<-dati[1,]
for (i in 2:n){
inov[i,]<-(dati[i,]-media[i-1,])
}
inov
}
}
```

```

D<-innov(dati)
inovU<-function(dati){
n<-nrow(dati)
m<-ncol(dati)
U<-matrix(,n,m)
for (i in 1:n){
U[i,]<-D[i,]*sqrt((i-1)/i)
}
U
}
U<-inovU(dati)
U

#####"accumulo di vettori"#####
##### utilizzando un modello MEWMA #####

#Z[j,]<-r*U[j]+(1-r)*Z[j-1,]
#Z0=(0,0,....,0)
#Ottengo la matrice Z,dei vettori cumulati con MEWMA

matZ<-function(r){
n<-nrow(dati)
m<-ncol(dati)
Z<-matrix(,n,m)
for (i in 1:n){
if (i==1)
Z[1,]<-rep(0,m)
else
Z[i,]<-r*U[i,]+(1-r)*Z[i-1,]
}
Z
}
Z<-matZ(r)

##### Stimo la matrice di covarianze S
#####
#si stima per j>=m+2 ed è la stima della matrice di cov per
#oss j+1
#Sj=var(dati[1:j,])

#####matrice covarianze(Z)#####

#la funzione "covZ" stima la matrice delle covarianze delle
Z
#per il vettore cumulato j-esimo
#si utilizza la matrice di covarianze stimata S(j-1)

covZ<-function(j,r){
sigmaZ<-r*(1-(1-r)^(2*j))/(2-r)*var(dati[1:j-1,])
sigmaZ

```



```

}

##### statistiche dei vettori non trasformati
#####

#la funzione "statU" calcola le statistiche EV , della
"carta #U" si possono calcolare per j>=m+2

statU<-function(r){
n<-nrow(dati)
m<-ncol(dati)
EV<-rep(0,n)
p<-m+2
for (j in p:n){
sigma<-covZ(j,r)
EV[j]<-t(Z[j,])%*(solve(sigma))%*Z[j,]
}
EV
}
ev<-statU(r)
##### Grafico della carta U #####
#la funzione "graf" disegna la "carta U",
#si devono dare i parametri: r=pesso per la oss.
corrente,h=UCL
#nel esempio di Sullivan(dati bivariati, ARL0=200),
h=20.719, r=0.1
#wmewma<-statU(0.1)
hu=20.719
r=0.1

graf<-function(r,hu){
wmewma<-statU(r)
lim<-hu
s<-max(c(lim,max(wmewma)))
i<-min(c(lim),min(wmewma))
matplot(cbind(wmewma,lim),ylab="ev",main="Carta U",
ylim=c(i,s),type="l",lty=1:2)
points(wmewma)
y<-data.frame(lim,wmewma)
y
}
graf(r,hu)

##### RL carta U #####
rlU<-function(ev,hu){
n<-length(ev)
ind<-(1:n)[ev>=hu]
rl<-ind[1]
rl
}

```

```

RLU<-rlU(ev,hu)
RLU

#####Statistiche di vettori trasformati#####

#la funzione "statT" calcola le statistiche ET della carta
T

statT<-function(r){
n<-nrow(dati)
m<-ncol(dati)
ET<-rep(0,n)
p<-m+2
F<-rep(0,n)
#ev<-statU(r)
for (j in p:n){
F[j]<-(j-m-1)/(m*(j-2))*ev[j]
ET[j]=sqrt(qchisq(pf(F[j],m,j-m-1),1))
}
ET
}
et<-statT(r)

##### RL carta T #####

rlT<-function(et,ht){
n<-length(et)
ind<-(1:n)[et>ht]
rl<-ind[1]
rl
}
RLT<-rlT(et,ht)
RLT

#####Grafico della carta T#####

#la funzione "grafT" disegna la "carta T",
#si devono dare i parametri: r=peso per la oss.
#corrente,h=UCL. Nell'esempio di Sullivan(dati bivariati,
#ARL0=200), #h=2.5082, r=0.1
#wmewma<-statU(0.1)

r=0.1
ht=2.5082

grafT<-function(r,ht){
wmewmaT<-statT(r)
lim<-ht
s<-max(c(lim,max(wmewmaT)))
i<-min(c(lim),min(wmewmaT))

```

```

matplot(cbind(wmewmaT,lim),ylab="et",main="Carta MEWMA per
vettori T",
ylim=c(i,s),type="l",lty=1:2)
points(wmewmaT)

```

```

y<-data.frame(lim,wmewmaT)
Y
}
grafT(r,ht)

```

```

#####Dati#####

```

```

library(MASS)
sigma<-Iden(2)
mu1<-rep(0,2)
mu2<-c(1,0)
n1<-10
X1<-mvrnorm(n1,mu1,sigma)
X2<-mvrnorm(30,mu2,sigma)
dati<-rbind(X1,X2)
dati

```

```

#####Parametro di noncentralita delta #####
shift<-mu2-mu1
delta<-sqrt(t(shift)%*%solve(sigma)%*%shift)
delta

```

```

# la funzione "Iden" da la matrice identità di ordine p
Iden<-function(p) {
V<-matrix(,p,p)
for (i in 1:p){
for(j in 1:p){
if (j==i) V[i,j]=1
else V[i,j]=0
}
}
V
}

```

```

#####Riassunto delle funzioni da
chiamare#####

```

```

n<-nrow(dati)
m<-ncol(dati)
r<-0.1
ht<-2.5082
media<-medie(dati)
D<-innov(dati)
U<-inovU(dati)
Z<-matZ(r)
ev<-statU(r)

```

```

et<-statT(r)
RLT<-rlT(et,ht)
RLT
grafT(r,ht)

#####Simulazioni#####

mul<-rep(0,2)
mu2<-c(1,0)
p<-length(mu1)
sigma<-Iden(p)

r=0.1
ht=2.5082
RLT<-c()

ARLT<-c()
ARL<-c()

n1<-c(10,20,30,50,75,100)
k<-length(n1)
for (i in 1:k){
for (j in 1:10000){
X1<-mvrnorm(n1[i],mul,sigma)
X2<-mvrnorm(50,mu2,sigma)
dati<-rbind(X1,X2)
n<-nrow(dati)
m<-ncol(dati)
media<-medie(dati)
D<-innov(dati)
U<-inovU(dati)
Z<-matZ(r)
ev<-statU(r)
et<-statT(r)
RLT[j]<-rlT(et,ht)
}
RL<-RLT-n1[i]
ARLT[i]<-mean(na.omit(RLT[RLT>0]))
ARL[i]<-mean(na.omit(RL[RL>0]))
}
ARLT
ARL
plot(n1,ARL,main="Carta T",type="b")

```

```
##### Dati per Esempio 1 #####
##### Holmes and Mergen (1993) #####
X1<-
c(5.4,3.2,5.2,3.5,2.9,4.6,4.4,5,8.4,4.2,3.8,4.3,3.7,3.8,2.6,2.7,
7.9,6.6,4,2.5,3.8,2.8,2.9,3.3,7.2,7.3,7,6.0,7.4,6.8,6.3,6.1,6.6,
6.2,6.5,6,4.8,4.9,5.8,7.2,5.6,6.9,7.4,8.9,10.9,8.2,6.7,5.9,8.7,6
.4,
8.4,9.6,5.1,5,5,5.9)
X2<-
c(93.6,92.6,91.7,86.9,90.4,92.1,91.5,90.3,85.1,89.7,92.5,91.8,91
.7,90.3,94.5,94.5,88.7,84.6,90.7,90.2,92.7,91.5,91.8,90.6,87.3,7
9,82.6,83.5,83.6,84.8,87.1,87.2,87.3,84.8,87.4,86.8,88.8,89.8,86
.9,83.8,89.2,84.5,84.4,84.3,82.2,89.8,90.4,90.1,83.6,88,84.7,80.
6,93,91.4,86.2,87.2)
dati<-cbind(X1,X2)
dati
```

```
#####Carta SSMEWMA Hawkins #####
```

```
#####Matrice Identità#####
```

```
Iden<-function(p) {
V<-matrix(,p,p)
for (i in 1:p){
for(j in 1:p){
if (j==i) V[i,j]=1
else V[i,j]=0
}
}
V
}
```

```
##### Matrice di covarianze #####
sigmaf<-function(cor,p) {
sigma<-matrix(,p,p)
for(i in 1:p){
for (j in 1:p){
sigma[i,j]<-cor^abs(i-j)
}
}
return(sigma)
}
```

```
##### Dati #####
#Dati:
mu1<-rep(0,5)
mu2<-c(1,0,0,0,0)

sigma<-sigmaf(0,5)  #=Iden(5)
```

```

library(MASS)
X1<-mvrnorm(30,mu1,sigma)
X2<-mvrnorm(30,mu2,sigma)
X<-rbind(X1,X2)
dati<-data.frame(X)

#####Parametro noncentralità#####

#necesario per scegliere il limite della carta (parametro
h)
#delta=1, ARL=500, m=5 -> h=17.186
#funzione "parametro"
parametro<-function(mu1,mu2,sigma)
{
delta<-mu2-mu1
paramnonc<-t(delta)%*%solve(sigma)%*%delta
paramnonc
}
parametro(mu1,mu2,sigma)

#####
n<-nrow(dati)
m<-ncol(dati)

##### Funzione R1 #####
#per ottenere prima colonna della matrice R dei residui
#r(1,i)=[x(1,i+1)-x(1,i)] / sqrt(1+1/i)

R1<-function(dati){
n<-nrow(dati)
r1<-c()
for (i in 1:n-1){
{
xlhat<-sum(dati[1:i,1])/i
r1[i+1]<-(dati[i+1,1]-xlhat)/sqrt(1+1/i)}
}
r1
}

##### Funzione R2 #####
##### Residui per X2 #####

#mod2 = modello di regressione lineare X2~X1
#residui per colonna X2:r2(i)= (x2(i)-
#hat.x2(i))/sqrt(1+h2(i))
#dove h2(i)=x[i,2]*solve(X'[i-1,2]X[i-1,2])*x'[i,1]
#x[i,1]:1,xil
#X[i-1,1:2]:matrice delle prime i-1 righe, prime due
colonne #di X (prima colonna di X è 1, seconda = prima
caratteristica #da monitorare)

```

```

R2<-function(dati) {
n<-nrow(dati)
r2<-c()
for (i in 2:n){
if (i<n)
{
mod2<-lm(X2~X1,data=dati[1:i,1:2])
x2hat<-predict(mod2,newdata=data.frame(X1=dati[i+1,1]))
a<-X[i+1,1:2]
b<-X[1:i,1:2]
h<-t(a)%*(solve(t(b)%*b))%*a
r2[i+1]<-(dati[i+1,2]-x2hat)/sqrt(1+h)
}
i<-i+1
}
r2
}

##### Funzione R3 #####
#####Residui per X3 #####

#mod3 = modello di regressione lineare X3~X1+X2
#residui per colonna X3:r3(i)= (x3(i)-
#hat.x3(i))/sqrt(1+h3(i))
#dove h3(i)=x[i,2]*solve(X'[i-1,2]X[i-1,2])*x'[i,2]
#x[i,2]:1, xi1, xi2
#X[i-1,3]:matrice delle prime i-1 righe,prime 3 colonne di
X
#(1,prima caratteristica, seconda caratteristica)

R3<-function(dati) {
n<-nrow(dati)
r3<-c()
for (i in 3:n){
if (i<n){
mod3<-lm(X3~X1+X2,data=dati[1:i,1:3])
x3hat<-predict(mod3,newdata=data.frame(X1=dati[i+1,1],X2<-
dati[i+1,2]))
a<-X[i+1,1:3]
b<-X[1:i,1:3]
h<-t(a)%*(solve(t(b)%*b))%*a
r3[i+1]<-(dati[i+1,3]-x3hat)/sqrt(1+h)
}
i<-i+1
}
r3
}

```

```

##### Funzione R4 #####
#####Residui per X4 #####

#mod4 = modello di regressione lineare X4~X1+X2+X3
#residui per colonna X4:r4(i)= (x4(i)-
#hat.x4(i))/sqrt(1+h4(i))
#dove h4(i)=x[i,3]*solve(X'[i-1,3]X[i-1,3])*x'[i,3]
#x[i,4]:1, xi1, xi2, xi3
#X[i-1,4]:matrice delle prime i-1 righe,prime 4 colonne di
X
#(1,prima caratteristica, seconda caratteristica,terza
#caratteristica)

R4<-function(dati){
n<-nrow(dati)
r4<-c()
for (i in 4:n){
if (i<n)
{
mod4<-lm(X4~X1+X2+X3,data=dati[1:i,1:4])
x4hat<-predict(mod4,newdata=data.frame(X1=dati[i+1,1],X2<-
dati[i+1,2],X3<-dati[i+1,3]))
a<-X[i+1,1:4]
b<-X[1:i,1:4]
h<-t(a)%*(solve(t(b)%*%b))%*%a
r4[i+1]<-(dati[i+1,4]-x4hat)/sqrt(1+h)
}
i<-i+1
}
r4
}

##### Funzione R5 #####
#####Residui per X5 #####

#mod5 = modello di regressione lineare X5~X1+X2+X3+X4
#residui per colonna X5:r5(i)= (x5(i)-
#hat.x5(i))/sqrt(1+h5(i))
#dove h5(i)=x[i,5]*solve(X'[i-1,5]X[i-1,5])*x'[i,5]
#x[i,5]:1, xi1, xi2, xi3,xi4
#X[i-1,5]:matrice delle prime i-1 righe,prime 5 colonne di
X
#(1,prima, seconda, terza, quarta caratteristica)

R5<-function(dati){
n<-nrow(dati)
r5<-c()
for (i in 5:n){

```



```

if (i<n)
{
mod5<-lm(X5~X1+X2+X3+X4,data=dati[1:i,1:5])
x5hat<-predict(mod5,newdata=data.frame(X1=dati[i+1,1],X2<-
dati[i+1,2],X3<-dati[i+1,3],X4<-dati[i+1,4]))
a<-X[i+1,1:5]
b<-X[1:i,1:5]
h<-t(a)%*(solve(t(b)%*%b))%*%a
r5[i+1]<-(dati[i+1,5]-x5hat)/sqrt(1+h)
}
i<-i+1
}
r5
}

```

```

##### matrice R dei residui#####
#le prime m righe della matrice non sono complete

```

```

Rf<-function(dati){
R<-matrix(,n,m)
R[,1]<-R1(dati)
R[,2]<-R2(dati)
R[,3]<-R3(dati)
R[,4]<-R4(dati)
R[,5]<-R5(dati)
R
}
R<-Rf(dati)

```

```

##### T , U #####

```

```

#T dei residui indipendenti ma non identicamente
distribuiti
#U dei residui iid normali standard
#U è completamente definita alla (m+2)-esima osservazione

```

```

Uf<-function(dati){
n<-nrow(dati)
m<-ncol(dati)
T<-matrix(,n,m)
U<-matrix(,n,m)
for (j in 1:m){
for (i in (j+2):n){
T[i,j]<-R[i,j]/sqrt((sum(R[(j+1):(i-1),j]^2))/(i-j-1))
U[i,j]<-qnorm(pt(T[i,j],i-j-1))
}
}
U
}
U<-Uf(dati)

```

```

##### U completa #####
UC<-U[(m+2):n,]
z<-nrow(UC)

#####SSMEWMA#####
# per ottenere i vettori MEWMA : Mn
#la è il peso per la osservazione corente
#h la costante per il limite per la carta , dipende da (m,
la, ARL)

la<-0.1
h<-17.186
mu0<-rep(0,m)

mewma<-function(dati,la,mu0)
{
M<-matrix(,n-m-1,m)

for (i in 1:(n-m-1))
{
if (i==1)
M[1,]<-la*dati[1,]+(1-la)*mu0
else
M[i,]=la*dati[i,]+(1-la)*M[i-1,]
}
i<-i+1
round(M,3)
}
Mn<-mewma(UC,la,mu0)

#####Mn2#####
#le statistiche Mn2=t(Mn[i,])*(Mn[i,])
#per la carta SSMEWMA

Mn2<-function(Mn){
Mn2<-c()
for (i in 1:nrow(Mn)){
Mn2[i]<-t(Mn[i,])%*%(Mn[i,])
}
Mn2
}
#Mn2(Mn)

#####Limiti della carta#####
#limiti per Mn2: (la/(2-la))*(1-(1-la)^(2*i))*h
# dalle tabelle Hawkins
# h=17.186

```

```

limiti<-function(la,h,Mn){
k<-c()
for (i in 1:nrow(Mn)){
k[i]<-(la/(2-la))*(1-(1-la)^(2*i))*h
}

k
}

#####Grafico#####
wmewma<-Mn2(Mn)
lim<-limiti(la,h,Mn)
s<-max(c(lim,max(wmewma)))
i<-min(c(lim),min(wmewma))
matplot(cbind(wmewma,lim),ylab="Statistica SSMEWMA",
ylim=c(i,s),type="l",lty=1:2)
points(wmewma)
legend
y<-data.frame(lim,wmewma)
Y
#####
ind<-(1:n)[wmewma>lim]
RLm<-ind[1]

##### RL #####
# la RL della carta
ind<-(1:(n-m-1))[wmewma>lim]
RL1<-ind[1]
RL1
RL2<-(n-m-1)-RL1
RL2
RL<-nrow(X2)-RL2
RL

##### Simulazioni RL,ARL #####
#####Simulazioni#####
RLsim<-c()
j<-1
for (j in 1:100){
X1<-mvrnorm(30,mu1,sigma)
X2<-mvrnorm(30,mu2,sigma)
X<-rbind(X1,X2)
dati<-data.frame(X)
n<-nrow(dati)
m<-ncol(dati)
R<-Rf(dati)
U<-Uf(dati)
UC<-U[(m+2):n,]
z<-nrow(UC)
mu0<-rep(0,m)
Mn<-mewma(UC,la,mu0)
wmewma<-Mn2(Mn)

```

```

lim<-limiti(la,h,Mn)
ind<-(1:(n-m-1))[wmewma>lim]
RL1<-ind[1]
RL1
RL2<-(n-m-1)-RL1
RL2
RLsim[j]<- nrow(X2)-RL2
RLsim[j]
j<-j+1
}
RLsim
ARL<-mean(na.omit(RLsim[RLsim>0]))
ARL

#####Simulazioni al variare di IC,
p=2#####

#p=2, shift=1, sigma=I2, ARL=200 (->h<-8.786)

mu1<-c(0,0)
mu2<-c(1,0)
sigma<-Iden(2)
h<-8.786
la<-0.1
Rf<-function(dati){
R<-matrix(,n,m)
R[,1]<-R1(dati)
R[,2]<-R2(dati)
R
}

n1<-c(10,20,30,50,75,100)
k<-length(n1)

RLsim<-c()
ARL<-c()
for (i in 1:k){
for (j in 1:100){
X1<-mvrnorm(n1[i],mu1,sigma)
X2<-mvrnorm(30,mu2,sigma)
X<-rbind(X1,X2)
dati<-data.frame(X)
n<-nrow(dati)
m<-ncol(dati)
R<-Rf(dati)
U<-Uf(dati)
UC<-U[(m+2):n,]
z<-nrow(UC)
mu0<-rep(0,m)
Mn<-mewma(UC,la,mu0)
wmewma<-Mn2(Mn)
lim<-limiti(la,h,Mn)

```

```

ind<- (1:(n-m-1)) [wmewma>lim]
RL1<-ind[1]
RL1
RL2<- (n-m-1) -RL1
RL2
RLsim[j]<- nrow(X2) -RL2
}
ARL[i]<-mean(na.omit(RLsim[RLsim>0]))
}
ARL
ARL= 23.89501 18.85024 16.60368 14.25763 12.70262
12.08891plot(n1,ARL,main="Carta SSMEVMA",type="b")

```

```
##### Confronto con metodo Sullivan #####
```

```

ARLS<-c(23.89501,18.85024,16.60368,14.25763,12.70262,12.08891)
ARLH<-
c(13.235294,11.685714,10.692308,10.015152,10.650794,9.222222 )
n1<-c(10,20,30,50,75,100)
matplot(n1,cbind(ARLS,ARLH),col=c(2,4),lty=c(2,4),type="l",main=
"Con fronto Sullivan - Hawkins",xlab="Periodo IC",ylab="ARL")
legend(50,20,c("Sullivan","Hawkins"),col=c(2,4),lty=c(2,4))

```


Bibliografia

Hawkins, D.M.(1981), “A cusum for a scale parameter”, *Journal of Quality Technology*,13,22-28

Hawkins, D.M.(1987), “Self-Starting CUSUM Charts for Location and Scale”, *The Statistician*,36,299-315

_____ (1993)”Regression Adjustment for Variables in Multivariate Quality Control”, *Journal of Quality Technology*,25,170-182

Hawkins, D.M., Olwell, D.H. (1997),*Cumulative Sum Charts and Charting for Quality Improvement*, New York:Sprinter-Verlag

Hawkins, D.M., Maboudou-Tchao, E.M.(2007),”Self –Starting multivariate Exponentially Weighted Moving Average control charting”,*Technometrics*,49,199-209

Jones, L.A., Champ, C.W., Rigdon, S.E.(2001), “The Performance of Exponentially Weighted Moving Average Control Chart with Estimated Parameters ”,*Technometrics*,43,156-167

Lowry, C.A., Woodall, W.H., Champ, C.W., Rigdon, S.E.,(1992),”A Multivariate Exponentially Weighted Moving Average Control Chart”, *Technometrics*,34,46-53

Prabhu, S.S, Ranger, G.C.,(1997),”Designing a Multivariate EWMA Control Chart”, *Journal of Quality Technology*,29,8-15

Quesenberry, C.P(1991), “SPC Q Charts for Start-Up Processes and Short or Long Runs”, *Journal of Quality Technology*,23,213-224

_____ (1993),”The Effect of Sample Size on Estimated Limits for X and X Control Charts”, *Journal of Quality Technology*,25,237-247

_____ (1995),”On Proprieties of Q Charts for Variables”, *Journal of Quality Technology*,27,184-203

_____ (1997),*SPC Methods for Quality Improvement*,New-York,Wiley

Roberts, S.W.,(1959),”Control Chart Tests Based on Geometric Moving Average”, *Technometrics*,1,239-250

Schaffer, J.R.(1998),”A Multivariate Application of the Q Chart”, presented at the 1998 Joint Statistical Meetings

Sullivan, J.H., Jones, L.A.,(2002), “A Self-Starting Control Chart for Multivariate Individual Observations”,*Technometrics*,44,24-33

Fallace, D.L.,(1959) Bounds on normal approximations to Student’s and the Chi-square distributions, *Annals of Mathematical Statistics*,30,1121-1130

Woodall, W.H.,and Montgomery, D.C.(1999), “Research Issues and Ideas in Statistical Process Control”, *Journal of Quality Technology*,31,376-386