

UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE

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



Tesi di laurea

**Carta di controllo per il coefficiente di variazione e carta
di controllo sintetica per la media e la varianza:
confronto tra i profili dell'ARL**

Relatore:

Prof. Giovanna Capizzi

Laureando:

Vouakouanitou Mpeni Arly Privat

Anno Accademico 2009-2010

A mio Padre **Vouakouanitou Jean-Pierre** e mia Madre
Banzouzi Pauline

INDICE

INTRODUZIONE	7
---------------------------	----------

CAPITOLO 1

IL Controllo statistico di processo	9
1.1: Introduzione	9
1.2: La Qualità.....	11
1.3: Cause di variabilità.....	11
1.4: Carta di Controllo Shewhart	12
1.4.1: Carta di Controllo per Variabili.....	13
1.4.1.1 : Carta di Controllo per la Media.....	14
1.4.1.2 : Carta di Controllo per controllare la variabilità del Processo.....	19

CAPITOLO 2

Carta di controllo per il coefficiente di variazione e carta sintetica per la media e la varianza	23
2.1 : Carta di controllo per il coefficiente di variazione	23
2.1.1: Distribuzione del coefficiente di variazione campionaria	24
2.1.2 : Presentazione della carta di controllo per il cv	27
2.1.3: I limiti di controllo per la carta di controllo per il cv	27

2.2: Carta sintetica per la media e la varianza	31
---	----

CAPITOLO 3

Confronto dell'efficienza della carta CV e della carta sintetica.....	33
--	-----------

3.1 Introduzione	33
------------------------	----

3.2 Piano delle simulazioni.....	34
----------------------------------	----

3.3 Analisi dei risultati.....	36
--------------------------------	----

CAPITOLO 4

Conclusioni.....	42
-------------------------	-----------

Appendice A

Funzioni R usati per le simulazioni dei profili dell'ARL per le due carte	44
--	-----------

A.1 Funzioni usati per il calcolo del profilo dell' ARL per la carta cv.....	44
--	----

A.1.1 arsd	44
------------------	----

A.1.2 RLcv	44
------------------	----

A.1.3 ARLcv	45
-------------------	----

A.1.4 ARLcv_shvar	46
-------------------------	----

A.2 Funzioni usati per il calcolo del profilo dell' ARL per la carta sintetica...	47
---	----

A.2.1 RLMR	47
------------------	----

A.2.2 ARLMR	48
-------------------	----

A.2.3 ARLMR_shvar.....	48
------------------------	----

A.2.4 ARLMR_shmed	49
-------------------------	----

Bibliografia51

INTRODUZIONE

Dalle prime applicazioni del Controllo Statistico di Processo (SPC) è stata riconosciuta l'importanza di sorvegliare, nel caso in cui la caratteristica di qualità osservata sia normalmente distribuita, due parametri del processo, la media e la varianza. Carte di controllo opportunamente disegnate per la sorveglianza di tali parametri, lanciano un allarme in presenza di cause speciali di variazione. Tradizionalmente questo è stato fatto usando separatamente una carta di controllo per la media e una per il range o per la deviazione standard. Recentemente molte sono state le carte proposte per controllare la media e la varianza di un processo usando una singola statistica di controllo (Hawkins and Deng, 2009; Sheu et. al., 2009; Costa e Rahim, 2006).

In tale ambito, Kang et al. (2007) investigano particolari processi, dal settore di produzione materiale all'educazione, alla salute e a vari domini di applicazione sociale, in cui la deviazione standard non è costante bensì è proporzionale alla media. Per questo tipo di processi, comuni ad esempio nell'ambito del controllo di qualità delle analisi cliniche di laboratorio, l'utilizzo di una carta Shewhart per la sorveglianza della media e di una per la sorveglianza della deviazione standard non è adeguato essendo queste basate sull'assunto di media e varianza costante. Nel caso particolare in cui la deviazione standard del processo σ sia direttamente proporzionale alla media, Kang et. al. (2007) suggeriscono dunque il disegno di una carta Shewhart per la sorveglianza del coefficiente di variazione $cv = \sigma / \mu$ (CV), che in tali contesti è invece assunto costante e pari a k , infatti $cv = (\kappa\mu) / \mu = \kappa$

Nell'ipotesi che il coefficiente di variazione sia costante, l'obiettivo in questa tesi è quello di studiare l'efficienza della carta introdotta da Kang et al. (2007), confrontandola non con la carta Shewhart congiunta (\bar{X}, R) ma con una carta singola che sorvegli simultaneamente la media e la varianza del processo. In particolare, useremo l'Average Run Length (ARL) per confrontare l'efficienza della carta Shewhart CV con quella della carta sintetica per la media e la varianza suggerita da Costa e Rahim (2006).

La tesi è divisa in quattro capitoli, nel primo capitolo introdurremo il concetto del controllo statistico di processo (SPC) in generale, parleremo della sua importanza e dell'uso delle carte di controllo per la sorveglianza della media e la varianza. Nel secondo capitolo parleremo specificamente delle due carte di controllo che in particolare saranno studiate: la carta di controllo per il coefficiente di variazione (Kang et al., 2007) e la carta di controllo sintetica per la media e la varianza (Costa and Rahim, 2006). Ciascuna delle due carte verrà definita e ne verrà illustrata l'implementazione. Nel terzo capitolo sarà spiegato il piano delle simulazione dei profili dell'ARL delle due carte di controllo. I risultati, presentati in tabelle saranno usati per confrontare l'efficienza delle due carte.

L'ultimo capitolo, sarà la conclusione della tesi, in cui in base ai risultati delle simulazioni si discuterà l'efficienza delle due carte e le situazioni in cui risultà più adeguato l'uso della carta di controllo per il coefficiente di variazione.

Capitolo 1

Il controllo statistico di processo

1. 1 Introduzione

Il problema di controllo di qualità, ha visto il giorno nel laboratorio della compagnia telefonica BELL negli Stati Uniti con lo statistico Walter Shewhart che nel 1920 introduce lo strumento statistico delle carte di controllo per aiutare a determinare quando un processo di produzione è fuori controllo. La diffusione delle carte di controllo è stata veloce, in meno di 20 anni le carte di controllo sono diventate uno strumento basilare nel controllo statistico del processo di produzione nel mondo industriale ma cominciano ad essere applicate anche in vari altri contesti quali quelli sanitari e sociali.

Il Controllo della Qualità è l'insieme delle procedure con le quali si stabilisce e si consegue uno standard della qualità in grado di soddisfare le esigenze del cliente. Quindi controllo di qualità vuol dire gestione di un processo per assicurare che tutti i suoi parametri rimangano il più possibile entro limiti definiti e per mantenere e migliorare nel tempo le prestazioni di quel processo. Diversi strumenti statistici sono usati per raggiungere tali obiettivi di qui il nome di Controllo Statistico di Processo.

Lo scopo della statistica consiste nel suggerire conclusioni in merito ad un'intera popolazione (o realtà), utilizzando tutte le informazioni contenute in un campione. Nell'ambito specifico del controllo della qualità, le tecniche statistiche sono usate per definire, analizzare e verificare le condizioni che determinano la variabilità della caratteristica di qualità in esame. In sostanza, SPC identifica il comportamento ideale del

processo (macchina) in presenza delle sole cause naturali di variazione; sceglie per tale processo delle caratteristiche di qualità delle quali misura l'evolversi nel tempo, tramite le carte di controllo, uno dei famosi sette strumenti di Deming E.(1982).

Deming (1982) nella sua filosofia per realizzare il miglioramento continuo del processo, ha stabilito il processo di produzione come un ciclo PDCA: *Plan* (progettare,pianificare), *do* (agire,realizzare), *check* (controllare) e *act* (stabilizzare o correggere e riavvio del ciclo di intervento). le sette strumenti di Deming (1982) che sono :

- il foglio raccolta dati
- l'istogramma
- il diagramma causa-effetto
- il diagramma di Pareto
- l'analisi per stratificazione
- l'analisi di correlazione
- la carta di controllo

Quelle elencate sono tecniche statistiche usati per controllare ogni tappa del ciclo produttivo, servono ad individuare le variabili o le cause più importanti ad, analizzarne nel modo più obiettivo il comportamento ed a prendere decisioni razionali.

2 La Qualità

In modo breve si può definire la qualità dicendo che è una o più caratteristiche che il prodotto o servizio devono possedere. Ogni organizzazione ha le proprie esigenze di qualità, cioè delle normative ben definite rispetto alle quali un prodotto o servizio di buona qualità deve essere conforme, e che sono specifiche per ogni organizzazione.

La qualità è diventata uno dei fattori più importanti nel processo di decisione del consumatore, quando si tratta di scegliere tra beni o servizi in concorrenza tra di loro. Di conseguenza comprendere, controllare e migliorare la qualità è un elemento decisivo :

- per raggiungere il successo nell'impresa produttiva, per la crescita e il miglioramento della propria posizione nel mercato,
- per prendere decisioni giuste e ragionevoli in caso degli esami in laboratorio clinica.

La qualità è misurata dalla variabilità della produzione o del processo produttivo, più il processo è variabile più è di scarsa qualità, quindi l'obiettivo principale del SPC è quello di migliorare la qualità riducendo la variabilità del processo.

1.3 Cause di variabilità

In SPC si distinguono due cause di variabilità del processo :

- **Cause Speciali** : sono cause assegnabili, non prevedibili, possono essere attribuite a un mal funzionamento dell'apparecchio, all'usura delle macchine, all'inesperienza dell'impiegato. Quindi è possibile risalire al motivo che ha provocato la variazione.

- **Cause Comuni** : sono cause casuali, o normali. Sono piccole variazioni, prevedibili, e inevitabili.

Un processo, soggetto una variazione dovuta a cause speciali è detto fuori controllo, ha un comportamento statisticamente non prevedibile,

Gli strumenti dell'SPC permettono di identificare la presenza di cause speciali. Una volta eliminate, la variabilità del processo dovuta solo a cause comuni presenta una distribuzione prevedibile è quindi controllabile. In questo caso il processo è detto in controllo.

1. 4 La Carta di Controllo Shewhart

La carta di controllo è uno strumento usato nel SPC, per la sorveglianza di una caratteristica di qualità X_t . La caratteristica della qualità, in controllo, presenta una certa distribuzione di probabilità. In questa tesi si assume che le osservazioni sulle caratteristiche di qualità siano indipendenti e identicamente distribuite con una distribuzione normale.

Sia $W_t = g(x_t)$ una opportuna funzione statistica delle osservazioni fatte sulla caratteristica di qualità. Il disegno di una carta Shewhart consiste nel calcolo di:

- una linea centrale (LC), livello ideale della caratteristica pari al valore atteso della caratteristica della qualità nell'ipotesi che il processo sia in controllo.

$$LC = \bar{W} = \frac{\sum_{t=1}^n w_t}{n}$$

- Un limite inferiore (LCL) e superiore (UCL) calcolato in base alla distribuzione in controllo della caratteristica di qualità .

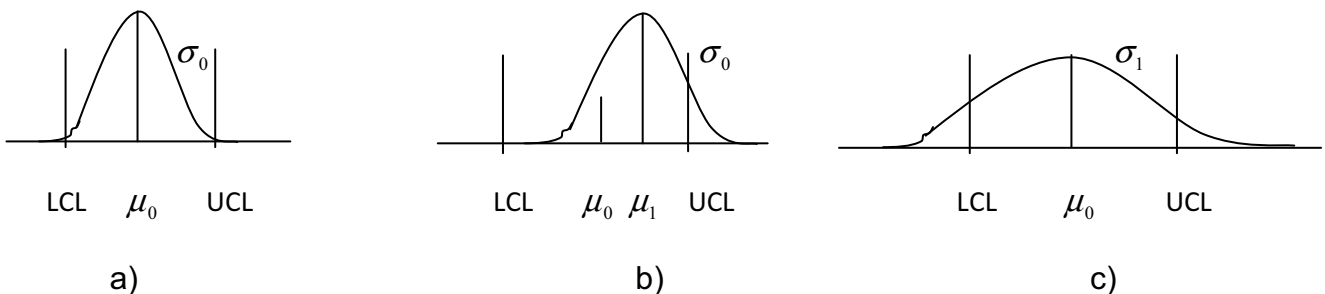
$$LCL = LC - L \sigma_{\bar{w}} ; \quad UCL = LC + L \sigma_{\bar{w}}$$

La costante L è il quantile di probabilità scelto in modo di avere una alta probabilità in controllo, di solito pari al 99%, che le osservazioni della caratterista di qualità siano dentro i limiti; mente $\sigma_{\bar{w}}$ è lo standard error della statistica di controllo.

1. 4. 1 Carta di Controllo per Varabili

Si definisce “variabile” una caratteristica della qualità misurabile numericamente come peso, dimensione, volume, ecc. In questo contesto di SPC, è necessario poter controllare la posizione e la variabilità del processo tramite carte di controllo specifiche.

Le figure seguente mostrano la distribuzione della caratteristica di qualità quando il processo è in controllo, quando il processo è fuori controllo a causa di un cambiamento avvenuto nella media, o nella variabilità del processo.



In particolare, i grafici descrivono le seguenti situazioni

- a) la media μ_0 e la deviazione standard σ_0 del processo sono al livello nominale (in controllo)
- b) la media del processo $\mu_1 > \mu_0$.
- c) la deviazione standard del processo $\sigma_1 > \sigma_0$.

1. 4.1.1 Carta di Controllo per la Media

Nella carta di controllo per la media si prende come caratteristica da controllare la media del processo. Supponendo il processo in controllo (presenza solo di una variabilità dovuta a cause comuni), con una distribuzione normale della caratteristica di qualità, ovvero $X \sim N(\mu, \sigma^2)$. La statistica di controllo \bar{W} = la media campionaria si distribuisce come una

$$N\left(\mu, \frac{\sigma^2}{n}\right).$$

Il disegno della carta di controllo per la media varia secondo lo stato dei parametri della distribuzione della variabile, se μ e σ del processo sono conosciuti o incogniti.

a) Caso in cui μ e σ sono noti

Quando μ e σ sono noti, la linea centrale e i limiti di controllo per la carta Shewhart per la media sono:

$$LCL = \mu - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad : \text{ il limite di controllo inferiore}$$

$$LC = \mu \quad : \text{ la linea centrale}$$

$$UCL = \mu + Z_{1 - \frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad : \text{ Il limite di controllo superiore}$$

dove $Z_{\frac{\alpha}{2}}$ è il quantile della normale che lascia a sinistra una probabilità pari a $\frac{\alpha}{2}$, e α

usualmente scelto in modo di avere 99% la probabilità che le osservazioni cadono dentro i limiti di controllo. Convenzionalmente si sceglie $Z_{\frac{\alpha}{2}} = Z_{1 - \frac{\alpha}{2}} \cong 3$ secondo la regola

chiamata del 3-sigma.

Nel caso in cui uno o entrambi i parametri del processo siano incogniti i limiti di controllo vengono stimati a partire da m sottogruppi razionali o campioni di n osservazioni ($2 \leq n \leq 10$), misurate nello stesso intervallo temporale e con gli stessi strumenti e rappresentative del processo in controllo.

b) Caso in cui solo σ è conosciuto, μ è incognita

Avendo nessuna informazione su μ , questa viene stimata facendo la media delle medie aritmetiche \bar{x}_i , $i = 1, \dots, m$, calcolate sugli m campioni Media delle medie aritmetiche:

$$\bar{x} = \frac{\sum_{i=1}^m \bar{x}_i}{m}$$

Quando il processo è in controllo la media campionaria si distribuisce come una normale

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Lo standard errore della media campionaria : $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$

I limiti di controllo a 3-sigma in questo caso sono :

$$LCL = \bar{x} - 3\sigma_{\bar{x}} = \bar{x} - 3\frac{\sigma}{\sqrt{n}}$$

$$LC = \bar{x}$$

$$UCL = \bar{x} + 3\sigma_{\bar{x}} = \bar{x} + 3\frac{\sigma}{\sqrt{n}}$$

c) **Caso in cui μ e σ sono incogniti**

Assumendo che n sia costante per ciascun campione. Se la varianza del processo è incognita allora uno stimatore corretto della varianza è la varianza campionaria

$$s^2 = \frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n-1}$$

La deviazione standard S non è uno stimatore corretto di σ , se la distribuzione sottostante è normale allora S è la stima di $c_4\sigma$. Due sono le stime della varianza in questo caso, e i limiti di controllo differiscono secondo la stima della varianza scelta

.c.1) stima della varianza basata sulla varianza campionaria (n > 2)

Sia S la deviazione standard, il suo valore atteso e la sua varianza sono pari a

$$E(s) = c_4 \sigma_x \text{ e } V(s) = \sigma_x (1 - c_4^2)^2, \text{ rispettivamente}$$

Uno stimatore corretto di σ_x è dato da dunque $\hat{\sigma} = \frac{E[\hat{s}]}{c_4} = \frac{\bar{s}}{c_4}$, dove $\bar{s} = \frac{\sum_{i=1}^m s_i}{m}$

con $S_i = \sqrt{\frac{\sum_{j=1}^n (x_{ij} - \bar{x})^2}{n-1}}$. In questo caso i limiti di controllo a 3-sigma diventano :

$$LCL = \bar{x} - 3 \hat{\sigma}_{\bar{x}} = \bar{x} - 3 \frac{\bar{s}}{c_4 \sqrt{n}}$$

$$LC = \bar{x}$$

$$UCL = \bar{x} + 3 \hat{\sigma}_{\bar{x}} = \bar{x} + 3 \frac{\bar{s}}{c_4 \sqrt{n}}$$

c.2) Stima della varianza basata sul range campionario ($2 < n < 10$)

Sia $R = X_{\max} - X_{\min}$ la statistica range il suo valore atteso e la sua varianza sono pari a

$E(R) = d_2 \sigma_x$ e $V(R) = d_3^2 \sigma_x^2 = d_3^2 \frac{E(R)^2}{d_2^2}$, rispettivamente. Uno stimatore corretto di

σ_x è $\hat{\sigma}_x = \frac{E[R]}{d_2} = \frac{\bar{R}}{d_2}$, con $\bar{R} = \frac{\sum_{j=1}^m R_j}{m}$. I Limiti di controllo a 3-sigma in questo

caso sono :

$$LCL = \bar{x} - 3 \hat{\sigma}_x = \bar{x} - 3 \frac{\bar{R}}{d_2 \sqrt{n}}$$

$$LC = \bar{x}$$

$$UCL = \bar{x} + 3 \hat{\sigma}_x = \bar{x} + 3 \frac{\bar{R}}{d_2 \sqrt{n}}$$

Le costanti d_2 , d_3 e c_4 sono tabulate in funzione di n , dimensione dei sottogruppi razionali (vedi Tabella 1)

Tabella 1 Costanti per il calcolo dei limiti della carta Shewhart

n	d_2	d_3	c_4
2	1,128	0,853	0,80
3	1,693	0,888	0,886
4	2,059	0,880	0,921
5	2,326	0,864	0,940
6	2,534	0,848	0,952
7	2,70	0,833	0,959
8	2,847	0,820	0,965
9	2,970	0,808	0,969
10	3,078	0,80	0,973

L'efficienza nell'usare il range per stimare la varianza decresce rapidamente all'aumentare della dimensione n . In particolare, la carta basata su s è più efficiente della carta basata su R quando $n > 10$

1.4.1.2 La carta di controllo per controllare la variabilità del processo

Controllare la sola media non è sufficiente, è molto importante avere anche sotto controllo la variabilità del processo. Due tipi di carta Shewhart sono usate per controllare la

variabilità del processo : la carta \bar{R} , basata sul range campionario R , e la carta S , basta sullo scarto tipo s (da preferire quando $n > 10$).

a) Carta \bar{R}

Per $n < 10$, la carta Shewhart basata sui range campionari viene usata per controllare la variabilità del processo. La statistica di controllo $W = R$ il range campionaria ha distribuzione $W \cong N(\mu_R, \sigma_R^2)$, con

$$\mu_R = E(R) = \bar{R} = d_2\sigma \quad \text{e} \quad \sigma_R = \sqrt{V(R)} = d_3\sigma$$

I limiti di controllo per la carta Shewhart che sorveglia la variabilità del processo, a seconda che la varianza in controllo sia nota o incognita, sono dati da

a.1) varianza nota

$$LCL = \bar{R} - 3d_3\sigma = d_2\sigma - 3d_3\sigma = (d_2 - 3d_3)\sigma$$

$$LC = \bar{R}$$

$$UCL = \bar{R} + 3d_3\sigma = d_2\sigma + 3d_3\sigma = (d_2 + 3d_3)\sigma$$

a.2) varianza incognita

$$\hat{E}(R) \pm 3 \hat{\sigma}_R = d_2 \hat{\sigma} \pm 3d_3 \hat{\sigma} \quad \text{con} \quad \hat{\sigma} = \frac{\bar{R}}{d_2}$$

$$LCL = \bar{R} - 3 \frac{\bar{R}}{d_2} d_3$$

$$LC = \bar{R}$$

$$UCL = \bar{R} + 3 \frac{\bar{R}}{d_2} d_3$$

b) Carta S

All'aumentare di n l'efficienza della carta basata sul range campionaria diminuisce, generalmente per $n > 10$ è preferibile per controllare la variabilità del processo usare

la carta S . La statistica di controllo $W = s$, sotto l'ipotesi della normalità della variabile di studio $X \sim N(\mu, \sigma^2)$, $W \cong N(\mu_s, \sigma_s^2)$, con $\mu_s = E(s) = c_4 \sigma$; $\sigma_s =$

$$\sqrt{V(s)} = \sqrt{(1-c_4^2)\sigma^2} = \sigma \sqrt{(1-c_4^2)}$$

I limiti di controllo per la carta Shewhart che sorveglia la variabilità del processo, a seconda che la varianza in controllo sia nota o incognita, sono dati da

b.1) Varianza nota

$$LCL = c_4 \sigma - 3 \sigma \sqrt{(1-c_4^2)}$$

$$LC = c_4 \sigma$$

$$UCL = c_4 \sigma + 3 \sigma \sqrt{(1 - c_4^2)}$$

b.2) Varianza incognita

In questo caso la varianza viene stimata con $\hat{\sigma} = \frac{\bar{S}}{c_4}$ e i limiti di controllo

$$LCL = \bar{S} - 3 \frac{\bar{S}}{c_4} \sqrt{1 - c_4^2}$$

$$LC = \bar{S}$$

$$UCL = \bar{S} + 3 \frac{\bar{S}}{c_4} \sqrt{1 - c_4^2}$$

Capitolo 2

Carta di controllo per il coefficiente di variazione e carta sintetica per la media e la varianza

2.1 Carta di controllo per il coefficiente di variazione

Assumiamo che la variabile di studio X sia distribuita normalmente $X \sim N(\mu, \sigma^2)$. La carta \bar{x} e la carta R o la carta S sono adeguatamente usate quando il processo sotto controllo presenta rispettivamente una media costante o una varianza costante.

Ci sono processi in controllo però che non presentano una media costante in questi casi il Controllo Statistico di Processo si riduce al monitoraggio della sola variabilità.

Un problema comune nei laboratori clinici consiste nel ripetere le misure di una caratteristica mantenendo la qualità della misurazione. Un esempio è quello del controllo del livello di un farmaco nel sangue dei pazienti. Essendo la media diversa per ogni paziente, sarà rilevante il monitoraggio solo della variabilità tra le misure ripetute.

In casi come questi la varianza è funzione della media e quindi non è stabile. Tuttavia esiste una relazione di proporzionalità tra la varianza e la media, $\sigma = \kappa\mu$, che rende costante il coefficiente di variazione, ovvero $cv = \sigma / \mu = \kappa$.

In questi contesti applicativi può essere conveniente sorvegliare la variabilità del processo prendendo sotto gruppi razionali, calcolando la media e la deviazione standard ed implementando una carta di controllo sul coefficiente di variazione campionaria $W = cv$.

2.1.1 Distribuzione del coefficiente di variazione campionaria W

Hendricks and Robey (1936), basandosi sulla distribuzione combinata della media e della deviazione standard di un campione di n elementi di una normale $N(\mu, \sigma^2)$ e operando con un cambiamento di variabile $v = \arctan(w)$, trovano la distribuzione del coefficiente di variazione campionaria W ,

$$f(v) = k(v) \int_{r=0}^{\infty} r^{n-1} \exp(-\{ar - b(v)\}) dr,$$

dove $a = \frac{\sqrt{n}}{(\sigma\sqrt{2})}$, $b(v) = a\mu \cdot \cos(v)$ e

$$k(v) = n^{n/2} \frac{\sin^{n-2}(v) \exp(-n\mu^2) \frac{\sin^2(v)}{2\sigma^2}}{2^{n/2-1} \sigma^n \sqrt{n} \Gamma\left\{\frac{[n-1]}{2}\right\}}$$

Un approccio alternativo è stato suggerito da Iglewicz(1967) e Verrill(2003), che trovano che la trasformazione

$$T = \sqrt{n} \frac{\bar{X}}{S} = \sqrt{n} W^{-1}$$

segue una distribuzione t non centrale con $(n-1)$ gradi di libertà con parametro di non centralità \sqrt{n}/κ . Quindi la funzione di distribuzione cumulata del coefficiente di variazione può essere calcolata a partire da una distribuzione non centrale t . Le formule esplicite di questa funzione di distribuzione cumulata sono date in Abramowitz and Stegun (1972).

Questa distribuzione è supportata da molti software statistici. Tuttavia nell'ambito comune delle carte di controllo, l'argomento e il parametro di non centralità sono grandi. Sotto questi condizioni le formule date da Abramowitz and Stegun (1972), implementate con molti algoritmi, come l'AS 243 (Lenth, 1989), non possono essere ottenute. R è ad esempio inaffidabile per i parametri di non centralità superiore a 37.62. SAS d'altra parte usa apparentemente un algoritmo diverso che può supportare anche parametri di non centralità grandi. Per queste ragioni è rilevante avere una forma esplicita per la funzione di densità e la funzione di distribuzione cumulativa di W .

Kang et. al. (2007) in particolare hanno sviluppato una forma canonica per la distribuzione del coefficiente di variazione nel modo seguente. Sia $Y \sim N(\kappa^{-1}, 1)$ e indipendentemente

$fV^2 \sim \chi_f^2$. Si definisce

$$U = \frac{V}{Y},$$

La funzione di densità di U è data esplicitamente come

$$f_1(u) = \frac{A(u)}{(1+u^2)^{(f+1)/2}} I_f \left(\frac{1}{\kappa(1+u^2)^{0.5}} \right) \quad \text{if } u \geq 0$$

$$f_2(u) = \frac{(-1)^{f-1} A(u)}{(1+u^2)^{(f+1)/2}} I_f \left(\frac{1}{\kappa(1+u^2)^{0.5}} \right) \quad \text{if } u < 0$$

dove :

$$A(u) = \frac{u^{f+1} \exp\left(-\frac{u^2}{2\kappa^2(1+u^2)}\right)}{\sqrt{2\pi} \sqrt{f} \Gamma\left(\frac{f}{2}\right) 2^{(f-2)/2}},$$

e

$$I_f(b) = \int_0^\infty q^f \exp\left[-\frac{1}{2}(q-b)^2\right] dq$$

è una funzione Airy (vedi par. 10.4 in Abramowitz e Stegun, 1972 e Iglewicz (1967)).

Si denoti con $C(cv, f)$ questa forma canonica e si ipotizzi di disporre di un campione di n osservazioni provenienti da una distribuzione normale con parametri noti, media μ e standard deviazione σ proporzionale alla media: $\sigma = \kappa\mu$. La media campionaria \bar{X} e la deviazione standard S si distribuiscono come una $N(\mu, (\kappa\mu)^2/n)$ e come $\kappa\mu\chi/\sqrt{f}$, rispettivamente.

Il coefficiente di variazione campionaria è $W = \frac{S}{\bar{X}}$. Scrivendo $V = S/(\kappa\mu)$ e $Y =$

$\sqrt{n}(\bar{X}/\kappa\mu)$ ricaviamo $W = \sqrt{n}(V/Y)$ dove $V \sim \kappa\mu\chi_{n-1}/\sqrt{n-1}$, e $Y \sim N(\sqrt{n}/\kappa, 1)$.

Quindi, il coefficiente di variazione campionaria W è \sqrt{n} volte una variabile canonica $C(\kappa/\sqrt{n}, n-1)$.

A parte la connessione esatta con la non centrale distribuzione t , ci sono numerose approssimazioni della distribuzione canonica del cv . McKay (1932) ha approssimato la distribuzione del cv con una chi quadrato con $(n-1)$ gradi di libertà. Con una grande numerosità campionaria, $n \geq 30$, i risultati dell'approssimazione di McKay (1932) sono

simili a quelli esatti e sono facile da calcolare usando i comuni software di statistica, ma il problema è che gruppi razionali di queste dimensioni sono raramente disponibili nelle applicazioni pratiche e quindi questa possibilità è di poco rilievo nel contesto considerato.

D'altro parte, Hendricks e Robey (1936) forniscono approssimazioni soddisfacenti per piccoli valori del CV , ma non per grandi.

2.1.2 Presentazione della carta di controllo per il CV

La costruzione della carta di controllo per il CV è molto simile a quelle delle carte Shewhart tradizionali, quali la carta \bar{x} e la carta R o la carta S .

Sempre considerando una variabile X distribuita normalmente, la statistica di controllo W consiste nel coefficiente di variazione campionaria. Due sono le assunzioni da verificare per potere usare la carta di controllo per il CV : che i campioni siano indipendenti e che la deviazione standard sia proporzionale alla media.

2.1.3 I limiti di controllo per la carta di controllo per il CV

Per stabilire i limiti della carta di controllo per il CV , l'approccio standard 3-sigma non può essere applicato visto che tutti i momenti del coefficiente di variazione campionaria sono infiniti. Infatti si consideri la forma canonica seguente:

$$E(U^k) = E(V^k) \cdot E(Y^{-k}),$$

con $Y \sim N(\mu, \sigma^2)$. I momenti $\kappa(\mu_y^k)$ della distribuzione si ricavano dalla formula

$$u_y^k = \int_a^b (y - y_0) f(y) d_y \quad \text{con} \quad f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \quad \text{oppure}$$

$$\mu_y^k = E(y^k) = E\{[y - E(y)]^k\} \rightarrow E(y^{-k}) = \frac{1}{E(y^k)}$$

Vista la simmetria della funzione di distribuzione della normale, tutti i suoi momenti di ordine dispari ($k = 2n + 1$) rispetto alla media ($E(y)$) sono nulli e di conseguenza tutti i

momenti negativi sono infiniti, infatti $E(y^{-k}) = \frac{1}{E(y^k)} = \frac{1}{0} = \infty$.

Kang et. al. (2007) suggeriscono, per il disegno della carta Shewhart per la sorveglianza del coefficiente di variazione, un approccio basato sui limiti di probabilità invece che uno basato sulla regola del 3σ .

In particolare la linea centrale (**LC**) della carta di controllo per il CV , viene posta uguale al valore del coefficiente di variazione in controllo, a differenza di quanto fatto per le tradizionali carte Shewhart in cui come linea centrale si userebbe il valore atteso del CV in controllo, ovvero la media dei coefficienti di variazione calcolati su campioni in controllo di ampiezza n . Nel caso della carta di controllo per il CV questo approccio non sarebbe però adeguato, infatti Connett and Lee (1990), che hanno investigato l'utilizzo della media dei CV da usare come linea centrale della carta di controllo, mostrano inefficienza statistica dell'uso della media dei CV .

Kang et. al. (2007), per un range di valori di n e di κ forniscono i limiti di controllo, inferiore e superiore, aventi ciascuno una probabilità pari a $\alpha/2 = 0.00135$, di essere superati quando il processo in controllo. Tali limiti sono riportati nella Tabella 2

Tabella 2: Limiti di probabilità carta **CV** per una selezione di K per differenti n

K	$n = 5$		$n = 10$		$n = 15$		$n = 20$	
	LCL	UCL	LCL	UCL	LCL	UCL	LCL	UCL
0.02	0.00325	0.04222	0.00743	0.03471	0.00957	0.03174	0.01090	0.03005
0.03	0.00488	0.06337	0.01114	0.05210	0.01435	0.04763	0.01634	0.04509
0.04	0.00650	0.08458	0.01485	0.06951	0.01913	0.06355	0.02178	0.06014
0.05	0.00812	0.10587	0.01855	0.08696	0.02391	0.07949	0.02722	0.07522
0.06	0.00975	0.12725	0.02226	0.10446	0.02868	0.09546	0.03266	0.09033
0.07	0.01137	0.14875	0.02595	0.12202	0.03344	0.11148	0.03809	0.10547
0.08	0.01298	0.17038	0.02965	0.13966	0.03820	0.12755	0.04351	0.12065
0.09	0.01460	0.19217	0.03333	0.15737	0.04296	0.14367	0.04892	0.13587
0.10	0.01621	0.21414	0.03701	0.17518	0.04770	0.15986	0.05432	0.15115
0.12	0.01943	0.25869	0.04435	0.21111	0.05716	0.19247	0.06510	0.18188
0.14	0.02264	0.30423	0.05165	0.24754	0.06657	0.22542	0.07583	0.21288
0.16	0.02583	0.35097	0.05892	0.28457	0.07594	0.25879	0.08651	0.24421
0.18	0.02900	0.39912	0.06613	0.32229	0.08525	0.29263	0.09713	0.27591
0.20	0.03215	0.44894	0.07330	0.36081	0.09450	0.32702	0.10768	0.30804
0.22	0.03529	0.50070	0.08042	0.40025	0.10368	0.36203	0.11816	0.34064
0.24	0.03840	0.55474	0.08747	0.44071	0.11279	0.39773	0.12856	0.37377
0.26	0.04148	0.61142	0.09447	0.48233	0.12183	0.43420	0.13888	0.40749
0.28	0.04455	0.67117	0.10140	0.52526	0.13078	0.47152	0.14911	0.44186
0.30	0.04758	0.73449	0.10826	0.56964	0.13965	0.50978	0.15925	0.47694

In pratica il disegno e l'implementazione della carta Shewhart consistono nella raccolta di un numero sufficientemente grande di osservazioni di Phase I da utilizzare per la verifica dell'attendibilità di un modello con coefficiente di variazione costante e per la stima del coefficiente di variazione in controllo (**LC**).

In particolare si estraggono dal processo, assunto in controllo m campioni di Phase I, ciascuno di dimensione n , non necessariamente costante. Sull' i -esimo campione si calcolano dunque \bar{X}_i , S_i e $W_i = \bar{X}_i / S_i$, ovvero la media, la deviazione standard e il coefficiente di variazione campionario. Tra le varie possibilità per combinare le stime ottenute sugli m campioni di Phase I, e dunque ottenere la linea centrale (**LC**), gli autori suggeriscono la media, la radice quadrata della media ponderata dei coefficienti di variazione campionari e lo stimatore di massima verosimiglianza. Di tali stime, indagate in termini di efficienza da Connett e Lee (1990), gli autori suggeriscono l'utilizzo della radice quadrata della media ponderata dei coefficienti di variazione al quadrato, ovvero

$$\hat{K} = \sqrt{\frac{\sum_i (n_i - 1) W_i^2}{\sum_i (n_i - 1)}}$$

L'altra condizione per potere usare la carta di controllo per il coefficiente di variazione, è di verificare la costanza del coefficiente di variazione nel processo, confrontando se esattamente la deviazione standard è proporzionale alla media oppure se varia in altri modi con questa. A tal fine si può usare un diagramma di dispersione dei coefficienti di variazione dei sottogruppi razionali verso la media corrispondente, possibilmente aggiungendo una retta o una curva per verificare l'andamento, o un formale test di regressione per verificare l'ipotesi che il coefficiente di variazione non dipenda dalla media.

Per una probabilità di eccedere i limiti di controllo in uno stato di controllo complessivamente pari a 0.0027, equivalente a una Average Run Length (ARL) pari a 370, Kang et. al. (2007) forniscono i valori dell'ARL fuori controllo, per alcune dimensioni campionarie, nell'ipotesi che il coefficiente di variazione subisca aumenti dello 0.25%, 50% e 100%. I risultati ottenuti dagli autori suggeriscono che la sensitività della carta Shewhart basata sul coefficiente di variazione aumenta all'aumentare della dimensione del cambiamento e risulta piuttosto stabile al variare del coefficiente di variazione.

2.2 Carta sintetica per la media e la varianza

Come sottolineato in precedenza non è sufficiente sorvegliare solo la media o solo la deviazione standard (o il loro range). Il controllo simultaneo della media e la varianza è sempre stato un problema molto rilevante nel SPC . I due parametri μ e σ , che danno informazioni sul processo, sono legati l'uno all'altro, un cambiamento in varianza ha un influenza notevole sulla carta \bar{X} , sebbene il viceversa non è vero. Un buon processo ha una varianza piccola e una media più o meno stabile. Quindi mantenere sotto controllo la media e la varianza del processo simultaneamente è un importante obiettivo nell'ambito del controllo statistico del processo.

Al fine di sorvegliare congiuntamente i due parametri possono utilizzarsi due carte separate o una singola carta e di conseguenza una singola statistica di controllo. Le carte Shewhart \bar{X} e R sono carte specifiche per individuare rispettivamente un cambiamento avvenuto nella media o nella varianza del processo. Questi carte prese separatamente ci danno solo informazioni sul comportamento della media per la carta \bar{X} e sul

comportamento della varianza del processo per la carta R senza nessuna connessione tra i due .

Per una maggiore efficienza in termini di risorse, il controllo simultaneo dei due parametri è una delle soluzioni, ad esempio in quei casi in cui sia importante identificare un aumento della variabilità del processo prima che questo influenzi la media.

La carta sintetica per la media e la varianza, discussa da Costa e Rahim (2006), combina la carta \bar{x} e la carta R o la carta S , eseguendo un controllo simultaneo della media e della varianza, e lanciando un allarme la prima volta che una delle due carte segnala la presenza di cause speciali.

Sugli m campioni indipendenti di n osservazioni ($2 \leq n \leq 10$) si calcolano le due statistiche di controllo $W_1 = \bar{x}$, media campionaria e $W_2 = R$, range campionario. I limiti di controllo e la linea centrale per ciascuno parametro sono calcolati come indicato nel capitolo 1, usando la distribuzione di ciascuna delle statistiche di controllo.

$$\bar{x} \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}) \text{ con } \mu_{\bar{x}} = \mu ; \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$R \sim N(\mu_R, \sigma_R) \text{ con } \mu_R = d_2\sigma ; \quad \sigma_R = d_3\sigma$$

Capitolo 3

Confronto dell'efficienza della carta CV e della carta sintetica

3.1 Introduzione

La carta di controllo per il coefficiente di variazione è disegnata sostanzialmente per la sorveglianza della sola variabilità del processo attraverso l'uso di un'unica statistica di controllo: il coefficiente di variazione, pari al rapporto tra la deviazione standard e la media del processo legate da una relazione di proporzionalità. La carta di controllo sintetica controlla invece sia la media sia la varianza del processo usando due statistiche di controllo, la media e il range.

Pur mostrando l'efficienza della carta di controllo Shewhart per il coefficiente di variazione, Kang et al. (2007) non effettuano alcuna comparazione della performance della carta da loro proposta con carte simili suggerite in letteratura. In questo elaborato si vuole indagare in quali contesti la carta Shewhart per il coefficiente di variazione risulta più efficiente di una carta disegnata per sorvegliare simultaneamente la media e la deviazione standard di osservazioni normalmente distribuite, quale appunto la carta sintetica proposta da Costa e Rahim (2006).

Nel seguito vengono presentati i risultati dello studio di simulazione svolto per confrontare l'efficienza delle due carte, in termini *Average Run Length* (ARL). L'ARL è la media della variabile casuale *Run Length* che indica la prima volta che la statistica di controllo supera i limiti di controllo. L'ARL dà informazione su quanto veloce sia una carta di controllo nel segnalare una variazione nel processo quando questo si è verificato realmente (ARL fuori controllo) e su quanto sia lenta a sbagliarsi nel segnalare una variazione nel processo che

non è mai avvenuta (ARL in controllo). Una carta di controllo è più efficiente di un'altra quando segnala più velocemente un cambiamento realmente avvenuto nel processo.

3. 2 Piano delle simulazioni

Per ottenere un confronto valido, nel seguito calcoleremo, via simulazione e sotto condizioni identiche: indipendenza, normalità, relazione di proporzionalità tra la deviazione standard e la media, e uguaglianza dell'ARL in controllo.

In particolare per ottenere un valore dell'ARL in controllo pari a 370, i limiti di probabilità utilizzati per la carta di controllo CV sono i seguenti :

$n = 5$		
CV	LCL	UCL
0,02	0,00325	0,04222
0,05	0,00812	0,10587
0,1	0,01621	0,21414
0,16	0,02583	0,35097
0,20	0,03215	0,44894
0,26	0,04148	0,61142
0,30	0,04758	0,73449

Mentre quelli della carta sintetica sono $\mu_0 \pm 3.190\sigma_0/\sqrt{5}$ e $5.397\sigma_0$ per la carta \bar{x} e R , rispettivamente.

Per tener conto della condizione di proporzionalità tra la deviazione standard σ e la media μ , tale che $\sigma = k\mu$ ottenendo $cv = (k\mu)/\mu = k$, le osservazioni usate per il calcolo dell'ARL delle due controllo sono state generate da una distribuzione normale i cui parametri μ e σ in controllo diano luogo al coefficiente di variazione costante.

Si osservi che, per ottenere una stima affidabile del cv e per potere controllare la sua costanza, è consigliato di usare un grande numero campioni m per la Phase 1 rispetto alla dozzina di sottogruppi razionali pensati finora adeguata. Nelle nostre simulazioni, per il calcolo della Run Length sono dunque stati simulati $m = 100000$ campioni di ampiezza $n = 5$.

Dunque, sono stati calcolati e confrontati per le due carte i profili dell'ARL, prima nel caso in cui solo la varianza del processo cambia da $N(\mu, \sigma^2)$ a $N(\mu, [\lambda\sigma]^2)$ e dopo, quando sia la varianza e la media del processo si assume cambino da $N(\mu, \sigma^2)$ a $N(\mu + \delta\sigma, [\lambda\sigma]^2)$. In questo contesto il processo sorvegliato si considera in controllo per $\delta = 0$ e $\lambda = 1$.

Avremo potuto esaminare anche la situazione in cui solo la media del processo cambia da $N(\mu, \sigma^2)$ a $N(\mu + \delta\sigma, \sigma^2)$, come fatto per la varianza ma non avrebbe avuto rilevanza in questo contesto poiché l'assunzione della costanza del coefficiente di variazione implica l'indipendenza di questo dalla media quindi un qualsiasi shift della media non ha effetto sull'ARL della carta di controllo sul coefficiente di variazione.

3. 3 Analisi dei risultati

Nella tabella 3.1, a parità di numerosità campionaria ($n=5$) e del coefficiente di variazione, si riportano i profili dell'ARL per le due carte sotto l'ipotesi in cui solo la varianza del processo cambia. Due aspetti sono rilevanti nei risultati ottenuti.

Il primo è che l'ARL della carta CV qualunque sia il valore del coefficiente di variazione oppure del coefficiente di proporzionalità K tra la deviazione standard e la media, risulta sempre più alto per tutti gli shift λ considerati rispetto all' ARL della carta sintetica. Infatti, tranne che nella situazione in controllo, $\lambda = 1$, l'ARL della carta per il coefficiente di variazione, segnala un aumento della variabilità del processo con un ritardo rispetto alla carta sintetica. Questo ritardo non è trascurabile, se guardiamo ad esempio la seconda riga della tabella 3.1, quindi quando la varianza del processo aumenta di 0.05 ($\lambda = 1.05$), vediamo che la carta per il coefficiente di variazione richiede in media 50 campioni di più rispetto alla carta sintetica. Questo risultato è meglio evidenziato nella Tabella 3.2, che riporta la media delle differenze tra l'ARL della carta CV e quella della carta sintetica, al variare dei cambiamenti avvenuti nella varianza del processo (ovvero al variare di λ). I risultati mostrano che tale differenza benché decrescente, è grande per piccoli valori di λ e piccola per grande valori di λ .

Il secondo risultato mostra che l'ARL per la carta CV, quando confrontata con l'ARL della carta sintetica, peggiora mano a mano che il coefficiente di variazione aumenta. Guardando la Tabella 3.1 si nota infatti come, prendendo in considerazione un qualsiasi valore di $\lambda > 1$, l'ARL della carta CV aumenta con il coefficiente di variazione. Ad esempio, per la carta Shewhart applicata al coefficiente di variazione e per $\lambda = 1.5$, abbiamo un'

ARL = 10.45 per $CV = 0.05$, un'ARL= 10.84, per $CV=0.1$ e così via fino ad arrivare un'ARL= 13.19 per $CV=0.3$.

Per evidenziare meglio questo aspetto la Tabella 3.2 riporta, per ogni valore del coefficiente di variazione, la differenza tra l'ARL della carta CV e quella della carta sintetica al variare di λ partendo da $\lambda = 1,05$. Si osserva che all'aumentare del coefficiente di variazione e quindi della costanza di proporzionalità κ la differenza tra i due profili diventa mano a mano più grande per un dato valore di λ .

Nella Tabella 3.3 infine sono riportati i profili per le due carte a confronto nel caso in cui λ e δ cambiano. Si vede che, come detto in precedenza, al variare di δ l'ARL della carta per il coefficiente di variazione non cambia. Le piccole fluttuazione nell'ARL, per diversi valori di δ per λ fissato, possono dunque essere considerati solo errori di simulazione. Infatti per la combinazione $\lambda = 1$, indipendentemente dal valore di δ , l'ARL per la carta CV è pari a 370, per $\lambda = 1.25$ l'ARL rimane pari a 43.6 indipendentemente da δ , e così via. Solo un cambiamento della varianza ($\lambda\sigma$) fa reagire la carta Shewhart per il coefficiente di variazione.

La carta sintetica è però tra le quella che segnala per prima quando la variabilità del processo aumenta, e questo per tutti i valori di λ e per differenti valori del coefficiente di variazione. Inoltre, a differenza dell'ARL per la carta Shewhart CV, per un dato valore di λ , l'ARL della carta sintetica non aumenta al crescere del valore del coefficiente di variazione, ma rimane semmai pressoché costante.

Alla luce di questi risultati diventa chiaro che la carta sintetica sembra più efficiente della carta per il coefficiente di variazione. La carta sintetica non solo mostra un guadagno di efficienza nel controllare la media e varianza di un processo ma è anche relativamente più

facile da usare perché non richiedendo la verifica di proporzionalità della deviazione standard con la media. In particolare i risultati inducono a suggerire l'utilizzo della carta per il coefficiente di variazione per valori piccoli del coefficiente di proporzionalità tra la deviazione standard e la media .

Al fine di investigare i profili dell'ARL al variare della direzione del cambiamento della varianza o del cv è stato anche esaminato il caso in cui la varianza decresce ($\lambda = 1, 0.95, 0.90, 0.80, \dots$) facendo decrescere anche il cv . Quello che si nota per la carta sintetica è una crescita continua dell'ARL, ad ogni riduzioni di 0.05 nella varianza si nota più di un raddoppio dell'ARL. Addirittura per $\lambda > 0.85$, la carta sintetica non si accorge quasi mai di questa riduzione mentre la carta CV mantiene una relativa sensitività anche rispetto alle riduzioni di variabilità. Sembra dunque che la carta CV mostra rispetto alla carta sintetica un qualche vantaggio nel caso di riduzioni della variabilità. Tuttavia l'individuazione di riduzioni nella variabilità del processo corrispondenti a un suo miglioramento ha una rilevanza minore nella maggior parte delle applicazioni pratiche.

Tabella 3.1 Profili dell'ARL carta per il coefficiente di variazione e la carta sintetica $\bar{X}-R$ per $n = 5$ per differenti valori di λ (aumento di λ)

λ	$CV = 0.02$ ARL		$CV = 0.05$ ARL		$CV = 0.1$ ARL		$CV = 0.16$ ARL		$CV = 0.20$ ARL		$CV = 0.26$ ARL		$CV = 0.30$ ARL	
	CV	$\bar{X}-R$	CV	$\bar{X}-R$	CV	$\bar{X}-R$	CV	$\bar{X}-R$	CV	$\bar{X}-R$	CV	$\bar{X}-R$	CV	$\bar{X}-R$
1	370.24	370.22	370	370	370	369.19	370.49	370.82	370.72	370.05	370.45	369	370.34	370.01
1.05	253.29	202.52	250.84	204.97	250.72	202.41	255.96	202.93	259.62	205.33	255.28	203.83	257.12	201.72
1.10	160.17	118.72	160.84	118.43	160.11	117.65	161.03	116.88	163.28	119.64	167.32	117.48	170.46	118.07
1.15	100.49	73.1	101.13	73.46	102.21	74.77	103.56	74.11	101.96	73.61	107.88	72.4	110.75	72.14
1.20	63.45	48.71	65.01	48.52	65.9	48.75	67.63	48.43	67.67	48.78	71.36	49	73.57	49.3
1.25	42.93	33.11	44.49	33	44.04	33.48	45.21	33.71	46.63	33.44	48.63	33.67	50.3	33.58
1.30	30.01	23.85	30.42	24.06	30.8	23.95	31.94	24.03	32.72	24.2	34.42	24.38	35.82	24.46
1.35	22.23	18.38	22.25	17.92	22.2	18.35	23.65	18.13	24.17	18.08	25.79	18.14	27.18	18.16
1.40	16.62	14.09	17.06	13.96	17.25	13.98	17.36	13.89	18.5	13.96	19.53	13.89	20.28	14
1.45	12.84	11.13	13.33	10.97	13.29	11.14	14.14	11.05	14.42	10.99	15.31	11.06	16.58	10.92
1.50	10.3	8.97	10.45	8.88	10.84	8.91	11.28	8.96	11.7	8.97	12.55	8.98	13.19	8.97
1.55	8.76	7.41	8.74	7.47	8.87	7.41	9.15	7.35	9.53	7.35	10.26	7.35	10.72	7.46
1.60	7.16	6.32	7.23	6.28	7.35	6.2	7.79	6.25	7.92	6.13	8.61	6.29	9.15	6.25
1.65	6.09	5.3	6.02	5.34	6.25	5.38	6.62	5.36	6.85	5.45	7.29	5.35	7.81	5.4
1.70	5.4	4.66	5.32	4.6	5.43	4.7	5.68	4.61	5.98	4.63	6.44	4.68	6.76	4.69
1.75	4.69	4.05	4.68	4.11	4.82	4.1	5.03	4.13	5.14	4.1	5.63	4.11	5.97	4.1
1.80	4.15	3.66	4.16	3.65	4.22	3.67	4.37	3.65	4.62	3.64	5.01	3.67	5.42	3.67
1.85	3.78	3.35	3.78	3.31	3.91	3.33	4.03	3.31	4.12	3.29	4.48	3.29	4.85	3.33
1.90	3.4	3.01	3.38	3.016	3.49	3.06	3.66	3.02	3.78	3.03	4.06	3.03	4.37	3.05
1.95	3.11	2.78	3.15	2.74	3.18	2.8	3.34	2.74	3.46	2.76	3.74	2.79	4	2.81
2	2.88	2.57	2.88	2.58	2.98	2.58	3.04	2.56	3.21	2.58	3.54	2.55	3.74	2.6

Tabella 3.2 Differenze tra L'ARL della carta per il cv e l'ARL della carta sintetica $\bar{X}-R$ per ogni valore del cv al variare di λ .

λ	<i>CV</i> = 0.02 Diff. ARL	<i>CV</i> = 0.05 Diff. ARL	<i>CV</i> = 0.1 Diff. ARL	<i>CV</i> = 0.16 Diff. ARL	<i>CV</i> = 0.20 Diff. ARL	<i>CV</i> = 0.26 Diff. ARL	<i>CV</i> = 0.30 Diff. ARL	Media Diff. ARL
	<i>CV</i> - ($\bar{X}-R$)	<i>CV</i> - ($\bar{X}-R$)	<i>CV</i> - ($\bar{X}-R$)	<i>CV</i> - ($\bar{X}-R$)	<i>CV</i> - ($\bar{X}-R$)	<i>CV</i> - ($\bar{X}-R$)	<i>CV</i> - ($\bar{X}-R$)	<i>CV</i> - ($\bar{X}-R$)
1.05	50.77	45.87	48.31	53.03	54.29	51.45	55.4	51.30
1.10	41.45	42.41	42.46	44.15	43.64	49.84	52.39	45.19
1.15	27.39	27.67	27.46	29.45	28.35	35.48	38.61	30.63
1.20	14.74	16.49	27.44	19.2	18.89	22.36	24.27	19.01
1.25	9.82	11.49	17.15	11.5	13.19	14.96	16.72	12.61
1.30	6.16	6.36	10.56	7.91	8.52	10.04	11.36	8.17
1.35	3.85	4.33	6.85	5.52	6.09	7.65	9.02	5.76
1.40	2.53	3.1	3.85	3.47	4.54	5.64	6.28	4.12
1.45	1.71	2.36	3.27	3.09	3.43	4.25	5.66	3.24
1.50	1.33	1.57	2.15	2.32	2.73	3.57	4.22	2.52
1.55	1.35	1.27	1.93	1.8	2.18	2.91	3.26	2.03
1.60	0.84	0.95	1.46	1.54	1.79	2.32	2.9	1.64
1.65	0.79	0.68	1.15	1.26	1.4	1.94	2.41	1.34
1.70	0.74	0.72	0.87	1.07	1.35	1.76	2.07	1.21
1.75	0.64	0.57	0.73	0.9	1.04	1.52	1.87	1.04
1.80	0.49	0.51	0.72	0.72	0.98	1.34	1.75	0.91
1.85	0.43	0.47	0.55	0.72	0.83	1.19	1.52	0.82
1.90	0.39	0.36	0.58	0.64	0.75	1.03	1.32	0.70
1.95	0.33	0.41	0.43	0.6	0.7	0.95	1.19	0.65
2	0.31	0.3	0.38	0.48	0.63	0.99	1.14	0.60

Tabella 3.3 Profili dell'ARL della carte per il cv e la carta sintetica al variare di λ e δ per $n = 5$.

$CV = 0.05$			
λ	δ	ARL $_{cv}$	Sintetica
1	0	370.2	370.50
	0.25	369.83	176.68
	0.5	370.56	48.42
	0.75	365.76	14.72
	1	372.27	5.85
	1.25	370.41	2.9
	1.5	370.78	1.78
	2	371.78	1.11
1.25	0	44.49	33
	0.25	43.63	26.62
	0.5	43.05	14.89
	0.75	43.53	7.71
	1	43.48	4.27
	1.25	43.84	2.56
	1.5	43.4	1.79
	2	43.34	1.17
1.5	0	10.45	8.88
	0.25	10.75	8.19
	0.5	10.59	6.34
	0.75	10.65	4.48
	1	10.71	3.09
	1.25	10.4	2.26
	1.5	10.56	1.71
	2	10.57	1.22
1.75	0	4.68	4.11
	0.25	4.62	3.96
	0.5	4.73	3.44
	0.75	4.71	2.91
	1	4.8	2.35
	1.25	4.7	1.91
	1.5	4.67	1.6
	2	4.65	1.24
2	0	2.88	2.58
	0.25	2.87	2.51
	0.5	2.84	2.31
	0.75	2.93	2.13
	1	2.9	1.89
	1.25	2.92	1.64
	1.5	2.92	1.47
	2	2.85	1.21

Capitolo 4

Conclusioni

Nei capitoli precedenti sono state presentate due carte di controllo, la carta per il coefficiente di variazione (Kang et al., 2007) e la carta sintetica per la media e la varianza (Costa, Rahim, 2006). E' stata illustrata la metodologia per la loro implementazione e descritte. La carta Shewhart per il coefficiente di variazione ha la particolarità di poter essere applicata su processi per i quali non è possibile utilizzare le carte Shewhart \bar{x} e R non possono essere applicate. Ad esempio quei processi che, anche in controllo presentano media e varianza non stabile. Nel caso in particolare in cui la deviazione standard σ del processo è proporzionale alla media μ , Kang et. al. (2007) suggeriscono la sorveglianza della sola variabilità del processo per mezzo di una carta Shewhart per il coefficiente di variazione.

Tuttavia, sempre nell'ipotesi che il coefficiente di variazione sia costante, Kang et. al (2007) non confrontano l'efficienza di questa carta con quella di altre carte simili.

In questo elaborato vengono dunque messe a confronto, in termini di Average Run Length, la performance delle carta Shewhart per il coefficiente di variazione e della carta sintetica. In particolare nel terzo capitolo, tre tabelle riportano i profili dell'ARL, delle due carte a confronto, simulati con criteri uguali per le due carte.

Il confronto dell'ARL delle due carte ha permesso di misurare l'efficienza di ciascuna carta. Si trova che a parità di condizioni, la carta sintetica per la media e la varianza risulta più efficiente della carta per il coefficiente di variazione disegnata per il controllo della sola

variabilità. Inoltre la carta per il coefficiente di variazione perde la sua efficienza in maniera progressiva mano a mano che aumenta il coefficiente di variazione, laddove questo non accade utilizzando la carta sintetica per la media e la varianza la cui performance non sembra variare al variare del coefficiente di variazione.

Kang et al. (2007) sottolineano che la sensibilità della carta Shewhart per il CV aumenta all'aumentare della dimensione dei singoli campioni. Possibili approfondimenti dei confronti tra carta Shewhart CV e carta sintetica riguardano l'esplorazione dell'effetto della dimensione campionaria, ovvero studiare per quale valore di n , maggiore di 5, la carta per il CV migliora la propria efficienza rispetto alla carta sintetica risulta più efficiente della carta sintetica a parità di CV . Inoltre è anche interessante esplorare, oltre che la loro dimensione, anche il numero m di campioni di Phase I da utilizzare a fine di avere una maggiore efficienza della carta per il CV.

Appendice A

Funzioni R usati per le simulazioni dei profili dell'ARL per le due carte

A.1 Funzioni usati per il calcolo del profilo dell' ARL per la carta cv

A.1.1 arsd

Questa funzione calcola la deviazione standard dato il cv e mu (la media in controllo) , è usata nella funzione nella funzione **RLcv** che segue per potere simulare delle normali con un cv costante.

```
arsd<-function(mu,cv) {  
sd=cv*mu  
sd  
}
```

A.1.2 RLcv

RLcv è la funzione che calcola il Run Length per la carta cv

```

RLcv<-function(mu,cv ,shiftm,shiftv,n,seq,LCL_cv,UCL_cv)
{
i=1
coef=as.numeric(seq)
mu=mu+shiftm
sigma=arsd(mu,cv)
sigma=sigma*shiftv
while(i <= seq)
{
x=rnorm(n,mean=mu,sigma)
coef[i]=sd(x)/mean(x)
if((coef[i]>LCL_cv)&&(coef[i]<UCL_cv)){
i = i+1
}
else {
break
}

}
i
}

```

A.1.3 **ARLcv**

Questa funzione calcola l'ARL della carta cv facendo la media delle nsim replicazione dei RLcv.

```

ARLcv<-function(mu,cv,shiftm,shiftn,n,seq,nsim,LCL_cv,UCL_cv)
{
  RL=as.numeric(nsim)
  for (i in 1:nsim)
    {
      RL[i]=RLcv(mu,cv,shiftm,shiftn,n,seq,LCL_cv,UCL_cv)
    }
  ARL_cv=mean(RL)
  ARL_cv
}

```

A.1.4 **ARLcv_shvar**

La funzione `ARLcv_shvar` è quella finale combinando tutte le altre precedenti ci ritorna il profilo dell'ARL della carta cv al variare di $\lambda = \text{shiftn}$

```

shiftn=seq(1,2,by=0.05)
k<-length(shiftn)
ARLcv_shvar<-
function(mu,cv,shiftm,shiftn,n,seq,nsim,LCL_cv,UCL_cv) {
  A=as.numeric(k)
  for(i in 1:length(shiftn)) {
    A[i]=ARLcv(mu,cv,shiftm,shiftn[i],n,seq,nsim,LCL_cv,UCL_cv)
  }
  p=data.frame(Shiftvar=shiftn,ARL.cv=A)
}

```

```
p
}
```

A.2 Funzioni usati per il calcolo del profilo dell' ARL per la carta sintetica

A.2.1 RLMR

La funzione RLMR calcola il Run Length per la carta sintetica .

```
RLMR<-function(mu,sigma,shiftm,shiftv,n,seq,LCLm,UCLm,LCLr,UCLr)
```

```
{
  i = 1
  m=as.numeric(seq)
  r<-as.numeric(seq)
  mu=mu+shiftm
  sigma=sigma*shiftv
  while( i <= seq )
  {
    x=rnorm(n,mean=mu,sigma)
    m[i]=mean(x)
    r[i]=range(x)[2]-range(x)[1]
    if((m[i]>LCLm)&&(m[i]<UCLm)&&(r[i]>LCLr)&&(r[i]<UCLr)){
      i = i+1
    }
  }
  else{
    break
  }
}
```

```

    }
  i
}

```

A.2.2 **ARLMR**

ARLMR calcola l'ARL per la carta sintetica, facendo la media dei nsim replicazione dei RLMR .

```

ARLMR<-
function(mu,sigma,shiftm,shiftv,n,seq,nsim,LCLm,UCLm,LCLr,UCLr)
{
  RL=as.numeric(nsim)
  for (i in 1:nsim)
  {
    RL[i]=RLMR(mu,sigma,shiftm,shiftv,n,seq,LCLm,UCLm,LCLr,UCLr)
  }
  ARL=mean(RL)
  ARL
}

```

A.2.3 **ARLMR_shvar**

Questa è la funzione che ci da il profilo del' ARL per la carta sintetica al variare di λ =shiftv

```
shiftv=seq(1,2,by=0.05)
k<-length(shiftv)
ARLMR_shvar<-
function(mu,sigma,shiftm,shiftv,n,seq,nsim,LCLm,UCLm,LCLr,UCLr) {
  A=as.numeric(k)
  for(i in 1:length(shiftv)){
A[i]=ARLMR(mu,sigma,shiftm,shiftv[i],n,seq,nsim,LCLm,UCLm,LCLr,UCLr)
  }
  p=data.frame(Shiftvar=shiftv,ARL.MR=A)
  p
}
```

A.2.4 **ARLMR_shmed**

Questa ultima ci calcola il profilo del' ARL per la carta sintetica al variare di δ = shiftm

```
shiftm=seq(0,2,by=0.05)
k<-length(shiftm)
ARLMR_shmed<-
function(mu,sigma,shiftm,shiftv,n,seq,nsim,LCLm,UCLm,LCLr,UCLr) {
  A=as.numeric(k)
  for(i in 1:length(shiftm)){
```

```
A[i]=ARLMR(mu,sigma,shiftm[i],shiftv,n,seq,nsim,LCLm,UCLm,LCLr,UCLr)
    }
    p=data.frame(Shiftmedia=shiftm,ARL.MR=A)
p
}
```

Bibliografia

Abramowitz, M. e Stegun, I. A. (1972) . Handbook of Mathematical functions. Dover, New York.

Connett, J.E. e Lee, W.W. (1990) . "Estimation of the Coefficient of Variation from Laboratory Analysis of Split Specimens for Quality Control in Clinical Trials". Controlled Clinical Trials 11, P. 24-36.

Costa, A.F.B. e Rahim M.A. (2006) A synthetic control chart for monitoring the process mean and variance Journal of quality in Maintenance Engineering, 12, p. 81-88.

Deming W. E. (1982) out of the Crisis. Massachusetts Institute Of Technology.

Douglas M. Hawkins, QiQi Deng (2009) Combined Charts for mean and Variance Information Journal of quality Technology, 41, p. 415-425.

Hendricks, W.A. e Robey, K. (1936). "The Sampling Distribution of the Coefficient of Variation". The Annals of Mathematical Statistics 7, p. 129-132.

Iglewicz, B. (1967). "Some Properties of the Coefficient of Variation". Ph.D. Dissertation, Virginia Polytechnic Institute.

Kang C. W., Lee M.S. , Seong Y. J., Hawkins D. M. (2007), A Control Chart for the Coefficient of Variation, Journal of Quality Technology, 39, p. 151-158.

Lenth, R. V. (1989). "Algorithm AS 243 Cumulative Distribution Function of the Non-Central t Distribution" .Applied Statistics 38, p. 185-189.

McKay. A. T. (1932). "Distribution of the Coefficient of Variation and the Extended t Distribution" . Journal of the Royal Statistical Society 95, p. 695 . 699.

Verrill, S. (2003). "Confidence Bounds for Normal and Lognormal Distribution Coefficients of Variation". Res. Pap. FPL-RP-609. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.