



UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE

CORSO DI LAUREA IN
STATISTICA E TECNOLOGIE INFORMATICHE

TESI DI LAUREA TRIENNALE

MODELLO CHANGE-POINT E OMNIBUS EWMA A CONFRONTO

RELATORE: Ch.ma Prof.ssa Capizzi Giovanna

Laureando: Scarso Nicola
Matricola: 499175

ANNO ACCADEMICO 2005/2006

Indice

Introduzione	i
Capitolo 1 Modello change-point	1
1.1 Formulazione del change-point	2
1.2 Change-point per la media del processo	3
1.3 Change-point per la varianza	4
1.3.1 Osservazione continua del processo per l'individuazione di un cambiamento in varianza	6
1.4 Limiti di controllo	7
1.5 Disegno delle carte	10
1.6 Performance	11
Capitolo 2 Omnibus EWMA	13
2.1 Proprietà dell'omnibus EWMA	14
2.2 Disegno dell'omnibus EWMA	14
2.3 Performance	15
Capitolo 3 Confronto e risultati	19
3.1 Comparazione tra il change-point per la varianza e l'omnibus EWMA	19
Conclusioni	25
Appendice	27
Bibliografia	41

Introduzione

Le proprietà delle carte di controllo tradizionali sono basate sull'assunzione che i valori della media e della varianza in controllo devono essere conosciuti esattamente. Questi normalmente dovrebbero essere usati per il disegno della linea centrale, dei limiti di controllo e degli intervalli di decisione.

Nella realtà, solo raramente sono noti e vengono stimati estraendo dei campioni di dimensione $m \cdot n$ da un processo che si presume in controllo.

I parametri che si ottengono dalla stima vengono quindi considerati come se fossero i veri parametri, anche se questo è lontano dal vero (fase I).

Di conseguenza gli errori casuali di tali stime, determinati anche prendendo larghi campioni nella fase I, portano a distorcere nella fase II la performance delle carte, ottenendo una incerta distribuzione della run length.

Un'alternativa a queste carte tradizionali è una carta che permette di individuare cambiamenti in media, in varianza o in entrambe per il caso in cui il valore nominale di tali parametri non è conosciuto. Questo eviterebbe, quindi, la necessità di una lunga raccolta di dati nella fase I per la stima dei parametri stessi.

Nel primo capitolo si descrivono le caratteristiche del metodo del "change-point" per un punto di cambio in media (Hawkins, Qui e Kang (2003)), e in varianza (Hawkins e Zamba (2005)), nel quale non viene fatta nessuna assunzione sulla conoscenza dei parametri.

Nel secondo capitolo si descrivono invece, le caratteristiche dell'omnibus EWMA (Domangue e Patch (1991)) che è capace di individuare cambiamenti sia in media sia in varianza, nella quale però vengono stabiliti precedentemente i valori di μ_0 e σ_0 .

Con il presente lavoro si vuole quindi mettere a confronto due metodi con caratteristiche che alla base sono ben diverse: uno che si basa sulla stima adattiva dei parametri e l'altro che si basa sui veri parametri.

Nel terzo capitolo vengono riportati i risultati ottenuti, tramite simulazione, cambiando la deviazione standard dalla 50-esima osservazione e scartando ogni

sequenza nella quale il test del change-point per la varianza o l'omnibus EWMA segnalava un falso allarme.

Nell'appendice viene riportato il codice C utilizzato per ottenere i diversi risultati riportati, mentre nella bibliografia si può trovare il riferimento ad articoli di riviste che possono essere utili per assistere il lettore nell'applicazione dei metodi descritti.

Capitolo 1

Modello change-point

Il controllo statistico di un processo produttivo (SPC, Statistical Process Control) è caratterizzato da un insieme di strumenti, utili per migliorare la produttività attraverso la riduzione della variabilità, e per raggiungere la stabilità del processo.

Il principale obiettivo delle carte di controllo si propone quindi di individuare gli istanti nel quale il processo è andato fuori controllo. Questo è di vitale importanza per investigare le cause assegnabili alla perdita di controllo, che solitamente possono essere di due tipi: le cause comuni, che rappresentano la variabilità naturale, o rumore di fondo, di tutte le risorse non evitabili che possono essere rimosse solo cambiando il sistema, e le cause speciali che risultano da alcune risorse, come macchinari non funzionanti, errori dovuti agli operatori o materiali grezzi difettosi, che sono potenzialmente identificabili e che possono essere rimosse.

Un processo viene detto in uno stato statistico di controllo quando la variabilità è provocata solo dalle cause comuni, mentre quando interviene una causa speciale verrà detto fuori controllo.

L'assunzione più comune e più semplice per un modello di SPC è, che in uno stato di controllo, le osservazioni consecutive del processo X_i siano indipendenti e seguano una distribuzione normale.

Gli strumenti tradizionali di SPC per individuare grandi salti di livello sono le carte Shewhart \bar{X} (per la media del processo), R (per i range campionari) o S (per la deviazione standard di gruppi razionali). Tali carte non sono sensibili a piccole modifiche del livello del processo, perché la loro caratteristica principale è rappresentata dal fatto che usano solo l'ultimo istante di osservazione del processo come fonte di informazione sullo stesso ignorando tutte le precedenti. Di conseguenza due valide alternative per individuare piccoli salti di livello in media sono la carta CUSUM (Location Cumulative Sum) e la carta EWMA (Exponentially Weighted Moving Average). Per controllare invece salti di livello nella deviazione

standard sono state presentate alcune estensioni e varianti sia per l'EWMA – si veda per esempio MacGregor e Harris (1993), Crowder e Hamilton (1992) - sia per la CUSUM – i dettagli sono in Hawkins e Olwell (1998).

L'approccio che presento nei paragrafi successivi, è invece un metodo meno familiare rispetto ai tradizionali metodi di SPC presentati, ma consente di affrontare la mancata conoscenza dei parametri della distribuzione in controllo.

1.1 FORMULAZIONE DEL CHANGE-POINT

La formulazione statistica del change-point è la seguente:

$$X_i \sim \begin{cases} N(\mu_1, \sigma_1^2) & \text{se } i \leq \tau \\ N(\mu_2, \sigma_2^2) & \text{se } i > \tau \end{cases}$$

dove $X_1, X_2, \dots, X_i, \dots$ sono osservazioni successive del processo, μ_1 la media in controllo, μ_2 il valore della media fuori controllo, σ_1^2 la varianza in controllo e σ_2^2 il valore della varianza fuori controllo.

Se esiste, τ rappresenta l'istante di tempo incognito in cui può realizzarsi un cambiamento in media, se $\mu_1 \neq \mu_2$ e/o in varianza, se $\sigma_1^2 \neq \sigma_2^2$.

L'uso di questo modello nell'SPC, ha due compiti: quello di stimare i parametri incogniti e quello di verificare la presenza di un cambiamento in media, in varianza o in entrambe, se effettivamente il cambiamento c'è stato.

Hawkins, Qui e Kang (2003, abbreviazione HQK) hanno proposto questa formulazione per un punto di cambio in media considerando la varianza costante, mentre Hawkins e Zamba (2005) hanno sviluppato una metodologia parallela per un punto di cambio in varianza, ma con nessuna assunzione sulla costanza o diversità delle medie.

Entrambe le formulazioni di campionamento fisso sono adattate all'impostazione della fase II, definendo un contesto ripetuto di verifica d'ipotesi nel quale, non appena arriva una nuova osservazione, il test del change-point viene eseguito per tutti i dati accumulati, mantenendo costante la probabilità dei falsi allarmi.

1.2 CHANGE-POINT PER LA MEDIA DEL PROCESSO

HQK (2003), in alternativa alla carta CUSUM e alla carta \bar{X} per la media hanno proposto l'uso della metodologia change-point per un punto di cambio in media assumendo la varianza costante:

$$X_i \sim \begin{cases} N(\mu_1, \sigma^2) & \text{se } i \leq \tau \\ N(\mu_2, \sigma^2) & \text{se } i > \tau \end{cases}$$

μ_1 , μ_2 , σ^2 e τ sono tutti incogniti.

Sia X_i la successione di osservazioni dal processo che seguano la formulazione HQK. Per $1 \leq j \leq n-1$ sono definite le seguenti statistiche di sintesi:

la media delle prime j osservazioni

$$\bar{X}_{jn} = \sum_{i=1}^j X_i / j$$

la media delle rimanente osservazioni

$$\bar{X}_{jn}^* = \sum_{i=j+1}^n X_i / (n-j)$$

e la somma dei quadrati residua

$$V_{jn} = \sum_{i=1}^j (X_i - \bar{X}_j)^2 + \sum_{i=j+1}^n (X_i - \bar{X}_j^*)^2$$

Basandosi sull'assunzione che c'è un solo punto di cambio j , le statistiche appena descritte delineano in tale istante le stime di massima verosimiglianza di μ_1 e μ_2 , mentre $\hat{\sigma}_{jn}^2 = V_{jn} / (n-2)$ rappresenta la stima di σ^2 .

Il test t a due campioni (two-sample t-statistic) per mettere a confronto i due insiemi è così definito:

$$T_{jn} = \sqrt{\frac{j \cdot (n-j)}{n}} \cdot \frac{\bar{X}_{jn} - \bar{X}_{jn}^*}{\hat{\sigma}_{jn}}$$

Sotto l'ipotesi nulla di assenza di cambiamento, per j scelto arbitrariamente, T_{jn} dovrebbe seguire una distribuzione t di Student's con $n-2$ gradi di libertà.

Se invece c'è la presenza di un punto di cambio, il test del rapporto di verosimiglianza generalizzato (abbreviato con GLR, Generalized likelihood ratio) consiste nel trovare $T_{\max,n}$, ossia il massimo di $|T_{jn}|$ tra tutti i possibili valori di j .

Il j trovato è quindi una stima di τ , mentre \bar{X}_j e \bar{X}_j^* sono le stime di massima verosimiglianza delle medie sconosciute μ_1 e μ_2 .

Questa formulazione può essere adattata all'impostazione della fase II come segue:

- dopo che l'osservazione n è stata aggiunta al record dei dati accumulati del processo, calcolare $T_{\max,n}$, ossia la statistica GLR per un punto di cambio in media avvenuto in qualche istante precedente;
- se $T_{\max,n} \leq h_n$, dove h_n è qualche limite di controllo adatto, si conclude che non c'è stato nessun cambiamento evidente in media;
- se $T_{\max,n} > h_n$ si conclude che c'è stato un evidente cambiamento in media.

1.3 CHANGE-POINT PER LA VARIANZA

Hawkins e Zamba (2005), come già accennato in precedenza, hanno sviluppato una metodologia parallela a quella di HQK per un punto di cambio in varianza con nessuna assunzione sull'andamento della media.

La formulazione statistica è la seguente:

$$X_i \sim \begin{cases} N(\mu_1, \sigma_1^2) & \text{se } i \leq \tau \\ N(\mu_2, \sigma_2^2) & \text{se } i > \tau \end{cases}$$

μ_1 , μ_2 , σ_1^2 , σ_2^2 e τ sono incogniti.

Sia X_i la successione di osservazioni del processo che seguano la formulazione del punto di cambio in varianza, per $0 \leq i < k \leq n$ sono definite le seguenti statistiche di sintesi:

$$\bar{X}_{i,k} = \sum_{j=i+1}^k X_j / (k-i)$$

$$V_{i,k} = \sum_{j=i+1}^k (X_j - \bar{X}_{i,k})^2$$

Supponendo che l'istante $\tau = k$ sia il punto di cambio, la stima dei rimanenti parametri è la seguente:

$$\hat{\mu}_1 = \bar{X}_{0,k}$$

$$\hat{\mu}_2 = \bar{X}_{k,n}$$

$$\hat{\sigma}_1^2 = V_{0,k} / (k-1)$$

$$\hat{\sigma}_2^2 = V_{k,n} / (n-k-1)$$

mentre sotto l'assunzione che la varianza sia costante, la stima di quest'ultima dovrebbe essere così determinata:

$$\hat{\sigma}^2 = (V_{0,k} + V_{k,n}) / (n-2)$$

Come si può notare le stime delle varianze $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}^2$ hanno l'usuale divisore dei gradi di libertà per renderle non distorte.

Il test statistico del rapporto di verosimiglianza per l'uguaglianza delle varianze è così definito:

$$G_{k,n} = \left[(k-1) \cdot \ln \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_1^2} \right) + (n-k-1) \cdot \ln \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_2^2} \right) \right] / C$$

dove C rappresenta il fattore di correzione di Bartlett (Bartlett e Kendall(1946), Bartlett (1937), Bartlett (1955)), ossia

$$C = 1 + [(k-1)^{-1} + (n-k-1)^{-1} - (n-2)^{-1}] / 3$$

$G_{k,n}$ sotto l'ipotesi nulla ha una distribuzione chi-quadro con un grado di libertà.

Invece, nella situazione di un possibile cambiamento in varianza, all'istante sconosciuto τ , il test GLR è dato dal massimo $G_{k,n}$, lungo tutti i possibili valori di k . Facendo riferimento proprio al range di questi possibili valori si può notare che G non è definito per $k=1$ o $k=n-1$. Di conseguenza il test statistico GLR per un cambiamento in varianza è il seguente:

$$G_{\max,n} = \max_{2 \leq k \leq n-2} G_{k,n}$$

1.3.1 OSSERVAZIONE CONTINUA DEL PROCESSO PER L'INDIVIDUAZIONE DI UN CAMBIAMENTO IN VARIANZA

Parallelamente a quanto descritto per la formulazione HQK per il controllo della media, si assume che un cambiamento in varianza possa esserci in qualsiasi istante. La procedura di SPC viene poi definita nel seguente modo:

- dopo che l'osservazione n è stata aggiunta al record totale dei dati accumulati, viene calcolata la statistica $G_{\max,n}$, ossia la statistica GLR per un punto di cambio in varianza avvenuto in qualche istante precedente;
- se $G_{\max,n} \leq h_n$, dove h_n è qualche limite di controllo adatto, si conclude che non c'è stato nessun cambiamento evidente in varianza;
- se $G_{\max,n} > h_n$, si conclude che c'è stato un evidente cambiamento in varianza.

Si possono notare alcune importanti differenze tra la formulazione del change-point e i più tradizionali metodi di SPC segnalati in precedenza. Una è la considerazione già fatta nell'introduzione di questo lavoro che riguarda la comune usanza di tarare le carte tradizionali basandosi sulle stime dei parametri ottenute dai dati raccolti nella fase I, e di trattare queste stime come se fossero i veri valori.

La formulazione change-point rimarca che le stime non sono più che stime e si propone di controllare il tasso dei falsi allarmi senza nessuna assunzione sulla conoscenza dei parametri.

Un'altra differenza è che, diversamente dall'impostazione convenzionale, dove l'acquisizione di informazioni sui veri parametri è limitata alla raccolta dei dati con la fase I, la formulazione change-point continua ad usare l'insieme dei dati progressivi accumulati per aggiornare le stime dei parametri del processo.

1.4 LIMITI DI CONTROLLO

Una questione importante da risolvere quando si deve disegnare una carta di controllo, è la scelta di una conveniente successione di limiti h_n .

HQK hanno proposto che tale sequenza deve essere definita da una proprietà intuitivamente desiderabile (Margavio et al. (1995)), ossia che mentre il processo è in controllo, la probabilità di un falso allarme è fissata dall'utente a qualche livello costante α , ossia

$$\Pr[T_{\max,n} > h_n \mid T_{\max,j} \leq h_j, j < n] = \alpha$$

Poiché non sembra possibile calcolare teoricamente questa sequenza di valori $h_{n,\alpha}$, HQK li hanno stimati via simulazione.

I valori di α presi in considerazione sono: 0.05, 0.02, 0.01, 0.005, 0.002 e 0.001. L'ARL in controllo corrisponde quindi a $1/\alpha$.

Dato che un minimo di due osservazioni per campione sono richieste per calcolare un test statistico, HQK pensano che pochi utenti comincerebbero dalla quarta osservazione ad applicare la verifica d'ipotesi.

Sebbene la formulazione del change-point non richiede una grande raccolta di dati nella fase I per stimare in modo accurato i parametri del processo, la pratica standard implica almeno una raccolta attenta di alcune osservazioni prima di applicare un test statistico.

Per questo motivo la simulazione di HQK incorpora questa possibilità, dando luogo a 17 tavole di valori $h_{n,\alpha}$, corrispondenti ai salti iniziali delle prime 3,4,...,19 osservazioni. La tabella completa è disponibile sul sito web www.stat.umn.edu/hawkins.

Un ragionevole approccio, ed è quello che ho adottato nel seguente lavoro, è di saltare le prime 9 osservazioni e cominciare con la verifica d'ipotesi dalla decima.

Usare dei limiti che si trovano in una tabella potrebbe essere un po' noioso da implementare, e quindi è da aiuto avere un modo più compatto, anche se approssimativo, di calcolare dei limiti adatti.

Basato su questo ragionamento, l'approccio suggerito da HQK per avere delle soglie adatte per il change-point in media è il seguente:

$$h_{n,\alpha} \approx h_{10,\alpha} \cdot \left(0.677 + 0.019 \cdot \ln(\alpha) + \frac{1 - 0.115 \cdot \ln(\alpha)}{n - 6} \right)$$

La tabella 1.1 mostra i limiti di controllo ottenuti applicando tale approssimazione per gli stessi α proposti da HKQ per n da 10 a 60.

n	α					
	0.05	0.02	0.01	0.005	0.002	0.001
10	3.662	4.371	4.928	5.511	6.340	7.023
11	3.255	3.902	4.413	4.950	5.718	6.353
12	3.091	3.691	4.161	4.654	5.355	5.933
13	2.974	3.540	3.982	4.443	5.097	5.633
14	2.886	3.426	3.847	4.285	4.902	5.408
15	2.818	3.338	3.743	4.162	4.751	5.233
16	2.763	3.268	3.659	4.063	4.631	5.093
17	2.718	3.210	3.590	3.982	4.532	4.978
18	2.681	3.162	3.533	3.915	4.449	4.883
19	2.649	3.122	3.485	3.858	4.380	4.802
20	2.622	3.087	3.443	3.810	4.320	4.733
21	2.599	3.057	3.408	3.767	4.268	4.673
22	2.578	3.030	3.376	3.730	4.223	4.620
23	2.560	3.007	3.348	3.698	4.183	4.574
24	2.544	2.986	3.324	3.669	4.148	4.533
25	2.530	2.968	3.302	3.643	4.116	4.496
26	2.517	2.951	3.282	3.620	4.087	4.463
27	2.505	2.936	3.264	3.598	4.061	4.433
28	2.495	2.922	3.248	3.579	4.038	4.406
29	2.485	2.910	3.233	3.562	4.016	4.381
30	2.476	2.898	3.219	3.546	3.997	4.358
31	2.468	2.888	3.207	3.531	3.978	4.337
32	2.460	2.878	3.195	3.517	3.962	4.318
33	2.453	2.869	3.184	3.505	3.946	4.300
34	2.447	2.861	3.174	3.493	3.932	4.283
35	2.440	2.853	3.165	3.482	3.918	4.267
36	2.435	2.846	3.156	3.472	3.906	4.253
37	2.430	2.839	3.148	3.462	3.894	4.239
38	2.425	2.832	3.141	3.453	3.883	4.227
39	2.420	2.826	3.133	3.445	3.873	4.215
40	2.416	2.821	3.127	3.437	3.863	4.203
41	2.411	2.815	3.120	3.430	3.854	4.193
42	2.408	2.810	3.114	3.423	3.846	4.183
43	2.404	2.806	3.109	3.416	3.837	4.173
44	2.400	2.801	3.103	3.410	3.830	4.164
45	2.397	2.797	3.098	3.404	3.822	4.156
46	2.394	2.793	3.094	3.398	3.815	4.148
47	2.391	2.789	3.089	3.392	3.809	4.140
48	2.388	2.785	3.085	3.387	3.802	4.133
49	2.385	2.782	3.080	3.382	3.796	4.126
50	2.383	2.778	3.076	3.378	3.791	4.119
51	2.380	2.775	3.073	3.373	3.785	4.113
52	2.378	2.772	3.069	3.369	3.780	4.107
53	2.375	2.769	3.065	3.365	3.775	4.101
54	2.373	2.766	3.062	3.361	3.770	4.095
55	2.371	2.764	3.059	3.357	3.765	4.090
56	2.369	2.761	3.056	3.354	3.761	4.085
57	2.367	2.759	3.053	3.350	3.757	4.080
58	2.365	2.756	3.050	3.347	3.753	4.075
59	2.364	2.754	3.047	3.344	3.749	4.071
60	2.362	2.752	3.045	3.340	3.745	4.066

Tabella 1.1 Limiti di controllo per change-point in media

Per quanto riguarda i limiti di controllo per il change-point in varianza è stata usata una procedura parallela.

L'approccio seguito dagli autori Hawkins e Zamba (2005) per avere delle soglie approssimate adatte per il change-point in varianza è invece il seguente:

- per $n = 10, 11, 12, 13, 14, 15$ si usano i limiti della tabella disponibile nel sito web www.stat.umn.edu/hawkins e comunque riportati in tabella 1.2;
- per $n > 15$ si usa la seguente approssimazione:

$$h_{n,\alpha} = \begin{cases} -1.38 - 2.241 \cdot \ln(\alpha) + \frac{1.61 + 0.691 \cdot \ln(\alpha)}{\sqrt{n-9}} & \text{se } 0.001 \leq \alpha < 0.05 \\ 5 + 0.066 \cdot \ln(n-9) & \text{se } \alpha = 0.05 \end{cases}$$

La tabella 1.2 mostra i limiti di controllo ottenuti applicando tale approssimazione sempre per n da 10 a 60 come nel caso precedente.

n	α					
	0.05	0.02	0.01	0.005	0.002	0.001
10	6.374	8.003	9.229	10.451	12.039	13.238
11	5.651	7.328	8.585	9.840	11.489	12.734
12	5.357	7.077	8.373	9.653	11.357	12.631
13	5.228	6.988	8.312	9.634	11.367	12.672
14	5.173	6.960	8.304	9.658	11.423	12.760
15	5.149	6.960	8.323	9.692	11.469	12.828
16	5.128	6.974	8.346	9.718	11.532	12.905
17	5.137	7.000	8.384	9.768	11.598	12.982
18	5.145	7.022	8.416	9.810	11.652	13.046
19	5.152	7.041	8.443	9.845	11.698	13.100
20	5.158	7.057	8.466	9.875	11.738	13.147
21	5.164	7.071	8.486	9.901	11.772	13.187
22	5.169	7.084	8.504	9.925	11.802	13.223
23	5.174	7.095	8.520	9.945	11.830	13.255
24	5.179	7.105	8.534	9.964	11.854	13.284
25	5.183	7.114	8.547	9.981	11.876	13.309
26	5.187	7.122	8.559	9.996	11.896	13.333
27	5.191	7.130	8.570	10.010	11.914	13.355
28	5.194	7.136	8.580	10.023	11.931	13.375
29	5.198	7.142	8.589	10.035	11.947	13.393
30	5.201	7.148	8.597	10.046	11.961	13.410
31	5.204	7.154	8.605	10.056	11.975	13.426
32	5.207	7.159	8.612	10.066	11.987	13.441
33	5.210	7.164	8.619	10.075	11.999	13.455
34	5.212	7.169	8.626	10.083	12.010	13.468
35	5.215	7.172	8.632	10.091	12.021	13.480
36	5.218	7.176	8.638	10.099	12.030	13.492
37	5.220	7.180	8.643	10.106	12.040	13.502
38	5.222	7.184	8.648	10.113	12.048	13.513
39	5.224	7.187	8.653	10.119	12.057	13.523
40	5.227	7.190	8.658	10.125	12.065	13.532
41	5.229	7.194	8.662	10.131	12.072	13.541
42	5.231	7.197	8.667	10.136	12.080	13.550
43	5.233	7.199	8.671	10.142	12.087	13.558
44	5.235	7.202	8.674	10.147	12.093	13.566
45	5.237	7.205	8.679	10.152	12.100	13.573
46	5.239	7.207	8.682	10.156	12.106	13.580
47	5.240	7.210	8.685	10.161	12.111	13.587
48	5.242	7.212	8.688	10.165	12.117	13.594
49	5.243	7.214	8.692	10.169	12.123	13.600

(Segue dalla pagina precedente)

n	α					
	0.05	0.02	0.01	0.005	0.002	0.001
50	5.245	7.216	8.695	10.173	12.128	13.606
51	5.247	7.218	8.698	10.177	12.133	13.612
52	5.248	7.220	8.700	10.181	12.138	13.618
53	5.250	7.222	8.703	10.184	12.142	13.623
54	5.251	7.224	8.706	10.188	12.147	13.629
55	5.253	7.226	8.708	10.191	12.151	13.634
56	5.254	7.227	8.711	10.194	12.155	13.639
57	5.255	7.229	8.713	10.197	12.159	13.644
58	5.257	7.231	8.716	10.201	12.163	13.648
59	5.258	7.232	8.718	10.203	12.167	13.653
60	5.260	7.234	8.720	10.206	12.171	13.657

Tabella 1.2 Limiti di controllo per il change-point in varianza

1.5 DISEGNO DELLE CARTE

Per calcolare $T_{\max,n}$ e $G_{\max,n}$ possono essere trovati due insiemi di valori: il totale delle osservazioni correnti $W_n = \sum_{i=1}^n X_i$ e la somma delle deviazioni al quadrato dalla media corrente $V_{0,n}$. Gli aggiornamenti di questi array sono facili da ottenere:

$$W_{n+1} = W_n + X_{n+1}$$

$$V_{0,n+1} = V_{0,n} + n \cdot (X_{n+1} - W_n/n)^2 / (n+1)$$

Le quantità che servono per calcolare le due statistiche sono le seguenti:

$$\bar{X}_{i,k} = (W_k - W_i) / (k - i)$$

$$V_{i,k} = V_{0,k} - V_{0,i} - i \cdot (k - i) \cdot (\bar{X}_{0,i} - \bar{X}_{i,k})^2 / k$$

In generale, gli aggiornamenti di W_n e $V_{0,n}$ sono veloci, ma trovare il valore k che massimizza $G_{k,n}$, o il valore j che massimizza $|T_{jn}|$, comporta rispettivamente il calcolo di $n-3$ e $n-1$ test statistici, in quanto la statistica $G_{k,n}$ non è definita per $k=1$, $k=n-1$ o $k=n$, mentre la statistica T_{jn} non è definita per $j=n$, come è già stato detto nei paragrafi precedenti.

Il calcolo di queste statistiche per valori grandi di n , potrebbe diventare un limite, in quanto al crescere di n , cresce anche il tempo necessario per ottenere i risultati.

Un rimedio che potrebbe essere utilizzato è quello di restringere la ricerca di k o j in una finestra di ampiezza M che contiene solo le osservazioni più recenti, ossia quando arriva una nuova osservazione, viene cancellata quella più vecchia. In questo modo il k che massimizza $G_{k,n}$, e il j che massimizza $|T_{jn}|$ sono confinati in una

finestra, nel senso che, anche con il crescere di n , il calcolo rimane limitato rispettivamente a $M - 3$ e $M - 1$ test statistici.

Hawkins e Zamba (2005) fanno notare che questo approccio è pienamente efficiente dal punto di vista statistico, in quanto le osservazioni che vengono cancellate quando ne arriva una nuova, sono incluse in tutti i numeri di sintesi. Ciò che viene perso è l'abilità di individuare un valore k che massimizza $G_{k,n}$, o un valore j che massimizza $|T_{jn}|$ a più di M intervalli di tempo nel passato.

1.6 PERFORMANCE

La performance delle carte di controllo è comunemente misurata dall'Average Run Length (ARL). L'ARL rappresenta il numero medio di campioni che devono essere osservati prima che un punto cada al di fuori dei limiti di controllo. Idealmente uno si aspetta un ARL grande quando il processo è in controllo e un ARL piccolo quando il processo è fuori controllo.

Solitamente, quando un processo viene monitorato, vengono usate in modo simultaneo due carte di controllo: una per individuare cambiamenti in media e una per individuare cambiamenti in varianza.

Partendo proprio da questo fatto, ho simulato 500 sequenze di 1500 osservazioni per valutare la combinazione tra il change-point per la media e il change-point per la varianza quando la distribuzione in controllo $N(0,1)$ diventa una $N(\delta, \lambda^2)$. Sia per la formulazione HQK, sia per quella di Hawkins e Zamba (2005) ho usato delle soglie con $\alpha = 0.002$, che darebbero ciascuna un ARL in controllo di 500.

La tabella 1.3 mostra i risultati che ho ottenuto, introducendo un cambiamento in media, in varianza o in entrambe alla 50-esima osservazione.

La riga che corrisponde a $\delta = 0$ mostra come lo schema combinato reagisce solo ai cambiamenti in varianza, mentre la colonna con $\lambda = 1$ mostra la risposta solo ai cambiamenti in media.

δ	λ		
	0.512	1	1.563
0	20.11	279.89	50.08
1	11.72	15.43	11.73
1.5	5.90	6.28	6.26

Tabella 1.3 ARL della combinazione delle carte

Hawkins e Zamba (2005) fanno notare che nell'impostazione del campionamento fisso, il test T_{jn} per le medie e il test G_{kn} per le varianze sono statisticamente indipendenti, e quindi uno potrebbe aspettarsi dalla combinazione delle carte un ARL in controllo vicino a 250. Il fatto che hanno ottenuto degli ARL in controllo un po' più alti, come è confermato anche dal risultato che ho ottenuto e riportato in tabella 1.3, lo giustificano affermando che la ripetizione della verifica d'ipotesi fatta nella formulazione change-point compromette l'indipendenza, creando una relazione positiva tra le run lengths delle due carte.

La tabella 1.4 mostra invece la percentuale dei falsi allarmi ottenuti dalla medesima combinazione, ossia quando la carta per la varianza o la carta per la media sono uscite dai limiti di controllo prima dell'istante 51. Questo rappresenta uno strumento utile per valutare l'efficacia congiunta delle due carte nell'individuare un cambiamento.

δ	λ		
	0.512	1	1.563
0	15.40	15.20	16.40
1	16.20	16.20	20.00
1.5	15.40	14.40	19.80

Tabella 1.4 Percentuale dei falsi allarmi

Capitolo 2

Omnibus EWMA

La maggior parte delle procedure di SPC sono disegnate per individuare cambiamenti in media del processo. In realtà, per una migliore qualità del monitoraggio dello stesso, c'è la necessità di accertare anche altri tipi di cambiamenti, che potrebbero portare la distribuzione delle osservazioni lontana dal valore in controllo (target).

L'omnibus EWMA, è stata proposta da Domangue e Patch (1991), come una carta capace di individuare cambiamenti sia in media sia in varianza.

In questo tipo di carta, a differenza di quanto detto nel capitolo 1 riguardo al change-point, devono essere stabiliti precedentemente il valore della media μ_0 e della deviazione standard σ_0 .

Di conseguenza il processo sarà considerato in uno stato di controllo quando $\mu = \mu_0$ e $\sigma = \sigma_0$, mentre tutte le altre condizioni sono da considerarsi fuori controllo.

Una procedura per il monitoraggio del processo è tipicamente implementata come segue:

- viene calcolata una determinata statistica;
- se la statistica calcolata cade entro determinati limiti, si assume che il processo è in controllo;
- se la statistica calcolata supera i limiti, si conclude che c'è stato un cambio in media o in varianza.

Il processo è in controllo quando $\mu_0 = 0$ e $\sigma_0 = 1$.

2.1 PROPRIETA' DELL'OMNIBUS EWMA

L'enfasi dello studio di Domangue e Patch (1991) si basa sulla seguente statistica:

$$A_i = r \cdot |Z_i|^\alpha + (1-r) \cdot A_{i-1} \text{ per } i = 1, 2, \dots \text{ e } 0 < r \leq 1$$

dove $Z_i = \sqrt{n} \cdot (\bar{X}_i - \mu_0) / \sigma_0$.

\bar{X}_i è la media dell' i -esimo campione, ma visto che questa procedura è disegnata per campioni con $n = 1$, implica che $Z_i = X_i$.

r denota la costante di peso, mentre A_0 il valore iniziale.

Solitamente nell'EWMA tradizionale il valore di $A_0 = 0$. Gli schemi che usano un valore diverso da 0, sono chiamati schemi di Fast Initial Response (FIR), e sono stati discussi da Lucas e Saccucci (1990).

Per semplicità gli autori hanno discusso $|Z_i|^\alpha$ solo con $\alpha = 0.5$ e $\alpha = 2$.

2.2 DISEGNO DELL'OMNIBUS EWMA

Può essere mostrato che per una variabile casuale normale standard Z , il k -esimo momento di $|Z_i|^\alpha$ è dato da:

$$E[|Z|^{\alpha k}] = \frac{2^{\alpha k / 2} \cdot \Gamma[(\alpha k + 1) / 2]}{\sqrt{\pi}}$$

Usando questo risultato e la statistica A_i , per campioni indipendenti provenienti da una distribuzione normale che ha media μ_0 e deviazione standard σ_0 , si ha

$$E[A_i] = \left(\frac{2^\alpha}{\pi} \right)^{1/2} \cdot \Gamma[(\alpha + 1) / 2] \cdot (1 - (1 - r)^i) + A_0 \cdot (1 - r)^i$$

$$\text{e } \text{var}[A_i] = \frac{2^\alpha \cdot r}{(2 - r) \cdot \pi} \cdot \left[\sqrt{\pi} \cdot \Gamma[\alpha + 0.5] - (\Gamma[(\alpha + 1) / 2])^2 \right] \cdot (1 - (1 - r)^{2i})$$

Se $E^*[A_i]$ e $\text{var}^*[A_i]$ denotano i limiti di $E[A_i]$ e $\text{var}[A_i]$, rispettivamente per $i \rightarrow \infty$, si arriva al seguente risultato:

$$E^*[A_i] = \left(\frac{2^\alpha}{\pi} \right)^{1/2} \cdot \Gamma[(\alpha + 1) / 2]$$

$$e \text{ var}^*[A_i] = \frac{2^\alpha \cdot r}{(2-r) \cdot \pi} \cdot \left[\sqrt{\pi} \cdot \Gamma[\alpha + 0.5] - (\Gamma[(\alpha + 1)/2])^2 \right]$$

In linea con quanto proposto da Domangue e Patch (1991), anche nel presente lavoro considero solamente i valori di $\alpha = 0.5$ e $\alpha = 2$.

In particolare, per $\alpha = 0.5$ il valore atteso e la varianza sono

$$E^*[A_i] = \left(\frac{2^{0.5}}{\pi} \right)^{1/2} \cdot \Gamma[(0.5 + 1)/2] \approx 0.8222$$

$$e \text{ var}^*[A_i] = \frac{2^{0.5} \cdot r}{(2-r) \cdot \pi} \cdot \left[\sqrt{\pi} \cdot \Gamma[0.5 + 0.5] - (\Gamma[(0.5 + 1)/2])^2 \right] \approx \frac{0.1219 \cdot r}{2-r}$$

mentre per $\alpha = 2$,

$$E^*[A_i] = \left(\frac{2^2}{\pi} \right)^{1/2} \cdot \Gamma[(2 + 1)/2] = 1$$

$$e \text{ var}^*[A_i] = \frac{2^2 \cdot r}{(2-r) \cdot \pi} \cdot \left[\sqrt{\pi} \cdot \Gamma[2 + 0.5] - (\Gamma[(2 + 1)/2])^2 \right] = \frac{2 \cdot r}{2-r}$$

Un segnale di fuori controllo è dato ogni volta che $A_i \geq E^*[A_i] + L \cdot \text{var}^*[A_i]^{1/2}$, dove L è una quantità specificata dall'utente.

Il valore iniziale A_0 potrebbe essere scelto uguale alla media asintotica di A_i , oppure, ed è la soluzione che ho adottato in questo lavoro, potrebbe essere implementata una caratteristica FIR, assumendo che $A_0 = E^*[A_i] + S \cdot \text{var}^*[A_i]^{1/2}$ per $0 \leq S < L$.

L'utilizzo di questa caratteristica rende all'inizio il processo più sensibile alla situazione di fuori controllo e compensa il fatto che l'esatta varianza della statistica possa essere molto minore della varianza asintotica, specialmente quando r e α sono piccoli.

2.3 PERFORMANCE

La performance di una carta, come è già stato detto nel paragrafo 1.6 del capitolo precedente, è comunemente valutata tramite l'ARL.

Nella tabella 2.1 ho riportato per diversi schemi omnibus EWMA, le ARL nelle quali $\alpha = 0.5$ per diversi valori di μ e σ . Tutti gli schemi sono stati disegnati per un ARL in controllo pari a 250.

A differenza di quanto riportato da Domangue e Patch (1991), che hanno ottenuto i calcoli dell'ARL tramite l'approssimazione delle catene di Markov, i risultati qui riportati sono stati ottenuti tramite la simulazione dalla distribuzione normale di 1000 sequenze, ciascuna di 2000 osservazioni.

Se si mettono a confronto i due metodi, si può concludere che per il calcolo dell'ARL per lo schema qui presentato sono equivalenti. Solo per $r=1$ e $\alpha=0.5$ i risultati sono diversi. Questo potrebbe dipendere dalla scelta di α stesso, visto che tale problema non si verifica per $\alpha = 2$.

Gli schemi dell'omnibus EWMA con $r=1$ sono equivalenti agli schemi Shewhart \bar{X} . In questo caso non si può includere la caratteristica FIR perché inciderebbe solo la prima osservazione. Inoltre per le diverse combinazioni di parametri, gli schemi nei quali $S=1$ hanno sempre un rendimento migliore di quelli nei quali $S=0$. Si può notare comunque, che per lo stesso ARL in controllo uno schema nel quale $S=1$ avrà un ARL un po' più basso rispetto a quando $S=0$.

Di conseguenza gli schemi con $S=0$ dovrebbero essere usati se si considerasse molto importante avere un ARL in controllo più grande quando il processo è in controllo.

		r	0.025	0.5	0.1	0.2	0.5	1	0.025	0.5	0.1	0.2	0.5
		s	0	0	0	0	0	0	1	1	1	1	1
μ	σ	L	1.587	1.944	2.231	2.432	2.536	2.688	1.760	2.027	2.267	2.447	2.540
0	1		251.28	250.53	249.03	252.47	250.44	479.49	251.56	250.31	252.89	253.81	252.62
0	1.2		42.17	44.22	44.95	49.29	53.95	99.36	29.62	33.97	39.65	44.08	52.21
0	1.4		20.7	20.33	20.6	20.64	20.88	38.53	13.59	14.52	16.23	18.35	20.27
0	1.6		13.72	13.95	13.35	12.58	12.35	18.65	8.73	9.04	9.82	10.88	11.28
0.5	1		67.8	70.71	76.85	78.45	92.95	213.45	51.04	59.75	68	77.99	92.82
0.5	1.2		26.66	27.84	28.57	28.81	31.14	65.89	18.57	20.77	23.68	26.08	31
0.5	1.4		17.58	16.63	16.22	15.16	15.93	26.44	10.53	11.55	12.97	13.19	15.07
0.5	1.6		12.69	12.02	11.75	10.77	9.73	16.02	7.47	7.9	8.05	8.33	9.83
1	1		17.02	16.93	16.26	16.81	19.85	59.08	10.83	11.73	12.26	14.43	19.37
1	1.2		13.71	13.67	12.67	12.34	13.04	25.73	8.53	8.8	9.33	9.93	11.97
1	1.4		11.51	10.91	10.26	9.49	9.66	15.69	7.01	6.93	7.46	7.56	7.89
1	1.6		9.84	9.35	8.13	7.7	7.33	9.77	5.62	5.99	5.93	6.1	6.22
1.5	1		8.64	7.72	7.08	6.48	7.05	18.4	4.73	5.01	5.22	5.21	6.03
1.5	1.2		7.94	7.51	6.83	6.22	6.07	11.3	4.63	4.91	4.86	4.99	5.58
1.5	1.4		7.92	7.37	6.55	5.92	5.17	8.19	4.44	4.54	4.5	4.51	4.63
1.5	1.6		7.28	6.67	6.14	5.3	4.75	6.01	4.2	4.33	4.24	4.03	4.27
2	1		5.49	5.06	4.43	3.89	3.82	7.02	3.09	3.18	3.1	2.93	3.11
2	1.2		5.53	5.14	4.48	4.11	3.55	5.49	3.29	3.24	3.1	3.02	3
2	1.4		5.51	5.1	4.53	4.03	3.39	4.66	3.18	3.27	3.15	3.01	2.87
2	1.6		5.48	5.05	4.45	3.79	3.31	4.07	3.15	3.22	3.13	2.98	2.81

Tabella 2.1 ARL omnibus EWMA $|Z|^{0.5}$

Nella tabella 2.2 ho riportato i medesimi calcoli della tabella 2.1 considerando $\alpha = 2$.

		r	0.025	0.5	0.1	0.2	0.5	1	0.025	0.5	0.1	0.2	0.5
		s	0	0	0	0	0	0	1	1	1	1	1
μ	σ	L	1.663	2.192	2.776	3.463	4.558	5.150	1.842	2.280	2.817	3.481	4.563
0	1		251.62	248.17	251.7	250.5	253.63	254.1	254.7	248.01	250.96	252.67	249.43
0	1.2		33.57	37.47	39.12	46.69	53.81	59.27	24.6	29.32	36.44	43.16	54.53
0	1.4		15.02	15.62	17.43	17.69	20.41	25.43	10.6	12.26	13.84	16.94	21.49
0	1.6		10.05	10.04	9.7	10.67	12.15	13.59	6.78	7.46	8.53	9.98	11.88
0.5	1		60.55	65.19	71.29	84.8	100.98	111.7	47.41	56.07	66.8	80.94	98.77
0.5	1.2		20.85	21.35	23.55	25.98	32.41	37.51	15.11	18.57	20.72	25.04	33.88
0.5	1.4		11.98	12.29	12.96	14.22	17.02	19.61	8.56	9.89	11.25	12.87	16.18
0.5	1.6		8.57	8.7	8.43	9.18	10.06	11.16	5.74	6.71	7.16	8.27	9.49
1	1		14.24	14.44	16.12	17.64	24.77	32.73	9.01	11.53	13.53	16.75	24.47
1	1.2		10.13	10.5	10.76	11.61	14.75	16.54	6.8	7.51	8.93	10.37	13.21
1	1.4		7.72	7.64	7.5	8.21	9.22	10.14	5.42	5.87	6.76	7.34	9.04
1	1.6		6.41	6.15	6.13	6.16	6.82	7.52	4.39	4.93	5.28	5.4	6.36
1.5	1		6.32	6.37	6.16	6.53	8.71	11.74	4.07	4.77	5.07	5.93	8.25
1.5	1.2		5.94	5.61	5.27	5.42	6.41	7.89	3.71	4.16	4.38	4.85	6.27
1.5	1.4		5.05	5.08	4.71	4.78	5.33	5.96	3.5	3.78	3.84	4.1	5.1
1.5	1.6		4.66	4.41	4.32	4.3	4.56	4.94	3.2	3.28	3.65	3.83	4.5
2	1		3.77	3.7	3.48	3.49	3.88	5.43	2.55	2.66	2.79	3.02	3.83
2	1.2		3.61	3.52	3.41	3.36	3.62	4.47	2.48	2.59	2.71	2.98	3.53
2	1.4		3.38	3.3	3.17	3.17	3.38	3.74	2.39	2.58	2.63	2.8	3.27
2	1.6		3.3	3.15	3.1	2.98	3.04	3.32	2.35	2.56	2.62	2.72	3.04

Tabella 2.2 ARL omnibus EWMA $|Z|^2$

Capitolo 3

Confronto e risultati

Hawkins e Zamba (2005), nell'articolo in cui descrivono il modello del change-point per la varianza e del quale sono state riportate le caratteristiche fondamentali nel capitolo 1, hanno presentato il confronto di tale modello con la cusum e l'EWMA.

Prima di presentare i risultati gli autori stessi hanno premesso che il confronto tra il loro approccio e i tradizionali metodi di SPC rappresenta una comparazione non soddisfacente data l'assenza di un'alternativa standard.

Per un disegno ottimale della cusum e dell'EWMA hanno quindi deciso di considerare 200 campioni di apprendimento in controllo per tarare le due carte, di simulare 10000 sequenze, cambiando la deviazione standard dalla 201-esima osservazione, e di scartare ogni sequenza nella quale il test del change-point segnalava un allarme prima dell'effettivo cambiamento. Inoltre, a differenza del campionamento fisso adottato per il change-point, per il disegno delle due carte poste a confronto hanno usato campioni di dimensione 5, prendendo in considerazione un ARL di 1000.

Si osservi però che l'uso della cusum e dell'EWMA con campioni di dimensione maggiore di 1 non rappresenta certo la migliore soluzione, come del resto è stato sottolineato anche da Reynolds e Stoumbos (2004b).

Di conseguenza un confronto "più diretto" con il change-point sarebbe quello di proporre uno schema che utilizzi campioni di dimensione pari a 1.

3.1 COMPARAZIONE TRA IL CHANGE-POINT PER LA VARIANZA E L'OMNIBUS EWMA

La scelta del confronto che presento è guidata dalle seguenti linee:

- confrontare il change-point con un modello che è stato costruito per la medesima dimensione campionaria, ossia $n = 1$;

- confrontare il change-point con un modello sensibile sia ai cambiamenti in media sia ai cambiamenti in varianza.

Ho simulato per alcune combinazioni di parametri 500 sequenze, ciascuna di 1500 osservazioni, cambiando la deviazione standard dalla 50-esima osservazione e scartando ogni sequenza nella quale il test del change-point per la varianza o l'omnibus EWMA segnalava un falso allarme, ossia un'uscita dal limite di controllo avvenuta prima dell'istante 51. L'ARL in controllo è fissato pari a 250.

Le figure 3.1 e 3.2 mostrano gli ARL dei due schemi posti a confronto per $\alpha = 0.5$. Ci sono diversi punti meritevoli di attenzione.

Innanzitutto in entrambi i grafici l'omnibus EWMA si comporta meglio del change-point sul range dei valori presi in considerazione, ma sempre meno per larghi shift. Infatti per le due combinazioni di parametri il change-point evidenzia una certa differenza per shift nella deviazione standard minori di 2.5, ma si comporta un po' meglio per larghi cambiamenti in variabilità.

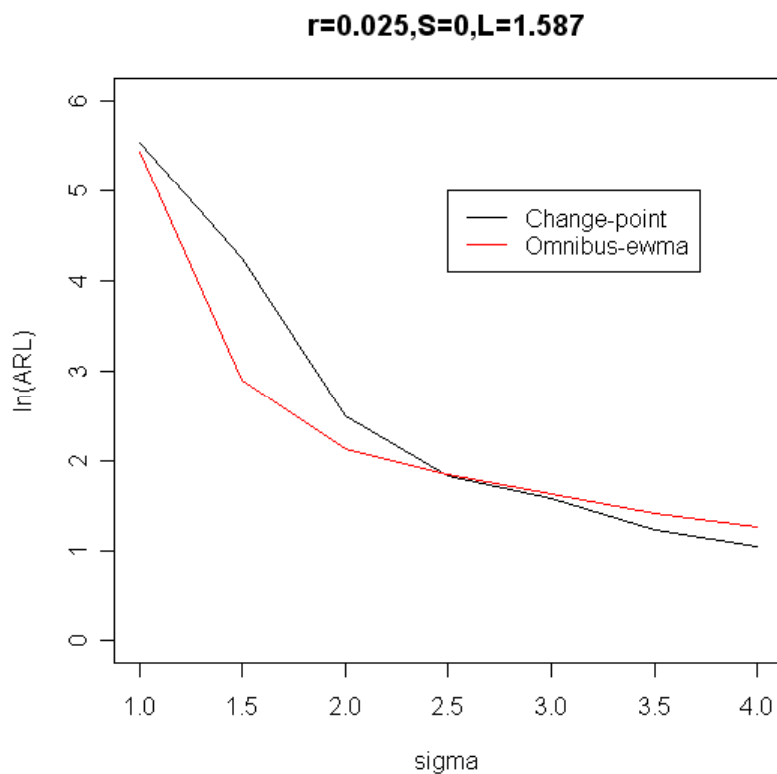


Figura 3.1 Change-point e omnibus EWMA a confronto

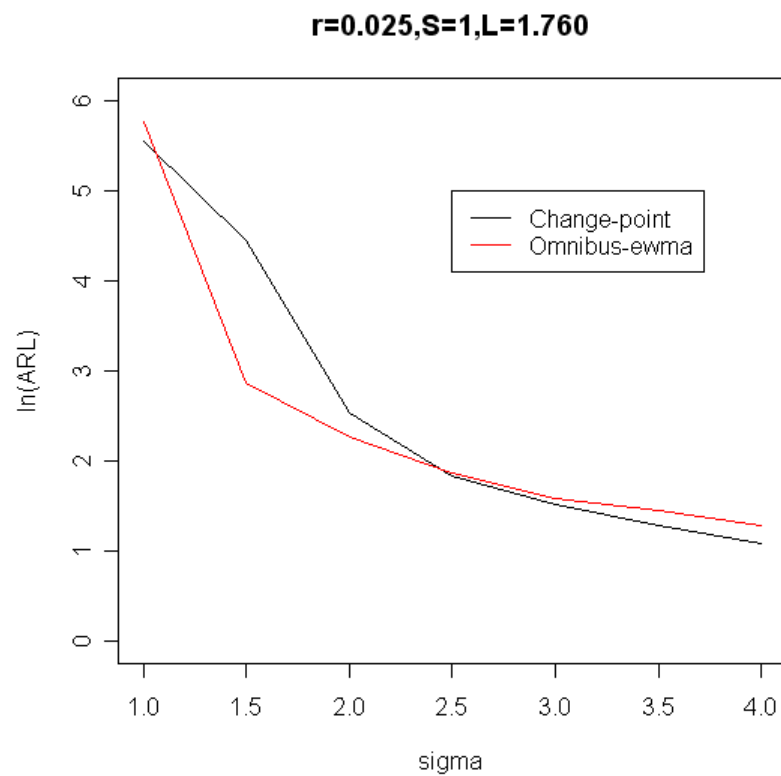


Figura 3.2 Change-point e omnibus EWMA a confronto

Le figure 3.3 e 3.4, fanno vedere un altro confronto del change-point con altre combinazioni di parametri per il disegno dell'omnibus EWMA, ossia dove $\alpha = 2$.

Si può notare che, a differenza di quanto emerso nella rappresentazione precedente, la scelta di $\alpha = 2$ evidenzia un comportamento diverso nell'individuazione di un fuori controllo.

In entrambe le situazioni riportate, per tutto il range dei valori l'omnibus EWMA si comporta meglio del change-point.

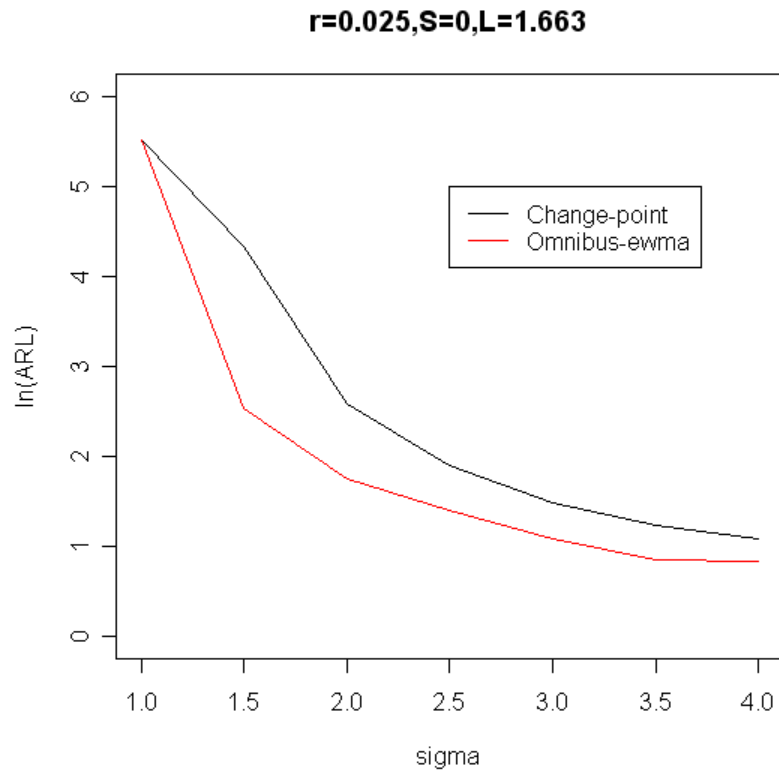


Figura 3.3 Change-point e omnibus EWMA a confronto

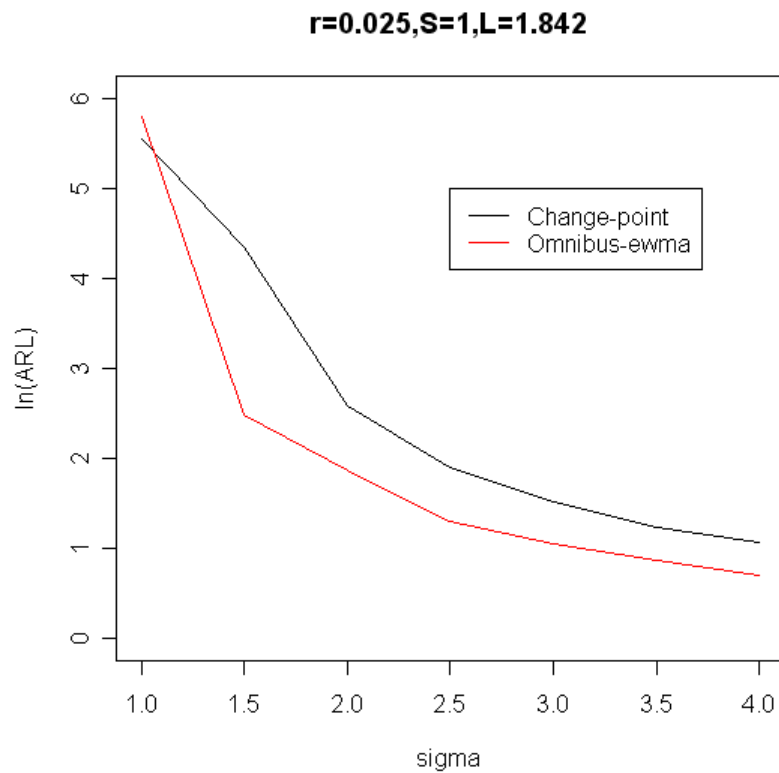


Figura 3.4 Change-point e omnibus EWMA a confronto

L'ARL in controllo dei diversi schemi proposti, non è mai risultato esattamente 250, ma solo approssimativamente. Questo mostra che stimare l'ARL tra le RL ottenute dalle 500 simulazioni potrebbe non essere sufficiente.

Inoltre c'è da segnalare che nella maggior parte dei casi, l'omnibus EWMA è più soggetta ai falsi allarmi rispetto al change-point, come si può osservare nelle tabelle 3.1 e 3.2 che riportano la percentuale dei falsi allarmi ottenuti dal confronto dei due metodi.

		Deviazione standard						
		1	1.5	2	2.5	3	3.5	4
r=0.025 S=0 L=1.587	Change-point	15.6	13.8	19.4	21.4	21.6	29.2	34
	Omnibus EWMA	16.8	19.2	16.8	19.2	20.6	22.4	27
r=0.025 S=1 L=1.760	Change-point	15.4	15.4	17.4	18.2	25.4	29.2	32.6
	Omnibus EWMA	32.2	35	32.4	36.8	37	39.4	38.4

Tabella 3.1 Percentuale dei falsi allarmi ($\alpha = 0.5$)

		Deviazione standard						
		1	1.5	2	2.5	3	3.5	4
r=0.025 S=0 L=1.663	Change-point	16.6	17	14.8	20.4	23.8	26	34.4
	Omnibus EWMA	20	20.4	26.2	32.6	39.4	44.6	47.6
r=0.025 S=1 L=1.842	Change-point	15.6	15.6	20	19.6	24.6	30.2	35.6
	Omnibus EWMA	33.6	35.6	41.8	43.4	50.4	51.4	57.6

Tabella 3.2 Percentuale dei falsi allarmi ($\alpha = 2$)

Conclusioni

Il presente lavoro si propone di mettere a confronto due metodi con caratteristiche alla base ben diverse: la formulazione del change-point per la varianza, che si basa sulla stima adattiva dei parametri e l'omnibus EWMA che si basa sui veri parametri.

I risultati ottenuti dimostrano che nella maggior parte dei casi, in particolare per tutte le combinazioni di parametri presentate dove $\alpha = 2$ e per shift nella deviazione standard minori di 2.5, nel caso dove $\alpha = 0.5$, l'omnibus EWMA si comporta meglio del change-point.

Per shift maggiori di 2.5 il change point per la varianza ha un rendimento migliore in termini di ARL. Nel confronto con un omnibus EWMA, con $\alpha = 0.5$, l'ARL è più bassa, mentre con $\alpha = 2$ l'ARL rimane sempre più alta, ma evidenzia una differenza minore rispetto a quella mostrata per shift minori di 2.5. Di conseguenza si può concludere che la formulazione del change-point, visto anche i risultati ottenuti da Hawkins e Zamba (2005) che hanno effettuato il confronto di questa metodologia con delle estensioni della CUSUM e dell'EWMA, si comporta abbastanza bene per grandi cambiamenti in variabilità.

La decisione di mettere a confronto l'omnibus EWMA, capace di individuare cambiamenti sia in media sia in varianza, con il change-point per la varianza di Hawkins e Zamba (2005) e non con il change-point per la media di HQK (2003) è dettata dal fatto che la metodologia di Hawkins e Zamba (2005) per il punto di cambio in varianza non fa alcuna assunzione sulla costanza o sulla diversità delle medie, mentre quella di HQK per il punto di cambio in media prende in considerazione una varianza costante.

Appendice

I programmi utilizzati per ottenere i risultati riportati nel presente lavoro sono stati scritti utilizzando il linguaggio di programmazione C.

Alcuni di questi programmi, in particolare "change_point.c" e "cp_oe.c" descritti di seguito, sono stati impostati in modo da ottenere 10 run lengths (RL) e la rispettiva media, ossia un ARL parziale, ogni volta che si ripeteva il programma. Questo per evitare di dover aspettare tempi troppo lunghi di esecuzione per calcolare l'ARL cercata in un unico tentativo con il crescere di n .

Man mano che ottenevo questi risultati parziali, gli inserivo in un foglio di calcolo (excel) per arrivare all'ARL definitiva, come è rappresentato nell'esempio sottostante:

		↓ Colonna 1														
← riga 1 →		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		1	2	3	4	5	6	7	8	9	10	n	tot	ARL	f.a	
2	1	10	10	10	10	10	10	10	10				7	70	10	3
3	2	1	1	1	1	1	1	1	1	1			8	8	1	2
4	3	5	5	5	5	5	5	5	5	5	5		10	50	5	0
...	...															
...	...															
...	...															
51	50															
52													L52	M52	N52	O52

Nella riga 1 sono riportati i numeri da 1 a 10 che indicano il numero di RL che si dovrebbero ottenere ogni volta che si ripete il programma.

n indica il numero di RL valide, tot la somma delle RL per riga, ARL il rapporto tra tot e n e f.a il numero di falsi allarmi in ogni ripetizione.

Nella colonna 1 è riportato il numero di volte che il programma è stato ripetuto, ossia se abbiamo ripetuto 3 volte il programma, significa che le RL calcolate sono state 30 (somma tra tutti gli n e i falsi allarmi).

- L52 rappresenta la somma di tutti gli n da L2 a L51;
- M52 rappresenta la somma di tutti i tot da M2 a M51;

- N52 rappresenta il rapporto tra M52 e L52, ossia l'ARL cercata;
- O52 rappresenta la somma di tutti gli f.a da O2 a O51.

Il codice dei programmi riportato nelle pagine seguenti può essere migliorato con semplici modifiche in base alle ricerche di studio che un utente potrebbe intraprendere, inserendo anche opportuni controlli sui parametri.

change_point.c

"change_point.c" è il codice che si riferisce al calcolo delle ARL della combinazione delle carte change-point per la media e change-point per la varianza, i cui risultati, ottenuti tramite simulazione, sono riportati nelle tabelle 1.3 e 1.4.

```
#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include <float.h>

float lim_var[2000];
float lim_media[2000];
double glr_var[2000];
double glr_media[2000];
double w[2000];
double v[2000];

#ifndef M_PI
#define M_PI          3.141592653589793238462643383280    /* pi */
#endif

static unsigned int I1, I2;

void set_seed(unsigned int i1, unsigned int i2)
{
    I1 = i1; I2 = i2;
}

double unif_rand(void)
{
    I1= 36969*(I1 & 0177777) + (I1>>16);
    I2= 18000*(I2 & 0177777) + (I2>>16);
    return ((I1 << 16)^(I2 & 0177777)) * 2.328306437080797e-10; /* in [0,1) */
}

static double BM_norm_keep=0.0;
double norm_rand(void)
{
    double s,theta, R;
    if(BM_norm_keep != 0.0) { /* An exact test is intentional */
        s = BM_norm_keep;
        BM_norm_keep = 0.0;
        return s;
    } else {
        theta = 2 * M_PI * unif_rand();
        R = sqrt(-2 * log(unif_rand())) + 10*DBL_MIN; /* ensure non-zero */
        BM_norm_keep = R * sin(theta);
        return R * cos(theta);
    }
}

int lim_v(int n)
```

```

{
    //calcola i limiti solo per alfa=0.004
    //con semplici modifiche si può estendere anche ad altri valori di alfa
    //(vedi capitolo 1)
    int i,h;
    float alfa=0.002;

    lim_var[1]=12.039;
    lim_var[2]=11.489;
    lim_var[3]=11.357;
    lim_var[4]=11.367;
    lim_var[5]=11.423;
    lim_var[6]=11.469;
    h=6;

    for (i=16;i<=n;i++)
    {
        lim_var[i-9]=-1.38-2.241*log(alfa)+(1.61+0.691*log(alfa))/sqrt(i-9);
        h++;
    }
    return h;
}

int lim_m(int n)
{
    int i,h;
    float alfa=0.002;

    lim_media[1]=6.340;
    h=1;
    for (i=11;i<=n;i++)
    {
        lim_media[i-9]=6.340*(0.677+0.019*log(alfa)+(1-0.115*log(alfa))/(i-6));
        h++;
    }
    return h;
}

double somma (double x[2000],int i)
{
    int j;
    double s=0;

    for (j=1;j<=i;j++)
    {
        s=s+x[j];
    }
    return s;
}

double diff_q (double x[2000], double m,int i)
{
    int j;
    double d=0;

    for (j=1;j<=i;j++)
    {
        d=d+pow((x[j]-m),2);
    }
    return d;
}

int glr_v(double x[2000], int n)
{
    double gkn[2000];
    double m,xik,xkn,vik,vkn,v1,v2,var;
    double g,b,max;
    int i,j,k,h=0;
    int p;

    for (i=1;i<=n;i++)
    {
        w[i]=somma(x,i);
        m=w[i]/i;
        v[i]=diff_q(x,m,i);
    }
}

```

```

    }
    for (j=10;j<=n;j++)
    {
        p=0;
        for (k=2;k<=(j-2);k++)
        {
            xik=w[k]/k;
            xkn=(w[j]-w[k])/(j-k);
            vik=v[k];
            vkn=v[j]-v[k]-(k*(j-k)*pow((xik-xkn),2))/j;
            v1=vik/(k-1);
            v2=vkn/(j-k-1);
            var=(vik+vkn)/(j-2);

            g=(k-1)*log(var/v1)+(j-k-1)*log(var/v2);
            b=1+(pow((k-1),-1)+pow((j-k-1),-1)-pow((j-2),-1))/3;
            gkn[k-1]=g/b;
            p++;
        }
        max=gkn[1];
        for (i=2;i<=p;i++)
        {
            if (gkn[i]>max)
            {
                max=gkn[i];
            }
        }
        glr_var[j-9]=max;
        h++;
    }

    return h;
}

int glr_m(double x[2000], int n)
{
    double t[2000],a,b,c;
    double xik,xkn,vik,vkn,v1,v2,var;
    double max;
    int i,j,k,h=0;
    int p;

    for (j=10;j<=n;j++)
    {
        p=0;
        for (k=1;k<=(j-1);k++)
        {
            xik=w[k]/k;
            xkn=(w[j]-w[k])/(j-k);
            vik=v[k];
            vkn=v[j]-v[k]-(k*(j-k)*pow((xik-xkn),2))/j;
            var=(vik+vkn)/(j-2);

            a=k*(j-k);
            b=sqrt(a/j);
            c=((xik-xkn)/sqrt(var));
            t[k]=fabs(b*c);
            p++;
        }
        max=t[1];
        for (i=2;i<=p;i++)
        {
            if (t[i]>max)
            {
                max=t[i];
            }
        }
        glr_media[j-9]=max;
        h++;
    }

    return h;
}

```

```

float calcola_media(int runl[],int y)
{
    int i,somma=0,h=0;

    if (y!=0)
    {
        for (i=0;i<y;i++)
        {
            somma=somma+runl[i];
            h=h+1;
        }

        return (float)somma/h;
    }
    else
    {
        printf("Nessun allarme segnalato\n");
        return 0;
    }
}

int main()
{
    int n=1500; //lunghezza di ogni sequenza
    int j,v,m,r,num=10; //num=10 per calcolare 10 rl alla volta
    int p;
    int valori_v,valori_m,minimo;
    int e,fe,exit[100];
    float arl;
    double x[2000];
    float mu,sd;
    int inizio,continua;

    //per cambiare il "punto di partenza" (della simulazione)
    srand(time(0));
    set_seed(rand(),rand());

    inizio=1;
    while (inizio>0)
    {
        printf("Valore di mu: ");
        scanf("%f",&mu);
        printf("Valore di sigma: ");
        scanf("%f",&sd);

        v=lim_v(n);
        m=lim_m(n);

        continua=1;
        while (continua>0)
        {
            if (m!=v)
            {
                printf("Limiti errati\n");
                return 0;
            }

            e=0;
            fe=0;
            for (r=1;r<=num;r++)
            {
                for (p=1;p<=49;p++)
                {
                    x[p]=norm_rand(); //simulazione da N(0,1)
                }

                for (p=50;p<=n;p++)
                {
                    x[p]=mu+sd*norm_rand(); //simulazione da N(mu,sd)
                }

                valori_v=glr_v(x,n);
                valori_m=glr_m(x,n);

                for (j=1;j<=m;j++)

```

```

        {
            if ((glr_var[j]>lim_var[j]) ||
                (glr_media[j]>lim_media[j]))
            {
                minimo=j+9;
                break;
            }
        }
    if (minimo>0)
    {
        if (minimo>50)
        {
            exit[e]=minimo-50;
            e=e+1;
        }
        else
        {
            fe++;
        }
    }
}

printf("Valori RL change-point combinato: ");
for (j=0;j<e;j++)
{
    printf(" %d",exit[j]);
    printf(" ");
}
printf("\n");

arl=calcola_media(exit,e);
printf("ARL approssimativo: %.2f\n",arl);
printf("Numero di falsi allarmi: %d\n",fe);
printf("MU: %f e sigma %f\n",mu,sd);

printf("\n");
printf("RICALCOLARE L'ARL CON GLI STESSI PARAMETRI?\n");
printf("Premi 1 per continuare ----- Premi 0 per cambiare
    parametri: ");
scanf("%d",&continua);
}

printf("*****\n");
printf("CAMBIARE PARAMETRI E CALCOLARE L'ARL?\n");
printf("Premi 1 per continuare ----- Premi 0 per terminare: ");
scanf("%d",&inizio);
}
}

```

cp_oe.c

"cp_oe.c" è il codice che si riferisce al confronto tra il change-point per la varianza e l'omnibus EWMA i cui calcoli sono riportati nel capitolo 3.

```

#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include <float.h>

float lim_var[2010];
double glr_var[2010];
double w[2010];
double v[2010];
float ai[2010];

#ifdef M_PI
#define M_PI 3.141592653589793238462643383280 /* pi */
#endif

```

```

static unsigned int I1, I2;

void set_seed(unsigned int i1, unsigned int i2)
{
    I1 = i1; I2 = i2;
}

double unif_rand(void)
{
    I1= 36969*(I1 & 0177777) + (I1>>16);
    I2= 18000*(I2 & 0177777) + (I2>>16);
    return ((I1 << 16)^(I2 & 0177777)) * 2.328306437080797e-10; /* in [0,1) */
}

static double BM_norm_keep=0.0;
double norm_rand(void)
{
    double s,theta, R;
    if(BM_norm_keep != 0.0) { /* An exact test is intentional */
        s = BM_norm_keep;
        BM_norm_keep = 0.0;
        return s;
    } else {
        theta = 2 * M_PI * unif_rand();
        R = sqrt(-2 * log(unif_rand())) + 10*DBL_MIN; /* ensure non-zero */
        BM_norm_keep = R * sin(theta);
        return R * cos(theta);
    }
}

int lim_v(int n)
{
    //calcola i limiti solo per alfa=0.004
    //con semplici modifiche si può estendere anche ad altri valori di alfa
    //(vedi capitolo 1)
    int i,h;
    float alfa=0.004;

    h=0;
    for (i=10;i<=n;i++)
    {
        lim_var[i-9]=-1.38-2.241*log(alfa)+(1.61+0.691*log(alfa))/sqrt(i-9);
        h++;
    }
    return h;
}

double somma (double x[2010],int i)
{
    int j;
    double s=0;

    for (j=1;j<=i;j++)
    {
        s=s+x[j];
    }
    return s;
}

double diff_q (double x[2010], double m,int i)
{
    int j;
    double d=0;

    for (j=1;j<=i;j++)
    {
        d=d+pow((x[j]-m),2);
    }
    return d;
}

int glr_v(double x[2010], int n)
{

```

```

double gkn[2010];
double m,xik,xkn,vik,vkn,v1,v2,var;
double g,b,max;
int i,j,k,h=0;
int p;

for (i=1;i<=n;i++)
{
    w[i]=somma(x,i);
    m=w[i]/i;
    v[i]=diff_q(x,m,i);
}

for (j=10;j<=n;j++)
{
    p=0;
    for (k=2;k<=(j-2);k++)
    {
        xik=w[k]/k;
        xkn=(w[j]-w[k])/(j-k);
        vik=v[k];
        vkn=v[j]-v[k]-(k*(j-k)*pow((xik-xkn),2))/j;
        v1=vik/(k-1);
        v2=vkn/(j-k-1);
        var=(vik+vkn)/(j-2);

        g=(k-1)*log(var/v1)+(j-k-1)*log(var/v2);
        b=1+(pow((k-1),-1)+pow((j-k-1),-1)-pow((j-2),-1))/3;
        gkn[k-1]=g/b;
        p++;
    }
    max=gkn[1];
    for (i=2;i<=p;i++)
    {
        if (gkn[i]>max)
        {
            max=gkn[i];
        }
    }
    glr_var[j-9]=max;
    h++;
}

return h;
}

float calcola_media(int runl[],int y)
{
    int i,somma=0,h=0;

    if (y!=0)
    {
        for (i=0;i<y;i++)
        {
            somma=somma+runl[i];
            h=h+1;
        }

        return (float)somma/h;
    }
    else
    {
        printf("Nessun allarme segnalato\n");
        return 0;
    }
}

void stampa(int runl[100],int e)
{
    int i;

    for (i=0;i<e;i++)
    {
        printf(" %d",runl[i]);
    }
}

```



```

        printf("\n");
    }

int main()
{
    int n=1500; //lunghezza di ogni sequenza
    int j,v,u,num=10; //num=10 per calcolare 10 rl alla volta
    int p,s,k,min_e;
    int inizio,continua;
    int q,qe,rl_ewma[100];
    int valori_v,minimo;
    int e,fe,rl_glr[100];
    float arl_cp,arl_oe;
    float alfa,e_ai,var_ai;
    float lim,a,r,L;
    double x[2010],z;
    float mu,sd;

    //per cambiare il "punto di partenza" (della simulazione)
    srand(time(0));
    set_seed(rand(),rand());

    inizio=1;
    while (inizio>0)
    {
        printf("Valore di mu: ");
        scanf("%f",&mu);
        printf("Valore di sigma: ");
        scanf("%f",&sd);

        printf("Il valore di alfa per statistica ai deve essere 0.5 o 2\n");
        printf("Inserisci uno dei due valori: ");
        scanf("%f",&alfa);

        printf("Valore di r: ");
        scanf("%f",&r);
        printf("Valore di s: ");
        scanf("%d",&s);
        printf("Valore di L: ");
        scanf("%f",&L);
        fflush(stdin); //svuotare il buffer

        if (alfa==2)
        {
            e_ai=1;
            var_ai=(2*r)/(2-r);
            lim=e_ai+L*sqrt(var_ai);
            a=e_ai+s*sqrt(var_ai); //utilizzo del FIR
        }

        if (alfa==0.5)
        {
            e_ai=0.8222;
            var_ai=(0.1219*r)/(2-r);
            lim=e_ai+L*sqrt(var_ai);
            a=e_ai+s*sqrt(var_ai); //utilizzo del FIR
        }

        v=lim_v(n);

        continua=1;
        while (continua>0)
        {
            e=0;
            fe=0;
            q=0;
            qe=0;
            for (u=1;u<=num;u++)
            {
                for (p=1;p<=49;p++)
                {
                    x[p]=norm_rand(); //simulazione da N(0,1)
                }

                for (p=50;p<=n;p++)

```

```

    {
        x[p]=mu+sd*norm_rand();//simulazione da N(mu,sigma)
    }

    valori_v=glr_v(x,n);

    minimo=0;
    for (j=1;j<=v;j++)
    {
        if (glr_var[j]>lim_var[j])
        {
            minimo=j+9;
            break;
        }
    }
    if (minimo>0)
    {
        if (minimo>50)
        {
            rl_glr[e]=minimo-50;
            e=e+1;
        }
        else
        {
            fe++;//conta falsi allarmi del change-point
        }
    }
    }

    z=fabs(x[1]);
    ai[1]=r*pow(z,alfa)+(1-r)*a; //ai_1

    for (j=2;j<=n;j++)
    {
        z=fabs(x[j]);
        ai[j]=r*pow(z,alfa)+(1-r)*ai[j-1];
    }

    min_e=0;
    for (k=1;k<=n;k++)
    {
        if (ai[k]>=lim)
        {
            min_e=k;
            break;
        }
    }
    if (min_e>0)
    {
        if (min_e>50)
        {
            rl_ewma[q]=min_e-50;
            q=q+1;
        }
        else
        {
            qe++;//conta falsi allarmi dell'omnibus-ewma
        }
    }
    }

    arl_cp=calcola_media(rl_glr,e);
    arl_oe=calcola_media(rl_ewma,q);
    printf("\n");
    printf("Valori RL change-point: ");
    stampa(rl_glr,e);
    printf("ARL approssimativo change-point: %.2f\n",arl_cp);
    printf("Numero di falsi allarmi: %d\n",fe);

    printf("*****\n");
    printf("Valori RL omnibus-ewma: ");
    stampa(rl_ewma,q);
    printf("ARL approssimativo omnibus-ewma: %.2f\n",arl_oe);
    printf("Numero di falsi allarmi: %d\n",qe);

    printf("\n");

```

```

        printf("RICALCOLARE L'ARL CON GLI STESSI PARAMETRI?\n");
        printf("Premi 1 per continuare ----- Premi 0 per cambiare
        parametri: ");
        scanf("%d",&continua);
    }

    printf("*****\n");
    printf("CAMBIARE PARAMETRI E CALCOLARE L'ARL?\n");
    printf("Premi 1 per continuare ----- Premi 0 per terminare: ");
    scanf("%d",&inizio);
}
}
}

```

oe.c

“oe.c” è il codice che si riferisce al calcolo dell’ARL per l’omnibus EWMA i cui risultati sono riportati nelle tabelle 2.1 e 2.2.

```

#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include <float.h>

float ai[2010];

#ifdef M_PI
#define M_PI          3.141592653589793238462643383280    /* pi */
#endif

static unsigned int I1, I2;

void set_seed(unsigned int i1, unsigned int i2)
{
    I1 = i1; I2 = i2;
}

double unif_rand(void)
{
    I1= 36969*(I1 & 0177777) + (I1>>16);
    I2= 18000*(I2 & 0177777) + (I2>>16);
    return ((I1 << 16)^(I2 & 0177777)) * 2.328306437080797e-10; /* in [0,1) */
}

static double BM_norm_keep=0.0;
double norm_rand(void)
{
    double s,theta, R;
    if(BM_norm_keep != 0.0) { /* An exact test is intentional */
        s = BM_norm_keep;
        BM_norm_keep = 0.0;
        return s;
    } else {
        theta = 2 * M_PI * unif_rand();
        R = sqrt(-2 * log(unif_rand())) + 10*DBL_MIN; /* ensure non-zero */
        BM_norm_keep = R * sin(theta);
        return R * cos(theta);
    }
}

float calcola_media(int rl_min[2010],int q)
{
    int i,somma=0,h=0;

    for (i=0;i<q;i++)
    {
        somma=somma+rl_min[i];
        h=h+1;
    }
}

```

```

        return (float)somma/h;
    }

int main()
{
    int n=2000; //lunghezza di ogni sequenza
    float alfa,e_ai,var_ai;
    float lim,a,r,L;
    float arl,mu,sd;
    int s,j,i,k,num=1000; //num=1000 numero delle RL
    int inizio,continua;
    int q,rl_min[2010];
    double x[2010],z;

    //per cambiare il "punto di partenza" (della simulazione)
    srand(time(0));
    set_seed(rand(),rand());

    inizio=1;
    while (inizio>0)
    {
        printf("Valore di alfa: ");
        scanf("%f",&alfa);
        printf("Valore di r: ");
        scanf("%f",&r);
        printf("Valore di s: ");
        scanf("%d",&s);
        printf("Valore di L: ");
        scanf("%f",&L);

        if (alfa==2)
        {
            e_ai=1;
            var_ai=(2*r)/(2-r);
            lim=e_ai+L*sqrt(var_ai);
            a=e_ai+s*sqrt(var_ai); //utilizzo del FIR
        }

        if (alfa==0.5)
        {
            e_ai=0.8222;
            var_ai=(0.1219*r)/(2-r);
            lim=e_ai+L*sqrt(var_ai);
            a=e_ai+s*sqrt(var_ai); //utilizzo del FIR
        }

        continua=1;
        while (continua>0)
        {
            printf("Valore di mu: ");
            scanf("%f",&mu);
            printf("Valore di sigma: ");
            scanf("%f",&sd);

            q=0;
            for (i=1;i<=num;i++)
            {
                for (j=1;j<=n;j++)
                {
                    x[j]=mu+sd*norm_rand();//simulazione da N(mu,sigma)
                }

                z=fabs(x[1]);
                ai[1]=r*pow(z,alfa)+(1-r)*a; //ai_1

                for (j=2;j<=n;j++)
                {
                    z=fabs(x[j]);
                    ai[j]=r*pow(z,alfa)+(1-r)*ai[j-1];
                }

                for (k=1;k<=n;k++)
                {
                    if (ai[k]>=lim)

```

```

        {
            rl_min[q]=k;
            q=q+1;
            break;
        }
    }

    printf("\n");
    arl=calcola_media(rl_min,q);
    printf("ARL approssimativo: %.2f\n",arl);
    printf("-----\n");
    printf("RICALCOLARE L'ARL CON GLI STESSI PARAMETRI?\n");
    printf("Premi 1 per continuare ----- Premi 0 per cambiare
           parametri: ");
    scanf("%d",&continua);
}
printf("-----\n");
printf("CAMBIARE PARAMETRI E CALCOLARE L'ARL?\n");
printf("Premi 1 per continuare ----- Premi 0 per terminare: ");
scanf("%d",&inizio);
}
}

```


Bibliografia

- [1] Bartlett M. S. (1937). "*Property of Sufficiency and Statistical Tests*", *Proceedings of the Royal Society of London, Series A* 160, 268-282
- [2] Bartlett M. S. e Kendall D. G. (1946). "*The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation*", *Supplement to the Journal of the Royal Society of London* 8, 128-138
- [3] Bartlett M. S. (1955). "*Approximate Confidence Intervals: A bias correction*", *Biometrika* 42, 201-204
- [4] Crowder S. V. e Hamilton (1992). "*An EWMA for Monitoring a Process Standard Deviation*", *Journal of Quality Technology* 24, 12-21
- [5] Domangue R. e Patch S. C. (1991). "*Some Omnibus Exponentially Weighted Moving Average Statistical Process Monitorino Schemes*", *Technometrics* 33, 299-313
- [6] Hawkins D. M. e Olwell D. H. (1998). "*Cumulative Sum Charts and Charting for Quality Improvement*", Springer Verlag, New York
- [7] Hawkins D. M., Qui P. e Kang (2003). "*The Change Point Model for Statistical Process Control*", *Journal of Quality Technology* 35, 355-365
- [8] Hawkins D. M. e Zamba K. D. (2005). "*A Change Point Model for a Shift in Variance*", *Journal of Quality Technology* 37, 21-31
- [9] Lucas J. M. e Saccucci M. S. (1990). "*Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements*", *Technometrics* 32, 1-12
- [10] MacGregor J. F. e Harris T. J. (1993). "*The exponentially weighted moving variance*", *Journal of Quality Technology*, 25
- [11] Margavio T. M., Coneriy M. D., Woodall W. H. e Drake L. G. (1995). "*Alarm Rates for Quality Control Charts*", *Statistic and Probability Letters* 24, 219-224
- [12] Montgomery Douglas C. (2000). "*Controllo statistico della qualità*", McGraw-Hill
- [13] Reynolds M. R. Jr. e Stoumbos Z. G. (2004b). "*Should Observations Be Grouped for Effective Process Monitoring?*", to appear in *Journal of Quality Technology*