

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



“Detección del Punto de Cambio Mediante el
Uso Secuencial de Cartas de Control y
Estimadores de Máxima Verosimilitud”

Por

ÁNGEL SALVADOR PÉREZ BLANCO

Como requisito parcial para obtener el grado de

DOCTOR EN CIENCIAS CON ORIENTACION EN
MATEMÁTICAS

Octubre 2013

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS
DOCTORADO EN CIENCIAS CON ORIENTACIÓN EN MATEMÁTICAS

Los miembros del comité de tesis recomendamos que la tesis

“Detección del Punto de Cambio Mediante el Uso Secuencial de Cartas de Control y
Estimadores de Máxima Verosimilitud”

realizada por el estudiante Ángel Salvador Pérez Blanco sea aceptada para su defensa como
opción al grado de Doctor en Ciencias con Orientación en Matemáticas.

El Comité de Tesis

Director/Asesor

Ph. D. Álvaro Eduardo Cordero Franco

Co-asesor

Ph. D. Víctor Gustavo Tercero Gómez

Revisor

Ph. D. José Fernando Camacho Vallejo

Revisor

Ph. D. Alberto Abelardo Hernández Luna

Revisor

Ph. D. José Guadalupe Ríos Alejandro

Vo. Bo.

Dr. José Luis Comparan Elizondo

División de Estudios de Posgrado

San Nicolás de los Garza, Nuevo León Octubre de 2013

RESUMEN

Ángel Salvador Pérez Blanco

Fecha de graduación: Septiembre, 2013

Título del estudio: "Detección del Punto de Cambio Mediante el Uso Secuencial de Cartas de Control y Estimadores de Máxima Verosimilitud"

Número de páginas 112

Candidato para el grado de Doctor en Ciencias con Orientación en Matemáticas

Área de estudio: Control estadístico de la calidad

Propósito y método de estudio: El objetivo final de esta tesis es presentar una metodología que apoye a hacer más efectivo y eficiente la inspección y la detección de cambios estructurales en el control estadístico de procesos. El control estadístico de procesos se utiliza en el monitoreo de los sistemas, ya sea de producción, servicio y/o naturales y se ha ido fortaleciendo mediante herramientas que ayudan a la detección de las causas que provocan una perturbación en el proceso, conocidas como causas asignables, y cuyo propósito es que una vez detectadas, se realicen acciones para su eliminación, todo esto considerando la factibilidad de la implementación de dichas medidas. La detección de cambios sostenidos auxilia en la implementación de las acciones correctivas y/o preventivas que redundarán en beneficios económicos, de tiempo y de esfuerzo que requieren los sistemas para mantenerse funcionando, bajo las mejores condiciones posibles para generar el mejor producto, en el sentido más amplio, para el que fueron diseñados.

Bajo este esquema general, se ha realizado un análisis global del tema y se ha centrado en el análisis del punto de cambio (CPA por sus siglas en inglés) haciendo una revisión literaria general del control estadístico de procesos (SPC por sus siglas en inglés) y particularizando en esta herramienta. También se propone una metodología consistente en aplicar de manera secuencial las cartas de control estadístico y el uso de estimadores de máxima verosimilitud. La herramienta propuesta ha sido desarrollada y sometida a pruebas considerando diferentes escenarios mediante el uso de simulaciones Monte Carlo con el fin de evaluar su desempeño frente a otros estimadores reportados en la literatura.

Contribuciones y conclusiones: La metodología propuesta es un enfoque de combinación de dos áreas del SPC, cartas de control y estimadores de máxima verosimilitud. Se utilizan las cartas de control para la detección de cambios en el proceso, y estimadores de máxima verosimilitud para estimar el punto de inicio de los mismos. El trabajo se realizó asumiendo series de observaciones independientes con distribución Normal y Gamma con parámetros iniciales y posteriores al cambio desconocidos, para el caso de la Normal, e iniciales conocidos con posteriores desconocidos para el caso de la Gamma

La razón de seleccionar la función de distribución Normal para la realización de la presente investigación es por su recurrencia y diversidad en la explicación de fenómenos que ocurren en la realidad y en los procesos de producción. Por otra parte, se eligió la función de distribución Gamma por ser representativa de los fenómenos cuya característica principal es que la ocurrencia de eventos tiene una periodicidad por unidad de tiempo, como el caso de llegadas de llamadas a un conmutador.

Las simulaciones, bajo diferentes escenarios, y en comparación con métodos ya utilizados, muestran resultados aceptables en términos de precisión y exactitud para la detección del punto de cambio, lo cual hace promisoria la investigación en esta línea y en las variantes futuras que se puedan considerar.

FIRMA DEL ASESOR: _____

PREFACIO

Este trabajo es el resultado de la investigación realizada durante el periodo Agosto 2010 a Julio 2013 dentro del programa “Doctorado en Ciencias con Orientación en Matemáticas” y se presenta como una tesis para obtener el grado de Doctor en Ciencias con Orientación en Matemáticas. Los resultados han sido publicados y presentados por los siguientes autores:

Publicaciones:

- 1 Tercero-Gómez, Víctor G.; Cordero-Franco, Álvaro E.; Hernández-Luna, Alberto A.; Pérez-Blanco, Ángel S.; “A Self-Starting CUSUM Combined with a Maximum Likelihood Estimator for the Time of a Detected Shift in the Process Mean”; *Quality and Reliability Engineering International*; publicado en línea: 24-May-2013; DOI: 10.1202/qre.1511; in press.
- 2 Pérez-Blanco, Ángel S.; Cordero-Franco, Álvaro E.; Tercero-Gómez, Víctor G.; Chávez-Valdez, María A.; “Análisis de punto de cambio dentro del control estadístico de procesos”; *Celerinet FCFM-UANL*, Vol 2 Julio-Diciembre 2013; en prensa.
- 3 Pérez-Blanco, Ángel S.; Cordero-Franco, Álvaro E.; Tercero-Gómez, Víctor G.; “Change point Estimation After the Signal of a Gamma Control Chart”; *ISERC 2013*; San Juan Puerto Rico; 2013

Presentaciones:

- 1 Pérez Blanco, Ángel S.; “Punto de Cambio: Estado Actual y tendencias”; *2do. Foro de Divulgación Científica y Tecnológica*, Monterrey, 2012.

AGRADECIMIENTOS

Primeramente a Dios por seguir teniéndome como uno de sus consentidos.

A mi familia: Lula el apoyo y compañerismo hecho realidad y por ser siempre una firme entusiasta de los nuevos retos. A Emiliano y Valeria que jamás se cansan de animarnos a seguir adelante.

A mis padres y hermana; Ángel[†], Zoyla y Aracely porque nunca me pierden de vista y siempre me han procurado lo mejor.

Al Ph. D. Eduardo Cordero por su mente siempre abierta a nuevas ideas y al Ph. D. Víctor Tercero porque siempre aportó la claridad en las metas. De ambos aprendí muchas cualidades de lo que es un investigador y quienes, a pesar de su juventud, son modelo a seguir.

Al Ph. D. José Fernando Camacho Vallejo, al Ph. D. José Guadalupe Ríos Alejandro y al Ph. D. Alberto Abelardo Hernández Luna por haberse tomado el tiempo de atender la revisión de mi tesis, pero sobre todo por sus valiosos comentarios y sugerencias que hicieron durante la pre-defensa.

A la Maestra Paty, la Maestra Azucena y al Maestro Rogelio, como los conocemos los estudiantes, por su apoyo incondicional y las facilidades que me dieron para llegar al día de hoy.

A mis chicas de escolar: Evelyn, Melly, Rosy, Carmen, Lalis y Gaby cuyo apoyo es invaluable.

A Rosy Gámez, Magdalena Sánchez, Martín Guerrero, Agustín, Lilia Guadalupe, Jesús Suarez, Francisco, Nora, la lista es enorme, amigos de la FCFM.

A mis compañeros de toda la vida: Aurora, Fernando y Sergio, “LA BANDA”, porque juntos hemos practicado el significado de incondicional.

TABLA DE CONTENIDO

TABLA DE CONTENIDO	vii
RESUMEN	x
LISTA DE TABLAS.....	xii
LISTA DE FIGURAS.....	xv
CAPÍTULO 1. INTRODUCCIÓN	1
1.1 Historia	1
1.2 Especificación del problema.....	3
1.3 Preguntas que Genera la Investigación.....	4
1.4 Hipótesis General	5
1.5 Formato de la Investigación	6
1.6 Propósito de la Investigación	7
1.7 Objetivo de la Investigación	7
1.8 Alcances y Limitaciones de la Investigación	8
1.9 Importancia del Presente Estudio	9
1.10 Productos y resultados de la investigación	10
CAPÍTULO 2. REVISIÓN DE LITERATURA	12
2.1 Introducción	12
2.2 Análisis de Punto de Cambio y el SPC.....	12
2.2.1 Carta de Control Estadístico	12
2.2.2 Punto de Cambio	14
2.3 Revisión de literatura del Análisis de Punto de Cambio.....	15
2.3.1 Análisis Bayesiano	15
2.3.2 Análisis Clásico.....	16
2.3.3 Análisis No-paramétrico	17
2.4 Resumen Histórico del SPC y del CPA.....	17
2.5 Comentarios Generales.....	24
CAPÍTULO 3. INVESTIGACIÓN 1	26

Detección Secuencial y Estimación de Cambios Sostenidos en la Media de Series de Tiempo Usando una Carta de Control con Auto-Inicio y el MLE para Observaciones con Distribución Normal	26
3.1 Introducción	27
3.2 Revisión Literaria	29
3.2.1 CUSUM y las Cartas de control de auto-inicio.....	29
3.2.2 Análisis del Punto de Cambio dentro del SPC	30
3.3 Modelo	31
3.3.1 CUSUM de auto-inicio	31
3.3.2 Estimación del Punto de Cambio.....	33
3.3.3 Análisis del Modelo Secuencial del Punto de Cambio.....	35
3.4 Desempeño de los Estimadores	36
3.4.1 Diseño de la Experimentación	36
3.4.2 Resultados de la Experimentación	37
3.5 Conclusiones y Trabajos Futuros.....	45
CAPÍTULO 4. INVESTIGACIÓN 2	46
Detección Secuencial y Estimación de Cambios Sostenidos en la Varianza de Series de Tiempo Usando una carta de Control con Auto-Inicio y el MLE para Observaciones con Distribución Normal.....	46
4.1. Introducción	47
4.2. Revisión Literaria	49
4.3 Modelo	50
4.3.1 CUSUM de auto-inicio	50
4.3.2 Estimación del Punto de Cambio.....	52
4.3.3 Análisis del Modelo del Punto de Cambio Secuencial.....	54
4.3.4 Diseño de la Experimentación	55
4.4 Resultados de la Experimentación	56
4.5 Conclusiones y Trabajos Futuros.....	64
CAPÍTULO 5. INVESTIGACIÓN 3	65
Estimación del Punto de Cambio Después de la Señal de una Carta de Control para la Gamma.....	65
5.1 Introducción	67
5.2 Trabajos Previos	67
5.3 Gráfica de Control EWMA	68
5.4 Estimación del Punto de cambio.....	69

5.5	Integración del Modelo	70
5.6	Ejemplo Numérico	71
5.7	Resultados de la simulación y la experimentación	74
5.8	Conclusiones y Trabajos Futuros	80
CAPITULO 6. CONCLUSIONES Y TRABAJO FUTURO.....		81
6.1	Conclusiones.....	81
6.2	Trabajo Futuro	86
Bibliografía.....		89

RESUMEN

Esta disertación sigue el formato de tres artículos. Los resúmenes de cada investigación se presentan a continuación:

Investigación 1: Detección Secuencial y Estimación de Cambios Sostenidos en la Media de Series de Tiempo Usando una Carta de Control con Auto-Inicio y el MLE para Observaciones con Distribución Normal.

En esta primera investigación se propone y prueba una metodología para la detección del punto de cambio para series de observaciones que siguen una función de distribución Normal y en los cuales se ha considerado que ocurre un cambio en el valor de la media (μ) que define la función de distribución hasta un momento desconocido específico (τ). La investigación supone una serie de observaciones que son caracterizadas por una función de distribución Normal con un valor inicial de la media (μ_0) y de la desviación estándar (σ_0) y que en un momento desconocido (τ) sufre un cambio en la media (μ_1) pero la desviación estándar se mantiene con el mismo valor ($\sigma_0 = \sigma_1$). Debido a que muchos fenómenos de la vida diaria pueden ser modelados mediante la función de distribución Normal es posible encontrar múltiples aplicaciones. En esta investigación se propone un modelo para estimar el punto de cambio y se realizan múltiples simulaciones para determinar la precisión y exactitud del método en la determinación del valor del parámetro buscado.

Investigación 2: Detección Secuencial y Estimación de Cambios Sostenidos en la Varianza de Series de Tiempo Usando una Carta de Control con Auto-Inicio y el MLE para Observaciones con Distribución Normal.

Esta segunda investigación propone y prueba una metodología para la detección del punto de cambio para series de observaciones que siguen una función de distribución Normal para las cuales se considera el caso en que el parámetro que cambia es la desviación estándar, las medias se mantienen constantes antes y después del momento desconocido τ ($\mu_0 = \mu_1$) mientras que el cambio se presenta en la desviación estándar ($\sigma_0 \neq \sigma_1$) a partir de ese momento. Se propone una metodología para obtener un estimador del punto de cambio (τ), la cual es probada mediante simulaciones y sus resultados son mostrados analizando la desviación del estimador encontrado con respecto al del valor real (sesgo) y la desviación estándar del estimador ($\hat{\tau}$).

Investigación 3: Estimación del Punto de Cambio Después de la Señal Producida por una Carta de Control para la Gamma.

En esta investigación se requiere determinar el punto de cambio considerado un conjunto de datos que siguen una función de distribución Gamma cuyos parámetros de forma y escala (α y β respectivamente) cambian en un momento desconocido (τ); $\alpha_0 \neq \alpha_1$ y $\beta_0 \neq \beta_1$. Se propone una metodología para obtener un estimador del punto de cambio y se evalúa su desempeño comparándolo con otros reportados en la literatura. Para ello se han realizado simulaciones con el fin de probar la eficacia de la metodología y se espera que sea útil para problemas donde la función de distribución Gamma caracterice a la serie de observaciones, esto ocurre particularmente en fenómenos donde el tiempo entre llegadas es un factor determinante, como el caso de las llegadas entre vuelos en un aeropuerto o bien en la bolsa de valores.

LISTA DE TABLAS

Tabla 1 Resumen histórico del SPC y del CPA	21
Tabla 2 Factores para medir el desempeño de la integración del CUSUM de auto-inicio y el MLE	36
Tabla 3 Los estimadores de desempeño sobre diferentes puntos de cambio τ de la carta de control CUSUM de auto-inicio para ARL, el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y el error estándar asociado del punto de cambio usando $\sigma(\tau_{MLE})$, tamaño del subgrupo de 1 ($m = 1$). Los valores fueron calculados usando 10,000 repeticiones.	38
Tabla 4 El desempeño sobre muestras de diferente tamaño de subgrupo (m) de la carta de control CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y la estimación del error estándar asociado para la estimación del punto de cambio $\sigma(\tau_{MLE})$ cuando $\tau = 50$. Los valores fueron calculados usando 10,000 repeticiones.....	39
Tabla 5 El desempeño de la carta de control CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y el error estándar asociado de la estimación del punto de cambio $\sigma(\tau_{MLE})$ cuando $\tau = 50$ y $m = 1$ utilizando diferentes valores de w datos extras.....	40
Tabla 6 El desempeño de la carta de control CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\sigma(\tau_{MLE})$ cuando $\tau = 100$ y $m = 1$ (tamaño del subgrupo) utilizando diferentes valores de w datos extras.	41
Tabla 7 El sesgo para la estimación del punto de cambio mediante el uso de la carta CUSUM de auto-inicio y del MLE, $E(\tau_{CUSUM} - \tau)$ y el error estándar asociado p cambio $\sigma(\tau_{CUSUM})$ para $m = 1$ utilizando diferentes valores de τ . Los valores fueron calculados usando 10,000 simulaciones.....	42
Tabla 8 Probabilidades de los diferentes valores de que el sesgo del estimador y el valor real sea menor a un rango de valores de $\{0, 1, 2, \dots, 24\}$ con $\tau = 100$, tamaño de subgrupo $m = 1$ y $w = 0$ datos extras.....	43
Tabla 9 Probabilidades de que el sesgo del estimador y el valor real sea menor a un rango de valores de $\{0, 1, 2, \dots, 24\}$ con $\tau = 100$, tamaño de subgrupo $m = 1$ y $w = 20$ datos extras.	44
Tabla 10 Factores para medir el desempeño de la integración del CUSUM de auto-inicio y el MLE.	55
Tabla 11 Desempeño estimado considerando diferentes puntos de cambio τ para el CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\sigma(\tau_{MLE})$ usando $m = 1$. Los valores fueron obtenidos realizando 10,000 simulaciones.	57

Tabla 12 Desempeño considerando diferentes muestras de tamaño (n) para el CUSUM de auto-inicio para el ARL , el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\sigma(\tau_{MLE})$ $\sigma(\tau_{MLE} - \tau)$ cuando $\tau = 50$. Los valores fueron calculados realizando 10,000 simulaciones. 58

Tabla 13 Desempeño del CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\sigma(\tau_{MLE})$ cuando $\tau = 50$ y $m = 1$ considerando diferentes valores de w datos extras. Los resultados fueron obtenidos realizando 10,000 simulaciones. 59

Tabla 14 El desempeño del CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $E(\tau_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\sigma(\tau_{MLE})$ cuando $\tau = 100$ y $m = 1$ considerando diferentes valores de w datos extras. Los resultados fueron obtenidos mediante la realización de 10,000 simulaciones. 60

Tabla 15 Comparación de las desviaciones estándar de las estimaciones obtenidas por el uso secuencial de la carta CUSUM de auto-inicio y el MLE para el punto de cambio y el equivalente en la CUSUM de auto-inicio. para $m = 1$ considerando diferentes valores de τ . Los resultados fueron obtenidos utilizando 10,000 simulaciones. 61

Tabla 16 Probabilidades de que el sesgo del estimador del parámetro este dentro de un rango de valores pre-establecidos $\{0, 1, 2, \dots, 24\}$ para diferentes valores de τ_{MLE} con $\tau = 100$, $m = 1$ (tamaño del subgrupo) y $w = 0$ datos extras. 62

Tabla 17 Probabilidades del sesgo del estimador para diferentes valores predefinidos $\{0, 1, 2, \dots, 24\}$ de τ_{MLE} con $\tau = 100$, $m = 1$ (tamaño del subgrupo) y $w = 20$ datos extras..... 63

Tabla 18 Cálculos requeridos para el EWMA y el MLE. Las primeras 15 observaciones provienen de una Gamma ($\alpha_1 = 1$, $\beta_1 = 1$) y las siguientes 6 de una Gamma ($\alpha_1 = 2$, $\beta_1 = 3$), el tamaño de subgrupo es 1, $\lambda = 0.20$ y $L = 2.962$. Las col. 3-5 corresponden a los cálculos para EWMA y las col. 6-8 corresponden a los cálculos para MLE..... 72

Tabla 19 Factores usados para evaluar el desempeño de τ_{EWMA} y τ_{MLE} para $\tau = 100$, $\alpha_0 = 1$, $\beta_0 = 1$, $\lambda = 0.20$ y $L = 2.962$ 74

Tabla 20 Desempeño considerando diferentes valores de α_1 y β_1 para las muestras de la Gamma, utilizando una carta de control EWMA para estimar el punto de cambio. Los parámetros de la simulación son: $\tau = 100$; 1,000 simulaciones; $\alpha_0 = 1$, $\beta_0 = 1$, $\lambda = 0.20$, $L = 2.962$; el ARL y desviación estándar SE_{EWMA} , son calculados..... 76

Tabla 21 Desempeño considerando diferentes valores de α_1 y β_1 para muestras de la Gamma utilizando el uso secuencial del EWMA y el MLE para estimar el punto de cambio. Los parámetros son $\tau = 100$;

1,000 simulaciones; $\alpha_0 = 1$, $\beta_0 = 1$, $\lambda = 0.20$ y $L = 2.962$; el sesgo ($\hat{t}_{MLE} - \tau$) y desviación estándar, S_{MLE} , son calculados..... 77

Tabla 22a Comparativa del desempeño de la carta de control EWMA y el uso secuencial de la EWMA y el MLE para el punto de cambio. Presenta variaciones en β de 1.00 a 2.00. Se han incluido el valor de μ y σ . Esta tabla está basada en las tablas 20 y 21 y permite comparar la detección del punto de cambio de ambos estimadores para diferentes parámetros α y β pero con misma μ o σ 78

Tabla 22b Comparativa del desempeño de la carta de control EWMA y el uso secuencial de la EWMA y el MLE para el punto de cambio. Presenta variaciones en β de 1.00 a 2.00. Se han incluido el valor de μ y σ . Esta tabla está basada en las tablas 20 y 21 y permite comparar la detección del punto de cambio de ambos estimadores para diferentes parámetros α y β pero con misma μ o σ 79

LISTA DE FIGURAS

Figura 1 Ejemplo de carta de control de Shewhart obtenida en Minitab tm usando los datos de [6] pág. 407.....	13
Figura 2 Ejemplo de carta de control de CUSUM bilateral obtenida en Minitab tm , usando los datos de [6] pag. 407	14
Figura 3 Ejemplo de carta de control de EWMA obtenida en Minitab tm , usando los datos de [6] pág. 407	14
Figura 4 Gráfica del número de artículos publicados sobre CPA en revistas especializadas vs año de publicación.	18
Figura 5 Gráfica comparativa entre las publicaciones de investigación y aplicación.....	19
Figura 6 Número de publicaciones.....	20
Figura 7 Número de publicaciones.....	20
Figura 8 Detección secuencial del punto de cambio y estimación con la CUSUM de auto-inicio y el MLE del punto de cambio para series de observaciones normales independientes.....	35
Figura 9 Detección secuencial del punto de cambio y su estimación con el CUSUM de auto-inicio y el MLE del punto de cambio para series de observaciones normales independientes.....	54
Figura 10 Detección secuencial del punto de cambio y estimación del mismo usando una gráfica de control para la Gamma y el MLE para el punto de cambio para una serie de observaciones independientes.....	70
Figura 11 Gráfica de dispersión de los datos del ejemplo numérico.	73
Figura 12 Carta de control EWMA para los datos del ejemplo numérico.....	73

CAPÍTULO 1. INTRODUCCIÓN

1.1 Historia

La calidad ha sido una preocupación desde que el hombre empieza a crear/producir satisfactores para su vida diaria; antiguos escritos, como el Código de Hammurabi, detallan como los castigos por un trabajo mal hecho, que no cumplía con los estándares para los que fue construido, podían ir desde la mutilación hasta la pena de muerte. Los tiempos han cambiado y el concepto de calidad ha ido evolucionado desde la idea donde el artesano, productor del bien, era quien definía la calidad del producto a su total arbitrio al tiempo presente en donde se ha extendido a considerar una serie de factores que involucran el aspecto económico, social, humano, entre otros, para conformar una definición basada en la relación productor-comercializador-cliente-satisfacción. Bounds et al. [1] detalla como el concepto de calidad ha transitado por diferentes etapas:

- 1º. la de inspección (siglo XIX),
- 2º. la del control estadístico del proceso (década de los treinta),
- 3º. la del aseguramiento de la calidad (década de los cincuenta) y
- 4º. la era de la administración estratégica por calidad total (la década de los noventa).

En los tiempos actuales (siglo XXI) según Cantú [2] “se está viviendo una etapa de innovación y tecnología, en la que la competitividad depende de la capacidad para responder a los cambios en el mercado y las fluctuaciones sociales, políticas, económicas y financieras con una alta velocidad soportada por la innovación rápida y el uso de la tecnología, tanto de procesos/operación como de información”.

Del párrafo anterior se observa que la evolución en la calidad ha sido exponencial a partir del siglo pasado, en particular a raíz del trabajo de Shewhart [3] quien detalló métodos estadísticos para ser usados en el control de la variabilidad de un proceso y estableció las primeras definiciones modernas de los conceptos de calidad, control, límite de control, variación, causas asignables, eliminación de causas y las consecuencias de seguir estas metodologías sentando las bases para lo que sería el control estadístico de la calidad. A la par con esto, la introducción del concepto de carta de control y la aceptación del control estadístico de la calidad (SPC), en diferentes industrias en América, Europa y Asia empezaron a utilizar el diseño de experimentos para el desarrollo de productos iniciando los primeros planes de muestreo. Este fue el albor de una filosofía que cambiaría la percepción mundial sobre la calidad y llamaría la atención de ingenieros y científicos sobre la mejor forma de controlar la variación en los procesos, detectando esa variación antes de que llegue a ser un problema y modificando los paradigmas de qué es calidad, incluyendo al cliente y su percepción del producto como parte de la definición de calidad y, desde luego, ligar todos estos elementos al costeo de la producción y la relación costo-beneficio-aceptación final del producto entre productor-comercializador-cliente.

Las cartas de control son de las primeras herramientas del SPC e implican la medición de una o más características de un proceso con el fin de dar un seguimiento a la variabilidad que se presenta en el mismo. Este seguimiento unido a los límites de control permitirá determinar cuándo un proceso se ha salido de control estadístico, entendiéndose con esto, que la variabilidad del proceso está más allá de los límites establecidos en la carta de control estadístico. Las primeras cartas de control, de Shewhart, fueron diseñadas para monitorear la variabilidad de un proceso y alertar cuando la variabilidad del mismo excedía los límites establecidos por el personal encargado de la calidad del producto. Son herramientas útiles para detectar cambios abruptos en la variabilidad del proceso pero cuando el cambio es lento y sostenido, estas cartas no son sensibles al mismo, indicando con esto que la detección es tardía. Para solucionar esta situación Page [4] presenta la carta de control de CUSUM (CUMulative SUMs) en 1954 la cual mantiene un monitoreo sobre la variabilidad del proceso pero es más sensible a los cambios lentos y sostenidos en el mismo; para este efecto las cartas de control CUSUM llevan un conteo sobre las tendencias recientes de los datos. En la misma dirección se encuentra la carta de control CUSUM de auto-inicio propuesta por Hawkins [5], cuya ventaja es poder realizar una carta de control utilizable directamente en la Fase II del SPC al no requerir los ajustes propios de la Fase I. La Fase I es considerada como las actividades y/o acciones en que se analizan las características del proceso para llevarlo a un estado de control estadístico y estimar los parámetros del proceso mientras que en la Fase II las cartas de control monitorean a éste [6]. Por otro lado Roberts en 1959 [7] presenta la carta de control EWMA (Exponentially Weighted Moving Average) cuya principal característica es dar distinta ponderación a las observaciones obteniendo una carta de control útil para detectar pequeños cambios sostenidos en la media del proceso. Con estas técnicas se están atacando dos problemas que se presentan en los cambios en la variabilidad de un proceso: los cambios repentinos y los cambios lentos y sostenidos. Las primeras cartas de control estadístico creadas permiten la detección de los cambios repentinos considerando principalmente los últimos datos recolectados, mientras que las cartas tipo CUSUM y EWMA son más sensibles a los cambios en el proceso al dar una valoración diferente a las observaciones. Al analizar el problema del punto de cambio es inevitable que surjan algunas interrogantes sobre la información disponible. Algunas de estas interrogantes son: ¿qué ocurre con las observaciones que empezaron a estar fuera de control (su variabilidad afecta al sistema en un plazo de tiempo) y no fueron detectadas? ¿Cómo se puede determinar el momento en que estas observaciones ya no son representadas por la función de distribución supuesta para los parámetros iniciales con los que se definieron los límites de control? Estas preguntas son compartidas por el autor de la presente tesis y autores de diferentes artículos reportados en la literatura y que son señalados a continuación.

Para contestar esas preguntas se han propuesto diferentes teorías, iniciando con Girshick y Rubin [8] en 1952 quienes sientan las bases para resolver el problema desde un enfoque bayesiano, su aportación sustenta las bases matemáticas para determinar la ubicación del

momento en que el proceso deja de estar en control estadístico. Otro frente de trabajo es utilizar el método de máxima verosimilitud para determinar el punto de cambio en un proceso, los primeros trabajos en esta dirección fueron de Hinkley [9] en 1970 donde el proceso de determinar el punto de cambio es desarrollado bajo el concepto de estimador de máxima verosimilitud (MLE por sus siglas en inglés). Un resultado reciente, que implica la creación de cartas de control, utilizando la formulación del problema de punto de cambio mediante el cociente de máxima verosimilitud generalizado (GLR, por sus siglas en inglés) ha sido desarrollada por Hawkins y Zamba [10] en 2005. Todos estos trabajos tienen como característica común que están basados en datos que siguen una distribución y se pueden catalogar como métodos basados en estadística paramétrica, mientras que los métodos que estudian los casos en que no se asume una distribución para los datos se conocen como no paramétricos y son de reciente desarrollo. Amin y Reynolds [11] presentan en 1996 una carta de control basada en el estadístico “regla del signo”, por su parte Das y Bhattachrya [12] en 2008 proponen una carta de control para monitorear la variabilidad utilizando rangos, finalmente Tercero et al. [13] en 2012 desarrollan un estimador de punto de cambio basado en la mediana para un conjunto de datos que no necesariamente tienen una función de distribución que los determine.

Con el desarrollo de la teoría del análisis de punto de cambio se han ido encontrando alternativas para implementarlo siendo una de ellas la que corresponde a la combinación de las cartas de control con el MLE para estimar el punto de cambio una vez que la carta de control ha enviado una señal de alerta. Es precisamente esta metodología la que toma la presente investigación y la particulariza para los casos en que los datos pueden caracterizarse mediante una función de distribución Normal o una función de distribución Gamma.

1.2 Especificación del problema

En la continua búsqueda de nuevas formas de mejorar el SPC, surge el análisis de punto de cambio que consiste en determinar la existencia de cambios estructurales y estimar el momento en que ocurrieron. Bajo el marco del control estadístico de procesos, la ocurrencia de un cambio estructural se refleja como un estado fuera de control donde los parámetros que se monitorean han cambiado. Considerando que la función de distribución sigue siendo la misma, un cambio define una modificación en los parámetros iniciales del proceso, y se hace necesario identificar su inicio para tener una mejor oportunidad de determinar las causas que lo provocaron y tomar las medidas pertinentes. Los tipos de cambios que pueden ocurrir bajo este esquema pueden ser: el cambio se presenta en uno, en varios o en todos los parámetros, en un momento único o en varios y desde luego las posibles combinaciones resultantes. La estimación de este momento, conocido como punto de cambio, facilita la búsqueda de las causas que lo motivan. Esto no es necesariamente reflejado en las cartas de control, ya que en éstas la señal del proceso fuera de control se produce de manera posterior al cambio en las observaciones del proceso. En la literatura [6] y [14] reportan que usualmente las cartas de control de Shewhart

tienen un pobre desempeño, en términos del promedio de longitud de corrida (ARL por sus siglas en inglés), frente a cambios sostenidos menores a 1.5 desviaciones estándares en comparación con las cartas de control CUSUM y EWMA. El mismo Shewhart [3] (pp 19) parece prever esta situación respecto a las cartas de control al escribir: “Esto muestra que cuando los puntos caen fuera de los límites, la experiencia indica que una causa asignable puede ser encontrada, pero eso no indica que cuando los puntos caen dentro de los límites establecidos, nosotros no podamos encontrar causas de variación.” Esta observación de Shewhart invita a reflexionar sobre la variante que ocupa el presente trabajo; una vez que se sabe que un proceso ha dejado de estar en control estadístico, ¿cuáles puntos, que estaban dentro de los límites de control, realmente corresponden al proceso con parámetros iniciales que dieron lugar a los actuales límites de control?. Se está suponiendo que la señal emitida por la carta de control no es una falsa alarma, que consiste en que la carta de control estadístico indique que el proceso ha salido de control estadístico cuando en realidad no es así. El problema de identificar el momento real en que el cambio ocurre es conocido como el problema de estimar el punto de cambio (τ), mientras que el proceso de resolverlo es llamado el análisis de punto de cambio (CPA por sus siglas en inglés). A través del CPA es posible desarrollar estimadores que asistan en la estimación de puntos de cambios una vez detectados por las cartas de control y así solventar esta debilidad que tienen las mismas.

Esta investigación analiza el problema de determinar el punto de cambio mediante la implementación secuencial de las cartas de control y el MLE para el punto de cambio. Se considera que los datos siguen un tipo de función de distribución en particular y que los parámetros iniciales, que la definen, cambian a partir de un momento específico en el tiempo. En particular se estudia la implementación de estimadores obtenidos a partir del MLE para punto de cambio con cartas de control CUSUM de auto-inicio para la función de distribución Normal y cartas de control EWMA para distribuciones Gamma.

1.3 Preguntas que Genera la Investigación

Investigación 1.- Considere una secuencia de observaciones independientes $x_1, x_2, x_3, \dots, x_{\tau-1}, x_{\tau}, x_{\tau+1}, \dots, x_{T-2}, x_{T-1}, x_T$ con distribución Normal con parámetros iniciales desconocidos (μ_0, σ_0) hasta un momento específico también desconocido τ a partir del cual μ_0 cambia a un valor desconocido μ_1 . Considere el estimador de punto de cambio ($\hat{\tau}$) obtenido mediante el uso secuencial de cartas de control y del MLE para punto de cambio ¿cuál es el desempeño de este estimador? Y ¿cuál es el desempeño de este estimador comparado con otros estimadores bajo las mismas condiciones?

Investigación 2.- Considere una secuencia de observaciones independientes $x_1, x_2, x_3, \dots, x_{\tau-1}, x_{\tau}, x_{\tau+1}, \dots, x_{T-2}, x_{T-1}, x_T$ con distribución Normal con parámetros iniciales desconocidos (μ_0, σ_0) hasta un momento específico desconocido τ a partir del cual σ_0 cambia a un valor desconocido σ_1 . Considere el estimador de punto de cambio $(\hat{\tau})$ obtenido mediante el uso secuencial de cartas de control y del MLE para punto de cambio ¿cuál es el desempeño de este estimador? Y ¿cuál es el desempeño de este estimador comparado con otros estimadores bajo las mismas condiciones?

Investigación 3.- Considere una secuencia de observaciones independientes $x_1, x_2, x_3, \dots, x_{\tau-1}, x_{\tau}, x_{\tau+1}, \dots, x_{T-2}, x_{T-1}, x_T$ con distribución Gamma con parámetros iniciales conocidos (α_0, β_0) hasta un momento específico desconocido τ a partir del cual los nuevos valores de los parámetros serán (α_1, β_1) , desconocidos. Considere el estimador de punto de cambio $(\hat{\tau})$ obtenido mediante el uso secuencial de cartas de control y del MLE para punto de cambio ¿cuál es el desempeño de este estimador? Y ¿cuál es el desempeño de este estimador comparado con otros estimadores bajo las mismas condiciones?

1.4 Hipótesis General

Las respuestas a las preguntas anteriores son respondidas en los capítulos 3, 4 y 5, para ello se hace el replanteo de las preguntas como hipótesis, las cuales se detallan a continuación:

- Investigación 1 (Capítulo 3):
 - Hipótesis 1: El estimador obtenido mediante la aplicación secuencial de una carta de control CUSUM de auto-inicio y un MLE para el punto de cambio en un conjunto de datos independientes caracterizado por una función de distribución Normal con parámetros iniciales μ_0 y σ_0 y que a partir del instante desconocido τ cambian a μ_1 y σ_1 con la característica que $\mu_0 \neq \mu_1$ y $\sigma_0 = \sigma_1$, todos ellos desconocidos, tiene un mejor desempeño, en términos de sesgo y variabilidad, que el estimador del punto de cambio obtenido por de la carta CUSUM de auto-inicio.
- Investigación 2 (Capítulo 4):
 - Hipótesis 2: El estimador obtenido mediante la aplicación secuencial de una carta de control CUSUM de auto-inicio y un MLE para el punto de cambio en un conjunto de datos independientes caracterizado por una función de distribución Normal con parámetros iniciales μ_0 y σ_0 que a partir del instante desconocido τ cambian a μ_1 y σ_1 con la característica que $\mu_0 = \mu_1$ y $\sigma_0 \neq \sigma_1$, todos ellos desconocidos, tiene un mejor desempeño, en términos de sesgo y variabilidad,

que el estimador del punto de cambio obtenido por la carta CUSUM de auto-inicio.

- Investigación 3 (Capítulo 5):

- Hipótesis 3: El estimador obtenido mediante la aplicación secuencial de una carta de control EWMA y un MLE para el punto de cambio en un conjunto de datos caracterizado por una función de distribución Gamma con parámetros iniciales conocidos α_0 y β_0 que a partir del instante desconocido τ cambian a α_1 y β_1 , ambos desconocidos, con la característica que $\alpha_0 \neq \alpha_1$ y/o $\beta_0 \neq \beta_1$, tiene un mejor desempeño, en términos de sesgo y variabilidad, que el estimador del punto de cambio obtenido por la carta EWMA

En cada caso el sesgo será medido como el promedio de las diferencias entre el valor del estimador ($\hat{\tau}$) y el valor real (τ) del punto de cambio. La variabilidad será medida como la desviación estándar del estimador $\hat{\tau}$.

1.5 Formato de la Investigación

El presente trabajo sigue un formato de tres investigaciones. Una introducción general y la revisión de literatura son presentadas en el capítulo dos, seguida de las tres investigaciones relacionadas con el análisis de punto de cambio y termina con las conclusiones generales. El capítulo dos presenta una revisión del análisis de punto de cambio (CPA) dentro del control estadístico de procesos (SPC) e incluye una revisión de los principales trabajos, a juicio del autor de esta tesis, realizados en el área. Esta revisión es hecha desde tres enfoques: bayesiano, clásico y no-paramétrico. En el capítulo tres se hace un análisis del estimador obtenido a partir del uso secuencial de una carta de control CUSUM de auto-inicio y el MLE para punto de cambio en el caso de una función de distribución Normal cuya media (μ) ha cambiado a partir de un momento específico desconocido (τ) mientras la varianza no lo hace. El capítulo cuatro presenta un análisis del estimador de punto de cambio obtenido mediante el uso secuencial de una carta de control CUSUM de auto-inicio y el MLE para punto de cambio en un conjunto de datos caracterizado por una distribución Normal cuya varianza (σ^2) ha cambiado a partir de un momento específico desconocido (τ) mientras las media (μ) no lo hace. El capítulo cinco hace un análisis del estimador de punto de cambio obtenido mediante el uso secuencial de una carta de control EWMA y el MLE para el punto de cambio en un conjunto de datos caracterizado por una función de distribución Gamma cuyos parámetros iniciales han cambiado a partir de un momento específico desconocido (τ). El capítulo seis contiene la descripción de los hallazgos de estas investigaciones y algunas posibles líneas de investigación para trabajos futuros.

1.6 Propósito de la Investigación

El propósito final de esta investigación es desarrollar herramientas que asistan a los analistas de procesos en la mejora continua de la calidad a través de la detección e identificación de causas asignables de variación en un menor tiempo y a un menor costo al que se requeriría utilizando exclusivamente cartas de control. Para lograr este propósito último, se propone el uso secuencial de una carta de control y el MLE para el punto de cambio y evaluar el desempeño de los estimadores así obtenidos para determinar si presentan un mejor desempeño que otros reportados en la literatura. Para llevar a cabo esta evaluación se han seleccionado los siguientes casos:

- 1) Una serie de observaciones independientes normales cuyos parámetros μ 's o σ 's han cambiado, solo uno de los dos, a partir de un momento específico desconocido τ . Los parámetros previos y posteriores al momento de cambio se consideran desconocidos.
- 2) Una serie de observaciones independientes gamma cuyos parámetros iniciales (α y β) son conocidos y a partir de un momento específico desconocido τ uno o ambos parámetros cambian a valores desconocidos. Es de esperar que los resultados obtenidos por la metodología propuesta puedan ser comparados con los resultados obtenidos por otras metodologías bajo las mismas condiciones.

1.7 Objetivo de la Investigación

El objetivo de esta investigación es diseñar un modelo para analizar el punto de cambio en los parámetros de una función de distribución que caracterice a una serie de observaciones independientes. Para conseguir este objetivo principal se proponen los siguientes objetivos particulares por investigación realizada:

Investigación 1: Diseñar un modelo para analizar el punto de cambio en la media en series con observaciones independientes y normales mediante la integración de cartas de control CUSUM de auto-inicio para medias y un estimador MLE para punto de cambio. El desempeño se evaluará mediante el cálculo del sesgo ($\hat{\tau} - \tau$) y de la desviación estándar de $\hat{\tau}$. Ambos valores serán comparados con el ARL y la desviación estándar obtenidos a partir de las estimaciones generadas por la CUSUM de auto-inicio. El modelo que representa a las observaciones es:

Sea una secuencia de variables aleatorias independientes $x_1, x_2, x_3, \dots, x_{\tau-1}, x_{\tau}, x_{\tau+1}, \dots, x_{T-2}, x_{T-1}, x_T$ con una función de distribución Normal con parámetros μ y σ , en particular considerando un cambio en los parámetros en el momento τ se puede representar el modelo como:

$$X_i \sim \begin{cases} N(\mu_0, \sigma_0), & 1 \leq i \leq \tau \\ N(\mu_1, \sigma_0), & \tau < i \leq T \end{cases} \quad (1)$$

Investigación 2: Diseñar un modelo para analizar el punto de cambio en la desviación estándar en series con observaciones independientes y normales mediante la integración de cartas de control CUSUM de auto-inicio para la desviación estándar y un estimador MLE para punto de cambio. El desempeño se evaluará mediante el cálculo del sesgo $(\hat{\tau} - \tau)$ y de la desviación estándar de $\hat{\tau}$. Ambos valores serán comparados con el ARL y la desviación estándar obtenidos a partir de las estimaciones generadas por la CUSUM de auto-inicio. El modelo que representa a las observaciones es:

Sea una secuencia de variables aleatorias independientes $x_1, x_2, x_3, \dots, x_{\tau-1}, x_{\tau}, x_{\tau+1}, \dots, x_{T-2}, x_{T-1}, x_T$ con una función de distribución Normal con parámetros μ y σ , en particular considerando un cambio en los parámetros en el momento τ se puede representar el modelo como:

$$X_i \sim \begin{cases} N(\mu_0, \sigma_0), & 1 \leq i \leq \tau \\ N(\mu_1, \sigma_1), & \tau < i \leq T \end{cases} \quad (2)$$

Investigación 3: Diseñar un modelo para analizar el punto de cambio en la desviación estándar en series con observaciones independientes y con distribución Gamma mediante la integración de cartas de control EWMA y un estimador MLE para punto de cambio. El desempeño se evaluará mediante el cálculo del sesgo $(\hat{\tau} - \tau)$ y de la desviación estándar de $\hat{\tau}$. Ambos valores serán comparados con el ARL y la desviación estándar obtenidos a partir de las estimaciones generadas por la carta EWMA. El modelo que representa a las observaciones es:

Sea una secuencia de variables aleatorias independientes $x_1, x_2, x_3, \dots, x_{\tau-1}, x_{\tau}, x_{\tau+1}, \dots, x_{T-2}, x_{T-1}, x_T$ con una función de distribución Gamma con parámetros α y β , en particular considerando un cambio en los parámetros en el momento τ se puede representar el modelo como:

$$X_i \sim \begin{cases} G(\alpha_0, \beta_0), & 1 \leq i \leq \tau \\ G(\alpha_1, \beta_1), & \tau < i \leq T \end{cases} \quad (3)$$

1.8 Alcances y Limitaciones de la Investigación

La metodología propuesta está enfocada a presentar y evaluar, mediante la comparación del desempeño de este y otros estimadores de punto de cambio, una alternativa para la estimación del punto de cambio particularizando en series de datos que pueden ser representados por funciones de distribución Normal o Gamma en donde los parámetros cambian a partir de un momento específico, desconocido y único τ , esto es, en ningún momento se consideran múltiples cambios en los parámetros. En particular para el caso de la Normal se considera el cambio en las medias o en las varianzas pero no en ambos parámetros de manera simultánea; esta variante de cambios ha sido tratada, incluyendo más de un cambio, por Garza et. al [15] en donde se aplica el método de máxima verosimilitud y se obtiene una función entera a optimizar

utilizando un algoritmo heurístico para aproximar las soluciones. Para el caso de la función de distribución Gamma los artículos reportados no son tan numerosos como los enfocados en la Normal. La necesidad de considerar a la función de distribución Gamma es que cuando las aplicaciones del SPC suponen que las observaciones del proceso se distribuyen siguiendo esta función de distribución no siempre pueden recurrir al Teorema del Límite Central (CLT por sus siglas en inglés) para buscar la normalidad ya que en muchas ocasiones se debe trabajar con observaciones individuales. Esto se debe en parte a las restricciones económicas y prácticas que lleva el proceso de muestreo en cada situación.. La industria química, la aeronáutica y el mercado de valores [16] son algunos casos donde esta circunstancia se presenta. En los artículos consultados la estimación del punto de cambio se realiza solo sobre el parámetro de escala (β) y no se considera ningún cambio en el parámetro de forma (α). Sin embargo, el supuesto de que sólo un parámetro cambia después de que el proceso recibe una perturbación no siempre se puede sostener. Por lo tanto, es más conservador asumir que en la práctica ambos parámetros pudieron haber cambiado. Para evaluar el desempeño se estima el sesgo ($\hat{E}[\hat{\tau} - \tau]$) y el error estándar del estimador ($\hat{\sigma}_{\hat{\tau}}$) de punto de cambio a través de simulaciones Monte Carlo. El sesgo se estima como el promedio de las diferencias entre las estimaciones del punto de cambio ($\hat{\tau}$) y el punto de cambio mismo (τ), mientras que el segundo se estima al obtener la desviación estándar de las estimaciones realizadas. Estos valores fueron obtenidos mediante 10,000 réplicas utilizando el método Monte Carlo para la Normal y de 1,000 para el caso de la Gamma (esta última cantidad es menor que la primera debido a la complejidad computacional que requieren la obtención de cada réplica). También han sido considerados los casos en donde se tienen más de una observación por muestra (solo Normal) y el caso en que se han incluido datos más allá del momento en que la carta de control manda la señal de fuera de control estadístico. Los capítulos 3, 4 y 5 detallan los resultados obtenidos para cada uno de estos casos mediante tablas que presentan los sesgos $\hat{E}[\hat{\tau} - \tau]$ y desviaciones estándar $\hat{\sigma}_{\hat{\tau}}$ de las estimaciones obtenidas.

1.9 Importancia del Presente Estudio

La aplicación de la metodología presentada aquí permite un mejor entendimiento del CPA dentro del marco del SPC, hace aportaciones a la teoría correspondiente y da una oportunidad para disminuir los costos que involucra la detección oportuna de un punto de cambio. Con esta metodología es posible ahorrar tiempo y costo en la detección de la causa asignable que provoca la variabilidad en el proceso que conduce a llevarlo a estar fuera de control estadístico, ya que indica el momento, casi preciso, en que el proceso empezó a generar variaciones que lo llevarán a estar en un estado fuera de control estadístico; es claro que esta información reduce el área de búsqueda de la(s) causa(s) de la variación.

Respecto a la contribución que hace a la teoría detrás del SPC, ésta radica en combinar dos herramientas del SPC de manera secuencial y contribuir a un enfoque que ayuda a romper paradigmas de cómo pueden interactuar para conformarse en una poderosa metodología dentro del mismo. Particularmente, el uso del MLE para el punto de cambio en los casos de la Normal y de la Gamma no había sido tratado bajo las condiciones y variaciones que presenta esta investigación.

La aplicación de esta metodología está ligada al uso de las cartas de control para procesos Normal o Gamma. Los estimadores aquí mostrados se utilizan como complementos de cartas de control, para facilitar la búsqueda de causas asignables de variación después de señalado un posible cambio en el proceso. Tal es el caso del problema de determinar la máxima vida útil (maximum service life, B_X life) de un componente, cuya vida sigue una distribución normal [17] o de determinar el cambio en el clima en el invierno en la ciudad de Seoul [18]. Por otra parte la Gamma tiene aplicaciones en el comportamiento del mercado de valores, [16] hace un análisis del comportamiento del índice Down Jones en el período de julio 1971 a julio de 1974 considerando los valores de cierre semanal reportados.

Los resultados que se detallan en esta disertación hacen pensar en la metodología propuesta como una herramienta útil en la estimación del punto de cambio y en el control estadístico del proceso en general, utilizando un concepto ya muy conocido como lo es la carta de control estadístico complementada con una herramienta que está mostrando su utilidad en el CPA como lo es el MLE de puntos de cambio. La dificultad que podría existir en la aplicación de la herramienta propuesta es debido a lo complejo que pudieran ser algunos cálculos, esta situación es zanjada por la potencia computacional, tanto en hardware como en software, actual y que ayuda a eliminar lo que podría ser una brecha entre un desarrollo teórico y la aplicación a casos cotidianos en la industria.

1.10 Productos y resultados de la investigación

El principal producto de esta investigación es un procedimiento para la obtención de estimadores del punto de cambio a partir del uso secuencial de una carta de control estadístico y el MLE para el punto de cambio. Esta herramienta auxiliara a las personas involucradas en el control de procesos, en la búsqueda de las causas asignables que llevan al proceso a estar fuera de control. En particular se espera obtener los siguientes resultados:

- 1) Los estimadores obtenidos por los procedimientos indicados para el caso de las observaciones independientes y normales no se afectan negativamente por el tamaño de las repeticiones para cada observación individual (tamaño del subgrupo) ni por considerar datos posteriores al momento en que la carta de control detecta el proceso fuera de control, esto último incluye al caso de las observaciones independientes y gamma.

- 2) Los estimadores obtenidos a partir de los modelos propuestos presentan un sesgo y desviación estándar menor, en la mayoría de los casos, a los obtenidos por el uso exclusivo de las cartas de control CUSUM y EWMA de auto-inicio.

CAPÍTULO 2. REVISIÓN DE LITERATURA

2.1 Introducción

El SPC es una herramienta que involucra el monitoreo de las variaciones en las mediciones de una o varias características de un producto, en el sentido más amplio, con el fin de facilitar la búsqueda para detectar la(s) causa(s) que la(s) provoca(n), si es que ésta(s) existe(n). Este fin implica una serie de pasos intermedios que se han ido tipificando, analizando, estudiando, proponiendo modificaciones y/o nuevos pasos para hacer del SPC una herramienta en continua mejora cumpliendo con las expectativas de ahorrar tiempo, trabajo y dinero en los procesos de producción de un bien. La frase referida en la página 4 del Capítulo 1 correspondiente a la opinión de Shewhart, la cual se puede interpretar como una visión de lo que es el concepto de punto de cambio y la importancia de tener una forma de estimarlo, implica que aunque las muestras sigan reportando que el proceso sigue en control estadístico, no es necesariamente cierto que esas muestras no sean parte de la serie de observaciones que definen un nuevo valor del parámetro de la función de distribución que lo caracteriza. El identificar este punto de cambio es relevante para la búsqueda de la causa asignable, desde luego, es necesario asumir que no se trata de una falsa alarma. En la búsqueda de una herramienta que permita estimar el valor real del punto de cambio se han desarrollado diversas herramientas, incluso importando resultados de otras áreas del conocimiento y planteándose nuevas formas de ver el problema de monitorear un proceso. Page [4], en 1954, hace la primera propuesta para determinar el punto de cambio al utilizar cartas de control y una prueba basada en sumas acumuladas. Box y Luceño [19] en 1979 demostraron que las cartas de control de Shewhart son óptimas para detectar puntos anormales aislados, mientras que las cartas de control CUSUM son óptimas para detectar cambios sostenidos en la media. Si se está interesado en ambos tipos de cambios, los dos procedimientos pueden ser usados para complementarse uno al otro. Actualmente el CPA ha incluido el uso combinado de las cartas de control, en general, con estimadores de punto de cambio para mejorar la estimación de cambios sostenidos en un proceso. Es en esta combinación de uso secuencial de ambas herramientas que se enfoca el presente trabajo por lo que la revisión de literatura ha convergido en ese tema.

2.2 Análisis de Punto de Cambio y el SPC

Para atender el CPA y el SPC es conveniente revisar brevemente los conceptos de carta de control estadístico y de punto de cambio.

2.2.1 Carta de Control Estadístico

Como ya se ha mencionado las cartas de control son una herramienta muy útil en el SPC ya que permiten, entre otras cosas, monitorear la variabilidad de un proceso y es en base a la forma en que realizan esta actividad en que han sido clasificadas por Koutras, Bersimis y Maravelakis [20] en tres grandes categorías:

- Cartas de Control de Shewhart, desarrolladas por Shewhart [3] en 1931, en las que es relativamente fácil detectar causas que conducen a grandes cambios, sostenidos o aislados. En la Fig. 1 se muestra un ejemplo de estas cartas de control. En ellas detalla el Límite Superior de Control (UCL, por sus siglas en inglés), el Límite Inferior de Control (LCL, por sus siglas en inglés) y la Línea Central (CL, por sus siglas en inglés).
- Cartas de Control de CUSUM, propuestas por Page [4] en 1954 cuya principal área de oportunidad son los casos donde los cambios son pequeños y sostenidos. En estas cartas se empieza a guardar una historia del proceso a diferencia de las cartas de control de Shewhart. Un ejemplo se muestra en la Fig. 2, esta figura presenta el UCL, el LCL y la CL.
- Carta de Control EWMA cuya aplicación es similar a la de CUSUM, con la variante que en ésta es posible darle mayor peso a las observaciones más recientes. Esto se logra mediante el uso del valor λ y fue desarrollada por Roberts [7] en 1959. Un ejemplo se muestra en la Fig. 3. En esta figura se detallan el UCL, el LCL y la CL.

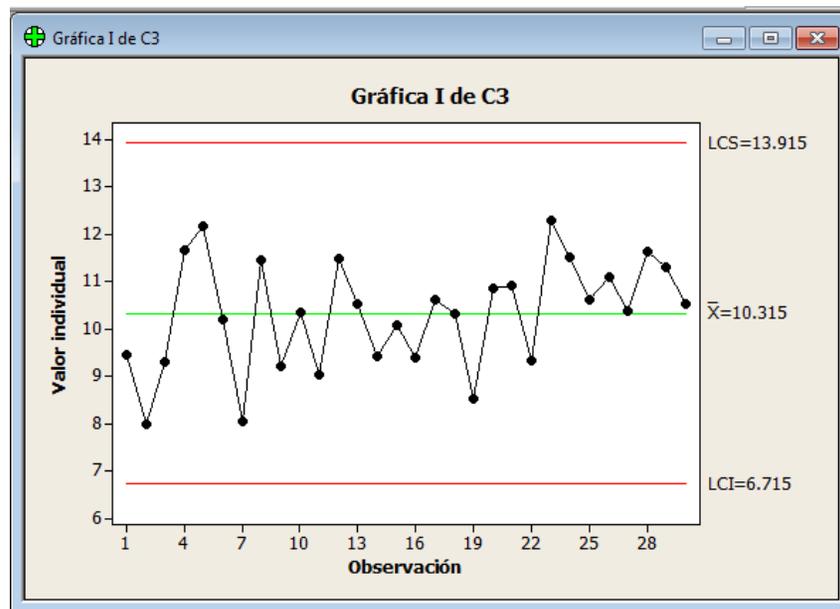


Figura 1 Ejemplo de carta de control de Shewhart obtenida en Minitabtm usando los datos de [6] pág. 407

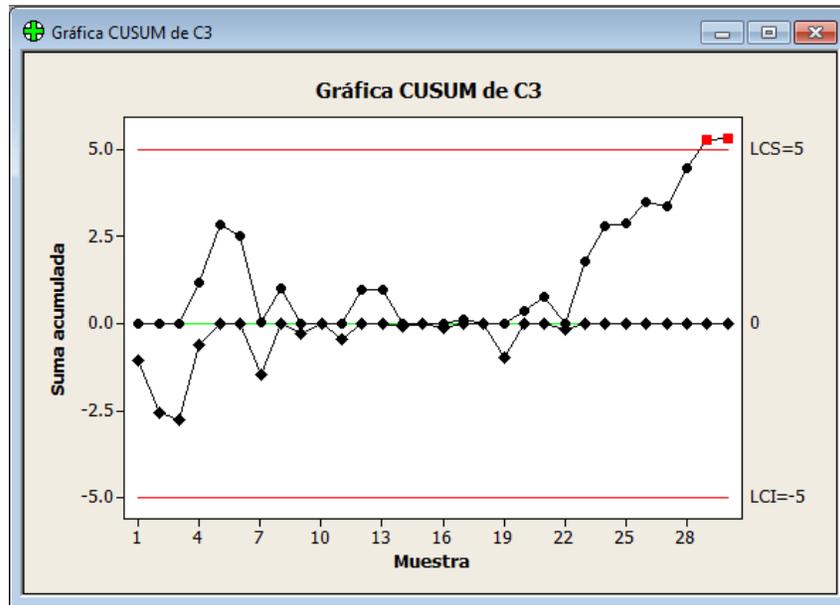


Figura 2 Ejemplo de carta de control de CUSUM bilateral obtenida en Minitabtm, usando los datos de [6] pag. 407

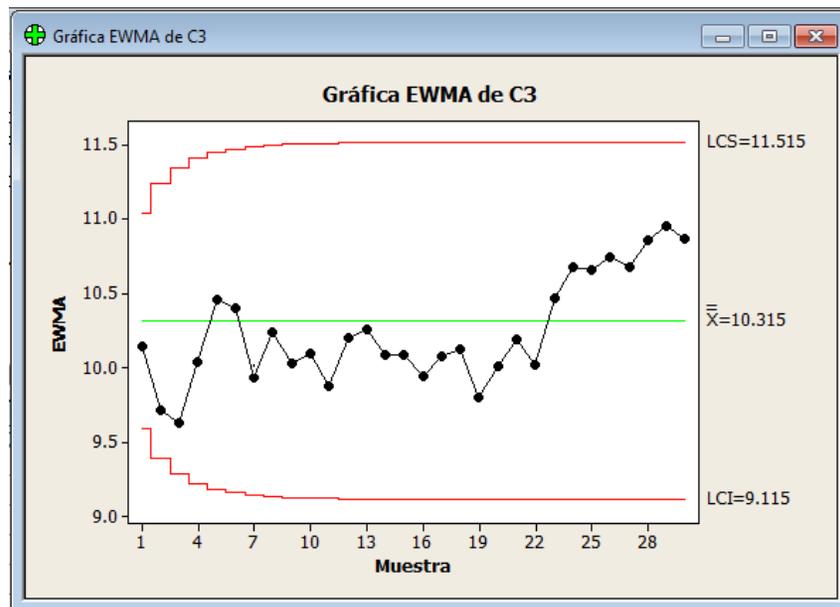


Figura 3 Ejemplo de carta de control de EWMA obtenida en Minitabtm, usando los datos de [6] pág. 407

2.2.2 Punto de Cambio

El punto de cambio es el momento en que al menos uno de los parámetros de la función de distribución que caracteriza al proceso, del que se está monitoreando la variabilidad, cambia. Page [21] en 1957 hace un tratado sobre los problemas que se presentan cuando cambia un parámetro en un momento desconocido y es una referencia frecuente en la literatura. La

conexión entre el CP y las cartas de control estadístico surge del hecho que un proceso que ha salido de control estadístico presenta un punto de cambio, suponiendo que no exista una falsa alarma, y si bien es cierto que la señal de alarma se debe a una carta de control también lo es el hecho que está señal sucede posteriormente al punto de cambio. Las primeras cartas de control estadístico, las de Shewhart, no fueron diseñadas para detectar el punto de cambio, esta situación fue atendida por el CPA modificando y creando nuevas cartas de control estadístico que atendieran esta situación, la CUSUM es una de ellas. La afirmación anterior no implica que esta sea la única herramienta del CPA, éste incluye más herramientas (pruebas de hipótesis, estimadores bayesianos entre otros). El CPA atiende el problema de estimar el punto de cambio desde tres enfoques:

- Estadística Clásica: Se supone conocida la función de distribución que sigue un proceso y en base a ello se procede a estimar el punto de cambio. Las técnicas de Estimador de Máxima Verosimilitud (MLE, Maximum Likelihood Estimator) y de Regresión Lineal (LR, Linear Regression) son propias de este enfoque.
- Estadística bayesiana: Utiliza las distribuciones a posteriori de los datos con el fin de hacer inferencias acerca del Punto de Cambio.
- Estadística No-Paramétrica: Utiliza transformaciones para mapear los datos en distribuciones particulares conocidas. Asume que los datos no presentan un comportamiento que pueda ser definido por una función de distribución.

2.3 Revisión de literatura del Análisis de Punto de Cambio

El CPA utiliza los parámetros que caracterizan el fenómeno en estudio para estimar el momento exacto, tan preciso como sea posible, en que éstos han cambiado. Los modelos del CPA varían de acuerdo a los casos que se estudian dadas las características de los mismos. Existen técnicas del CPA con enfoque clásico o bayesiano y desde luego herramientas que contemplan regresiones, cartas de control de auto-inicio, redes neuronales, etc. Es un tema que ha sido aplicado con éxito en el control estadístico de la variabilidad de los procesos, y especialistas de otras áreas han tomado sus resultados, los han adaptado, y han efectuado sus propios desarrollos para presentar diferentes métodos de implementación. Los que han trabajado con las cartas de control estadístico, saben que la evaluación de éstas requiere de una cantidad considerable de cálculos y que la implementación de las mismas es vista en función de la facilidad que ofrecen para llevarlas a la práctica. De acuerdo a la clasificación dada en la sección anterior sobre los enfoques con que el CPA atiende el problema de estimar el punto de cambio; a continuación se hace una breve descripción de los hallazgos más importantes en la literatura (a juicio del autor).

2.3.1 Análisis Bayesiano

La literatura del análisis bayesiano acerca del problema del punto de cambio se enfoca en estimar el cambio en la función de distribución de una serie de observaciones ordenadas en el tiempo basándose para ello en una muestra de tamaño fijo. Este tipo de análisis es llamado “retrospectivo o fuera-de-línea”; a diferencia del análisis en línea, como lo hace las cartas de control donde cada observación es considerada para tomar una decisión conforme es recolectada, el análisis retrospectivo requiere que primero sea definida y/o recolectada

totalmente la muestra en la que se va a aplicar esta herramienta. Si bien el análisis bayesiano no presenta tantos trabajos en el CPA como el análisis clásico, este enfoque es el primero que atiende el problema del punto de cambio en 1952 con Girshick y Rubin [8] quienes sientan las bases para resolver el problema del CP desde este punto de vista. En 1963 Shiryaev [22] presenta una opción para estimar el punto de cambio considerando que se tiene información a priori acerca de la función de probabilidad que caracteriza el modelo. Un trabajo que muestra el enfoque bayesiano basado en la probabilidad a posteriori de los posibles puntos de cambio aplicado a la distribución binomial y normal es presentado por Smith en 1975 [23] y cinco años después publica una aplicación sobre la aplicación de este enfoque en conjunto de datos sobre un trasplante renal [24]. En 1990 Gombay y Hobart [25] desarrollan trabajos sobre la prueba de máxima verosimilitud (MLT) donde estudian las propiedades asintóticas del mismo y posteriormente presentan una aplicación al punto de cambio [26]. Sobre la utilización de este enfoque en las funciones de distribución discretas se encuentra el trabajo de Raferty y Akman [27] quienes lo aplican a la función de distribución de Poisson. Recientemente en 2010 Lai y Xing [28] presentan un enfoque bayesiano para el CP cuando los parámetros de la función de distribución son desconocidos antes y después del cambio.

2.3.2 Análisis Clásico

El análisis clásico del problema del punto de cambio implica el uso de herramientas de la estadística descriptiva e inferencial para realizar las estimaciones del punto de cambio. En este enfoque se considera conocida la función de distribución de los datos aun cuando los parámetros, de la misma, no lo sean. Se puede decir que uno de los iniciadores del CPA desde el enfoque clásico es Page [4] quien en 1954 presenta una carta de control basada en la suma acumulada de observaciones, la CUSUM, la cual presenta una mayor sensibilidad a los cambio pequeños y sostenidos que las anteriores carta de control de Shewhart, en 1961 el mismo Page [29] presenta un artículo donde discute el futuro y las aplicaciones de estas cartas de control. Por su parte en 1959 Roberts [7] presenta la carta de control EWMA cuya característica es que los últimos datos tienen una influencia mayor en la decisión de si un proceso está en control estadístico o no, para ello asigna un mayor peso a las últimas observaciones. Estas cartas son las primeras en estimar el punto de cambio, ya que al ser más sensibles el momento en que indican que el proceso sale de control estadístico se considera una estimación por sí misma. La forma en que la CUSUM consigue esto es definiendo un valor conocido como umbral k en función a la media y haciendo un conteo de los últimos valores que superan ese umbral, mientras que la carta EWMA maneja un parámetro (λ) exponencial el cual da más peso a las observaciones más recientes provocando que los valores anteriores tengan un peso menor. En 1970 Hinkley [30] sienta las bases para utilizar el MLE para determinar el punto de cambio lo cual implica el uso de derivadas para determinar el valor máximo o mínimo. El concepto de utilizar sumas acumuladas vuelve a aparecer en 1987 con Hawkins [5] con una modificación llamada auto-inicio cuya ventaja es que la carta de control será aplicable directamente al proceso en la Fase I. En 1992 Nishima [31] compara las cartas de control estadístico EWMA, CUSUM y SMA desde una perspectiva del punto de cambio y concluye que no existe una diferencia significativa entre ellas para lo que fueron creadas, sin embargo recomienda las CUSUM basándose en un enfoque de punto de cambio. En 1997 Box y Luceño [19] hace un estudio que demuestra que las cartas de

control de Shewhart son óptimas en la detección de puntos aislados mientras que las CUSUM son mejores para la detección de cambios sostenidos. En 1998 Samuel, Pignatiello y Calvin [32] presentan una alternativa para determinar el punto de cambio a partir del uso del MLE para punto de cambio después de que una carta de control envía el mensaje que el proceso ha salido de control estadístico. En 2006 Shao, Hou y Wang [33] utilizan de manera secuencial una carta de control S y el MLE para punto de cambio al uso en procesos caracterizados por una función de distribución Gamma. En el 2010 el tema de auto-inicio es utilizado por Li, Zhang y Wang [34] para proponer una carta de control que indique cuando el proceso está en control estadístico monitoreando la media y la varianza simultáneamente. En 2011 Shao y Hou [35] utilizan una carta EWMA y el MLE para punto de cambio para estimar el punto de cambio de una distribución Gamma para el caso en que el parámetro de forma (α) cambia.

2.3.3 Análisis No-paramétrico

El enfoque No-paramétrico supone que la función de distribución no puede ser definida a priori ya que son los datos los que la determinan, esto implica una dificultad al momento de realizar pruebas para saber si el proceso que se está monitoreando se encuentra en control o no ya que los elementos de la estadística inferencial paramétrica ya no son válidos. Uno de los primeros en trabajar este enfoque fue Page en 1955 [36] proponiendo una prueba basada en la redefinición de una variable y la distribución binomial. Fueron Bhattacharyya y Johnson [37] los que iniciaron el desarrollo de los procedimientos no-paramétricos en la forma como los conocemos en nuestros días. Sus propuestas de estimación usaron la Prueba de Signo de Wilcoxon. Pettit [38] en 1979 trabajó con este enfoque para las funciones de distribución discretas Bernoulli y Binomial. Dentro de los trabajos realizados con el enfoque No-paramétrico el trabajo de Hinkley y Schechtman [39] en 1987 consistió en construir un procedimiento basado en una retoma de muestras aplicado al cambio en la media y su método fue aplicado al problema de los desbordamientos del Río Nilo [40]. Dos procedimientos generales para pruebas no paramétricas en el problema del punto de cambio son presentadas por Koziol [41] en 1987. Guverich y Vexler [42] revisan las políticas de punto de cambio retrospectivas y proponen nuevos procedimientos. Hovarth y Kokoszka [43] utilizan la regresión no paramétrica para determinar el punto de cambio. Una aplicación interesante en la estadística no paramétrica es la distribución del error y la detección del punto de cambio en el mismo, este problema es atendido por Neumeyer y Van Keilegom [44] en 2009. Finalmente Tercero [13] analiza el punto de cambio en problemas definidos por caminatas aleatorias.

2.4 Resumen Histórico del SPC y del CPA

EL SPC y el CPA han tenido un crecimiento acelerado en las últimas décadas, tanto en el desarrollo de herramientas como en el número de aplicaciones. Este crecimiento es presentado por Pérez, Cordero y Tercero [45] mediante estadísticas descriptivas. Algunas de ellas han sido reproducidas aquí con el fin de dar una idea del desarrollo del CPA y SPC

La Figura 4 muestra la forma en que se ha incrementado el número de artículos sobre CPA en revistas especializadas. De esta gráfica se puede observar que a partir de los 80's el número de artículos publicados creció de manera significativa. El autor considera que este crecimiento fue

debido a la facilidad de disponer de poder de cómputo por parte de la gente involucrada en el SPC y el CPA, ya que los cálculos que se requieren y el volumen de información que se maneja, implica un esfuerzo manual considerable. La gráfica inicia en 1970 ya que Pérez-Blanco et al [45] encontraron que previo a ese año los trabajos publicados anteriormente fueron significativamente pocos. Una posible razón para explicar el crecimiento acelerado de la publicación de artículos a partir de los 80's es el poder de computo individual, que también empezó a tener un crecimiento importante en esos años, y que auxilia en los cálculos que requiere el CPA para hacer sus estimaciones del punto de cambio y que suelen ser cuantiosos y algo complejos.

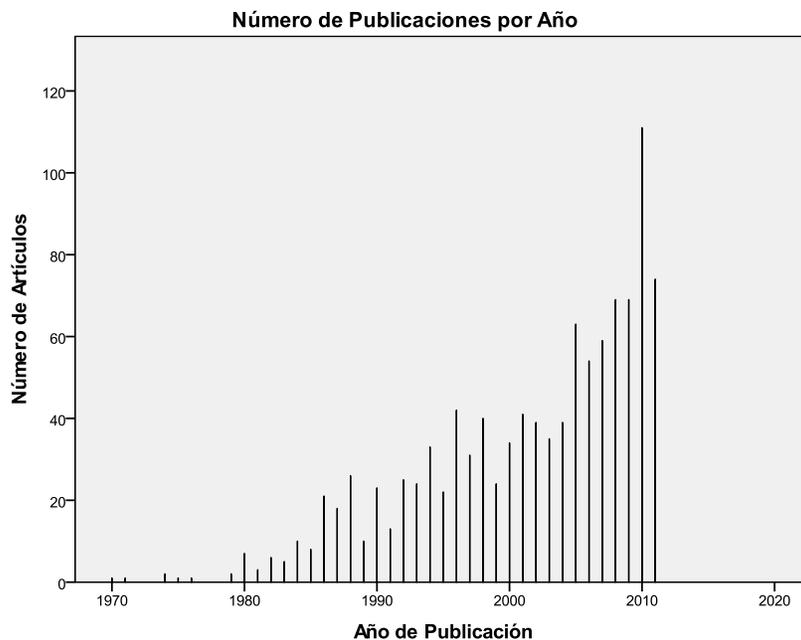


Figura 4 Gráfica del número de artículos publicados sobre CPA en revistas especializadas vs año de publicación.

La Figura 5 presenta una gráfica que incluye las publicaciones de investigación (barra clara) y las de aplicación (barra obscura) a partir de la cual se puede apreciar que las primeras superan a las segundas en una relación de 3 a 1. El periodo considerado por Pérez-Blanco et al [45]. De nueva cuenta se considera el periodo de 1970 a 2001.

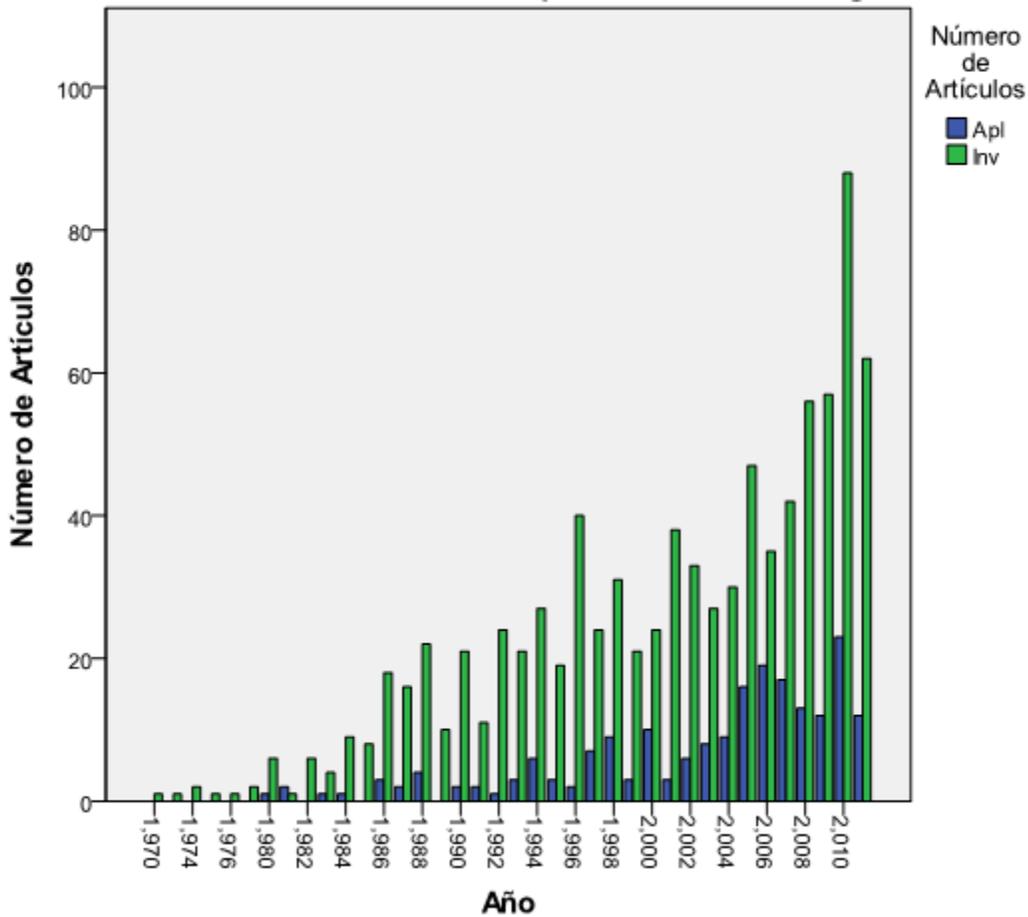


Figura 5 Gráfica comparativa entre las publicaciones de investigación y aplicación

Las Figuras 6 y 7 indican la cantidad de artículos publicados a partir de los 70's por revista. Se puede observar que al menos 4 revistas, de las 10 con mayor cantidad de artículos publicados, no son exclusivas del área de estadística y que las revistas especializadas ya cuentan con más de 20 artículos publicados, en el periodo que abracaron los autores de las gráficas (1970 -2011).



Figura 6 Número de publicaciones.

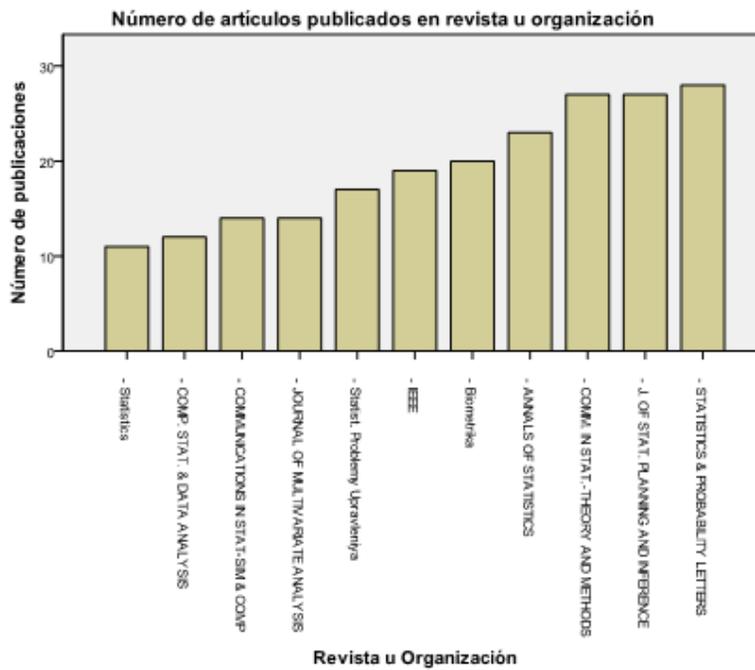


Figura 7 Número de publicaciones.

Complementando las gráficas anteriores se presenta la siguiente tabla que muestra un resumen histórico del SPC y el CPA señalando los artículos más relevantes, a juicio del autor, de los mismos en los diferentes enfoques, haciendo énfasis en los que han utilizado técnicas combinadas de carta de control y MLE en la estimación del punto de cambio.

Tabla 1 Resumen histórico del SPC y del CPA

Año	Autor y aportación	Enfoque al CPA
1931	Shewhart presenta sus resultados en control estadístico de la calidad [3]	Clásico
1952	Girshick y Rubin sientan las bases para resolver el problema de CP con un enfoque de Bayes [8]	Bayesiano
1954	Page presenta la carta de control CUSUM [29]	Clásico
1955	Page propone una prueba no paramétrica para estimar el CP [36]	No-paramétrico
1959	Roberts presenta una carta que da ponderación a las observaciones EWMA [7]	Clásico
1963	Shiryayev hace un estudio completo sobre métodos de detección rápida y establece una solución óptima para el caso de información completa a priori de una función de distribución [22]	Bayesiano
1964	Chernoff y Zacks analizan el problema del CP para una distribución normal desde un enfoque de Bayes y definen un estadístico de prueba T_n [46]	Bayesiano
1966	Kander y Zacks analizan las características operativas del estadístico T_n [47]	Bayesiano
1968	Bhattacharyya y Johnson utilizan la Prueba de Signo de Wilcoxon [37]	No-paramétrico
1970	Hinkley sienta las bases matemáticas para utilizar el MLE en el CPA [30]	Clásico
1975	Smith ofrece un enfoque de Bayes al problema del CP basado en la Binomial y la Normal y presenta un procedimiento secuencial informal para detectarlo [48]	Bayesiano
1979	Pettit presenta técnicas no-paramétricas para las funciones Bernoulli y Binomial	No-paramétrico
1981	Hawkins sugiere estandarizar las cantidades para obtener una mejor sensibilidad a la varianza [49]	Clásico
1984	Nelson establece una serie de reglas para incrementar la sensibilidad de la detección basadas en características de la carta de datos y no en la inferencia formal [50]	Clásico
1987	Hawkins presenta la CUSUM de auto-inicio [5]	Clásico
1987	Koziol presenta dos procedimientos generales para pruebas no paramétricas en el CPA [41]	Bayesiano
1987	Hinkley y Schechtman construyen un procedimiento basado en la retoma de muestras [39]	No-paramétrico
1990	Gombay y Horvath estudian las propiedades asintóticas de las pruebas usando MLT para cambios en la media y dan razones de porque son efectivos [25]	Clásico
1991	Quesenberry sugiere cartas de control utilizando funciones de distribución Binomial e Hipergeométrica. Publica 3 artículos en ese año [51] [52] [53]	Clásico
1993	MacGregor y Harris presentan la cartas Exponentially Weighted Mean Squared Deviation (EWMS) [54]	Clásico
1994	Gombay y Horvath presentan una aplicación del MLT al problema del punto de cambio [26]	Bayesiano
1995	Lai analiza el CP en los sistemas dinámicos [55]	Clásico
1995	Nikiforov considera por primera vez el problema de detectar múltiples	Clásico

	puntos de cambio [56]	
1996	Gombay y Horvarth prueban que el MLT es asintóticamente normal si hay un cambio en los parámetros en un momento desconocido [57]	Bayesiano
1996	Sullivan y Woodall proponen una carta de control basada en una prueba del cociente de máxima verosimilitud (LRT) [58]	Bayesiano
1997	Box y Luceño demuestran que las cartas de control de Shewhart son óptimas en la detección de puntos aislados mientras que las CUSM son mejores para la detección de cambios sostenidos [19]	Bayesiano
1998	Hawkins y Olwell presentan estudios sobre el monitoreo de la media y la varianza usando CUSUM [59]	Clásico
1998	Acosta presenta un estudio sobre la reducción de la variabilidad usando el rango [60]	Clásico
1998	Samuel, Pignatiello y Calvin utilizan el MLE para determinar el punto de cambio en un proceso normal después de que la carta de control asociada indica que está fuera de control [32] [61]	Clásico
1998	Timmer, Pignatiello y Longnecker utilizan el análisis de regresión para el CP [62]	Clásico
1999	Acosta y Pignatiello realizan comparación entre diferentes cartas de control para monitorear la dispersión en un proceso [14]	Clásico
2000	Nedumaran, G., Pignatiello, J.J., Calvin, J.A. utilizan el MLE para determinar el punto de cambio en combinación con cartas de control de χ^2 [63]	Clásico
2000	Lai generaliza el problema de detección de CP múltiples en el caso de procesos estocásticos [64]	Bayesiano
2001	Dabye y Kutoyans hacen una estimación del punto de cambio para un proceso Poisson [65]	Bayesiano
2001	Kumar y Wu presentan un procedimiento integrado para la detección del CP basado en lógica difusa [66]	Bayesiano
2001	Pignatiello y Samuel consideran el uso del MLE para el punto de cambio en lugar del estimador construido para la carta CUSUM y/o EWMA [67]	Clásico
2002	Sullivan y Jones desarrollan una carta de control con auto-inicio para observaciones individuales multivariadas, sin necesidad de observaciones previas, para estimar los parámetros iniciales [68]	Clásico
2002	Horvarth y Kokoszka utilizan la regresión no-paramétrica para determinar el punto de cambio [43]	No-paramétrico
2002	Nedumaran, Pignatiello y Calvin utilizan cartas de control basadas en χ^2 para detectar el punto de cambio [63]	Clásico
2003	Gombay considera una situación general en la detección y estimación del CP y propone un estimador sesgado del mismo [69]	Bayesiano
2004	Park y Park proponen un estimador de máxima verosimilitud para el caso de una distribución normal con parámetros iniciales desconocidos junto con cartas de control para \bar{X} y S [70]	Clásico
2005	Hawkins y Zamba crean cartas de control utilizando el GLR [71]	Clásico
2005	Ramanayake presenta un estimador para la Gamma cuando se tiene un cambio en el parámetro de la forma [72]	Clásico
2006	Shao, Hou y Wang combinan la carta de control de S y el MLE para determinar el CP en una función de distribución Gamma [33]	Clásico

2007	Costa y Machado proponen el estadístico VMAX basado en la varianza estandarizada de las muestras para p características de calidad para construir una carta de control multivariada [73]	Clásico
2007	Hawkins y Mauboudou crean una carta de control con auto-inicio multivariada con promedio móvil exponencialmente ponderado [74]	Clásico
2007	Schechtman, Bandner y Meginy comparan métodos estadísticos y de SPC para monitorear la varianza en un proceso. [75]	No-paramétrico
2007	Zou crea una carta de control con auto-inicio basada en el monitoreo recursivo de los residuos [76]	No-paramétrico
2007	Zou, Zhou y Wang presentan una carta de control con auto-inicio para datos con perfil lineal [77]	Clásico
2009	Neumeyer y Van Keilegom utilizan la distribución del error en el CPA [44]	No-paramétrico
2010	Brodsky, B. presenta un método secuencial para determinar el CP utilizando el CUSUM no-paramétrico y una prueba de Kolmogorov – Smirnov [78]	No-paramétrico
2010	Lai y Xing presentan un enfoque de Bayes para el CP cuando los parámetros, antes y después del momento de cambio, son desconocidos [28]	Bayesiano
2010	Guverich y Vexler revisan las políticas de punto de cambio retrospectivas y proponen nuevos procedimientos [42]	No-paramétrico
2010	Li, Zhang y Wanga desarrollan una carta de control con auto-inicio que controla simultáneamente la media y la varianza cuando los parámetros iniciales son desconocidos. [34]	Clásico
2010	Chen y Gupta publican el primer libro sobre análisis de punto de cambio [79]	Bayesiano y Clásico
2011	Maboudou y Hawkins proponen una carta de control multivariada con auto-inicio para un vector de medias, una matriz de covarianza y ambos al mismo tiempo. [80]	Clásico
2011	Shao y Hou utilizan de la EWMA y el MLE para estimar el punto de cambio en una función de distribución Gamma para el caso en que el parámetro de forma cambia. [35]	Clásico
2012	Tercero y Temblador utilizan la mediana para trabajar el PCA en estadística no-paramétrica [13]	No-paramétrico
2012	Cordero desarrolla el MLE para procesos normales con parámetros desconocidos [81]	Clásico
2012	Pérez, Cordero y Tercero realizan un estudio sobre las publicaciones concernientes al SPC y CPA entre 1931 y 2011 [45]	Bayesiano, Clásico y No-paramétrico
2012	Tercero desarrolla una carta de control y un estimador para el punto de cambio en la varianza basado en los p valores usando la prueba F [82]	Clásico
2013	Moshe y Krieger presentan un artículo donde prueban que la carta de control de Shewhart es óptima para la detección del punto de cambio bajo ciertas características.	Clásico

La Tabla 1 no es exhaustiva pero presenta los artículos que iniciaron esta área de estudio y los que han sido relevantes, a juicio del autor y considerando que esta investigación se basa en el

enfoque clásico, marcando la dirección en que los diferentes investigadores del mundo han atendido el problema del punto de cambio.

Finalmente es importante comentar que el mismo concepto en que se basa el CPA ha permitido que incursione en una variedad muy amplia de ciencias, desde la geología (terremotos) hasta médica (estudios de VIH) o incluso económica (valores bursátiles). El hecho que los fenómenos y/o procesos que estudia varíen a lo largo del tiempo, en forma continua o discreta ayuda a generar esta diversidad de aplicaciones. Aun cuando es una línea de investigación relativamente joven y que inició con pocas publicaciones, a partir de los años 80's muestra un crecimiento exponencial, tanto en investigación como en aplicaciones, manteniendo una relación de 3 a 1 entre ellas. Es de esperarse que en el futuro cercano siga siendo de interés y las publicaciones de investigación y aplicación no disminuyan.

2.5 Comentarios Generales

Para atender el problema del CP en el SPC lo ideal sería conocer la respuesta a dos preguntas ¿cómo detecto un punto de cambio? y ¿cómo estimo el punto de cambio? En la búsqueda de respuestas a estas interrogantes se han desarrollado las cartas de control permitiendo monitorear un proceso y creando estimadores, entre otras metodologías. A los pocos años de que Shewhart inicia la revolución del SPC surgen los primeros intentos de responder ambas preguntas; es una inquietud que fue trabajada y analizada en los 50's en donde se empieza a considerar la oportunidad de utilizar dos metodologías de manera secuencial para finalmente llegar a una que integre a ambas. Esto se logra mediante un proceso en dos pasos, primero se utiliza una carta de control hasta que manda una señal de que el proceso ha salido de control estadístico y entonces se retoma toda la información para aplicar un procedimiento único que resulta en una estimación del momento en que empezó a ocurrir ese cambio, el primer valor que ya no cumple con los valores iniciales del proceso, es importante tener en cuenta que se está considerando que la señal de la carta de control no es una falsa alarma. Como se comentó anteriormente, usualmente el valor en que la carta de control manda la señal indicando que el proceso salió de control estadístico es posterior al momento en que el punto de cambio ha ocurrido. Con la estimación del punto de cambio se auxilia a los encargados de los procesos en la búsqueda de la causa asignable que provocó que el proceso saliera de control estadístico ya que se busca que esta estimación sea cercana al momento en que realmente ocurrió y con ello tener la oportunidad de ahorrar en tiempo, esfuerzo y/o recursos. Esta oportunidad de reducir gastos es una de las razones por las que el CPA es útil. Para el caso particular de esta investigación se han considerado dos distribuciones de probabilidad que aparecen recurrentemente en fenómenos cotidianos. La función de distribución Normal tiene aplicaciones variadas y una de ellas puede ser consultada en [83] donde se describe un proceso normal cuya varianza cambia y que caracteriza a los impulsos eléctricos que se mandan a los músculos, para los autores de la referencia es importante estimar el momento en que el cambio

ocurre. Respecto a la función Gamma su aplicación puede ser encontrada en situaciones donde el factor del tiempo entre eventos es relevante, tal es el caso de la referencia [84] la cual hace una investigación sobre la intensidad de los ciclones en la década de los 80's y cuyos resultados han ayudado a la predicción de estos fenómenos y su intensidad en años posteriores. Finalmente un elemento que resulta interesante es que, dado que la función de distribución Gamma ha sido menos estudiada que la Normal en el SPC, se presenta una oportunidad para saber si las metodologías utilizadas en la función de distribución Normal funcionarán para la función de distribución Gamma. Ese punto es atendido en este trabajo tomando la función de distribución Normal y aplicando de manera secuencial una carta de control y el MLE para el punto de cambio para analizar el desempeño del estimador del punto de cambio y analizar los resultados para posteriormente hacer lo propio con la función de distribución Gamma, aplicar la misma metodología y observar los resultados. Es importante señalar que el desempeño de los estimadores de punto de cambio propuestos es evaluado en base a si los mismos son útiles para lo que fueron creados y ello se hará considerando las estimaciones del punto de cambio obtenidas y comparándolas con las propuestas por otros métodos mediante la medición del sesgo y la desviación estándar de los mismos.

CAPÍTULO 3. INVESTIGACIÓN 1

Detección Secuencial y Estimación de Cambios Sostenidos en la Media de Series de Tiempo Usando una Carta de Control con Auto-Inicio y el MLE para Observaciones con Distribución Normal

En esta primera investigación considera la inspección de procesos que siguen una distribución Normal y en los cuales se ha considerado que ocurre un cambio en el parámetro correspondiente a la media (μ). La investigación supone una serie de observaciones que siguen una distribución Normal con un valor inicial de la media (μ_0) y de la desviación estándar (σ_0) y que en un momento específico (τ) sufre un cambio en la media (μ_1) pero la desviación estándar se mantiene con el mismo valor ($\sigma_0 = \sigma_1$). Se propone el uso secuencial de cartas de control CUSUM de auto-inicio con estimadores de máxima verosimilitud para el monitoreo de este proceso. Se realizan múltiples simulaciones para comparar el desempeño del estimador de máxima verosimilitud propuesto y el estimador de la carta de control de auto-inicio, basado en el sesgo y la desviación del estimador ($\hat{\tau}$) respecto al valor real τ .

Este artículo fue publicado en: *Quality and Reliability Engineering International*; publicado en línea: 24-May-2013; DOI: 10.1202/qre.1511; in press.

Detección Secuencial y la Estimación de Cambios Sostenidos en la Media a través de Series de Tiempo Usando la Carta de Control CUSUM de auto-inicio y MLE para Observaciones Normales

Víctor G. Tercero-Gómez, Ph. D., Ph. D.

Álvaro E. Cordero-Franco, Ángel S. Pérez-Blanco, M.C.,

Universidad Autónoma de Nuevo León- Facultad de Ciencias Físico Matemáticas

San Nicolás de los Garza, Nuevo León, México

y Ph. D. Alberto A. Hernández-Luna

Instituto Tecnológico y de Estudios Superiores de Monterrey

Monterrey, Nuevo León, México

La detección de una causa especial de variación y la estimación del momento en que ocurre son dos actividades importantes en cualquier estrategia de mejora de un proceso que este bajo control estadístico. La detección de cambios en el proceso puede ser hecho mediante el uso de las cartas de control. Una de estas es la carta de control CUSUM de auto-inicio, la cual fue creada para detectar pequeños cambios con la posibilidad de ser implementada sin una Fase I o conocimiento previo de los parámetros del proceso. Para estimar el momento inicial del cambio, un estimador del punto de cambio basado en la carta de control CUSUM puede ser usado, aunque debe señalarse que en experimentos recientes de análisis retrospectivo se ha encontrado que el MLE correspondiente tiene un menor sesgo y error estándar. Este documento propone el uso secuencial de la carta de control CUSUM de auto-inicio y el MLE del punto de cambio en series de observaciones normales e independientes. El desempeño es estudiado mediante simulaciones Monte Carlo; mostrando que el uso del MLE reduce el sesgo de la estimación del punto de cambio. También es demostrado que al tomar observaciones extras, después de que el cambio es detectado, mejora la estimación del punto de cambio.

Palabras clave: Estimación del punto de cambio; CUSUM de auto-inicio; MLE; parámetros desconocidos; distribución Normal.

3.1 Introducción

Para vigilar un sistema, el control estadístico de procesos (SPC) ofrece las cartas de control. Éstas ayudan al monitoreo del comportamiento de observaciones de procesos estocásticos, los cuales son necesarios para detectar la aparición de causas especiales de variación y, al mismo tiempo, ayudar a los analistas en identificar las causas raíz que pueden ser corregidas con el fin de mejorar el sistema que se está estudiando. Para detectar cambios pequeños y sostenidos, las

cartas de control de sumas acumuladas (CUSUM) han sido reconocidas como herramientas efectivas del control de calidad. Esta carta de control fue originalmente creada por Page [4], y varios autores han estudiado su desempeño y extendido sus aplicaciones más allá de los supuestos de series de datos independientes provenientes de una distribución Normal y un conocimiento previo de los parámetros de la función de distribución de probabilidad. Para evitar la necesidad de la última suposición, Hawkins [5] desarrolló la carta de control CUSUM de auto-inicio para la ubicación y la escala, bajo los supuestos de un proceso normalmente distribuido. Estas cartas de control son relativamente buenas en la detección de cambios sostenidos mientras que para estimar el punto de cambio, los estimadores del punto de cambio necesitan ser usados.

Conocer el momento exacto de un cambio ayuda en el proceso de buscar causas especiales de variación, reduciendo el tiempo y costo asociado con la búsqueda. La estimación es parte de lo que se conoce como el problema del punto de cambio. Varios estimadores del punto de cambio han sido desarrollados para cuando han ocurrido cambios en la media o en la varianza de un conjunto de datos que pueden ser modelados por una función de distribución Normal. Samuel et al. [32] desarrolló el estimador de máxima verosimilitud (MLE), para el cambio en procesos normales con cambios en la media con parámetros iniciales conocidos. Usando el mismo enfoque, Cordero et al. [81] desarrolló el MLE para procesos normales con parámetros desconocidos, considerando tres posibles escenarios: un cambio en la media, un cambio en la varianza y un cambio en ambos parámetros. Estos estimadores atienden la misma clase de problemas que la carta CUSUM de auto-inicio, sin un conocimiento previo de los parámetros, por lo tanto, parece apropiado utilizar ambas herramientas de manera conjunta.

Esta investigación está enfocada en el uso secuencial de la carta CUSUM de auto-inicio para la detección de un cambio y la MLE para la estimación del punto de cambio en el momento en el que el cambio ocurre en un proceso normal con parámetros desconocidos. Esta “integración” deriva en un tiempo real de control estadístico capaz de detectar y estimar el punto de cambio en procesos normales donde hay un desconocimiento acerca de los valores de los parámetros iniciales. Las siguientes secciones del artículo están organizadas de la manera siguiente: la Sección 2 revisa la literatura de los estimadores del punto de cambio CUSUM y MLE; la Sección 3 describe las dos herramientas y presenta una estrategia para integrarlas secuencialmente; la Sección 4 muestra e interpreta el desempeño de los resultados de la experimentación Monte Carlo de la técnica propuesta; y finalmente, la Sección 5 ofrece una conclusión acerca de los resultados teóricos y su aplicación incluyendo posibles futuros trabajos en el campo del SPC.

3.2 Revisión Literaria

3.2.1 CUSUM y las Cartas de control de auto-inicio

Las cartas de control CUSUM fueron creadas originalmente por Page [4]. Él desarrolló una carta de control que tiene más poder que la carta de control Shewhart para detectar cambios sostenidos en la media y varianza poblacional. Cuando un cambio ocurre en el proceso, aun si es menor (1.5 derivaciones estándares o menores), el efecto del cambio es acumulado a través de una suma acumulada de desviaciones del valor objetivo, haciendo la carta de control CUSUM más sensible a pequeños cambios sostenidos que los enfoques tradicionales, los cuales no toman en cuenta la historia del proceso. Teóricamente los resultados de las propiedades de CUSUM del monitoreo de la media y la varianza en el proceso pueden ser encontrados en Page [36] [21], Hawkins y Olwell [59], Acosta Mejía [60], Acosta Mejía, Pignatiello y Rao [14].

La carta de control CUSUM asume un conocimiento de los parámetros, i.e. la carta de control CUSUM es usada para la Fase II. Para tratar con los problemas propios de la Fase I varias cartas de control con auto-inicio han sido desarrolladas, usando estimadores para parámetros desconocidos. Hawkins [5] creó una carta de control CUSUM de auto-inicio para detectar cambios en la media y en la varianza de los parámetros para procesos normales con parámetros desconocidos, usando la media móvil y la desviación estándar móvil como sustitutos para la media y la desviación estándar desconocidas. La experimentación muestra un buen desempeño aun cuando se compara con el caso en el cual los parámetros son conocidos. Un enfoque similar fue propuesto por Quesenberry [53], él desarrolló varias cartas de control del tipo Shewhart para la media y la varianza de series de observaciones normales independientes donde no hay información previa de los parámetros. Las cartas de control fueron creadas para diferentes combinaciones de parámetros conocidos o no conocidos. Quesenberry [51] sugirió el uso de una carta de control para el parámetro desconocido p (casos favorables entre casos posibles) que define a la función de distribución Hipergeométrica como un estimador insesgado de mínima varianza de la función de distribución binomial. El método converge en la función Binomial conforme el número de muestras se incrementa. También trabaja con un tamaño de muestra variable. Quesenberry [52] consideró dos casos, donde el parámetro de la distribución Poisson es conocido y desconocido. Él usó las cartas de control normalizadas conocidas como cartas de control Q que podrían ser usadas con cartas de control Z.

Sullivan y Jones [68] desarrollaron una carta de control con auto-inicio para observaciones individuales multivariadas sin la necesidad de observaciones previas para la estimación de los parámetros iniciales, usando la derivación de cada vector de observaciones del promedio de todas las observaciones previas y graficando estas derivaciones en una carta de control T^2 . Hawkins y Maboudou-Tchao [74] crearon una carta de control con auto-inicio multivariada con promedio móvil exponencialmente ponderada. Esta carta de control transforma los parámetros

desconocidos del vector del proceso en vectores con parámetro conocido de la misma dimensión. Esta carta tiene las mismas propiedades de control, tal como si el valor verdadero de la media del proceso y la matriz de covarianzas fueran perfectamente conocidas. Más tarde, Zou et al. [85] creó una carta de control con auto-inicio basada en el monitoreo de los residuos recursivos cuando los parámetros del proceso son desconocidos. La carta de control propuesta puede ser usada para detectar cambios en la pendiente, la intercepción o la desviación estándar. Recientemente, Li, Zhang y Wang [34] desarrollaron una carta de control con auto-inicio que controla simultáneamente la media y la varianza cuando los parámetros iniciales son desconocidos. Esta carta de control está basada en la prueba de cociente de verosimilitud y en el procedimiento del promedio móvil exponencialmente ponderado. Finalmente, Maboudou-Tchao y Hawkins [80] propusieron una carta de control multivariada con auto-inicio para un vector de medias, una matriz de covarianza, y ambos al mismo tiempo. Éste sigue el enfoque de la carta de control Múltiple EWMA integrando un proceso studentizado para incorporar la estimación del error cuando se definen los límites del control.

3.2.2 Análisis del Punto de Cambio dentro del SPC

El análisis del punto de cambio estudia el método de detección y estimación de cambios sostenidos en series de tiempo creando pruebas y estimadores respectivamente. Los primeros en ofrecer una forma de detectar el punto de cambio fueron Girshick y Rubin [8] haciéndolo desde un punto de vista Bayesiano, y después Page [4] desde la perspectiva clásica de la estadística. La historia del SPC y el análisis del punto de cambio se intersectan dentro del desarrollo de cartas de control para cambios sostenidos. Hinkley [30] construyó los MLEs y las pruebas de cociente de verosimilitud para detectar puntos de cambio. Usando esta metodología se han desarrollado varios MLEs para el punto de cambio para diferentes distribuciones con parámetros iniciales conocidos, algunos pueden ser consultados en Dabye y Kuroyants [65], Nedumaran, Pignatiello y Calvin [63], Pignatiello y Samuel [67], Piagnatello y Samuel [86], Samuel, Pignatiello y Calvin [61] [32], y Timmer y Pignatiello [62]. Estos estimadores fueron hechos considerando conocido el parámetro inicial, lo cual no siempre es posible en aplicaciones de la vida real. Cordero et al. [81] extendió el alcance del MLE mediante la obtención y evaluación de estimadores del punto de cambio, usando esta herramienta, cuando los parámetros iniciales son desconocidos en una serie de datos independientes provenientes de un proceso Normal.

Para detectar si el cambio ocurre, Timmer, Pignatiello, y Longnecker [62], Zamba y Hawkins [87], y Batsidis [88] han desarrollado pruebas LR para diferentes condiciones en una serie de tiempo. Hawkins y Zamba [10] propusieron un enfoque diferente para tratar con el análisis del punto de cambio para la varianza en un proceso normal con parámetros iniciales desconocidos. Ellos desarrollaron una carta de control usando el radio de verosimilitud general (GLR) basado en las pruebas estadísticas de Bartlett. Sin embargo, este modelo crece en complejidad

conforme más observaciones eran añadidas dentro de las mismas series de tiempo. Reynolds y Jianying [89] desarrollaron una carta de control GLR para cambios en la media en un proceso normal con los parámetros iniciales conocidos, usando una ventana móvil para tratar con la complejidad del GLR para series de tiempo grandes. Tercero et al. [82] desarrolló una carta de control y un estimador para el punto de cambio en la varianza basado en los $p - valor$ de la prueba F para un proceso normal en la Fase I usando el enfoque del GLR y la ventana móvil. Estas cartas de control fueron desarrolladas desde el enfoque del punto de cambio y pueden ser consideradas con auto-inicio, ya que no requieren un análisis de la Fase I. Usando un enfoque similar, pero para la Fase II, Cordero et al. creó una carta de control basado en los $p - valor$ de una prueba cuadrada Chi cuadrada para monitorear cambios en la varianza de un proceso normal.

3.3 Modelo

El modelo propuesto usa el CUSUM de auto-inicio para detectar cambios sostenidos durante el monitoreo del proceso. Una vez que el cambio es detectado, el punto de cambio τ es estimado usando el MLE para un conjunto de muestras independientes normales $X_1, \dots, X_\tau, X_{\tau+1}, \dots, X_T$ como se muestra en la ecuación de abajo (1)

$$X_i \sim \begin{cases} N(\mu_0, \sigma_0), & 1 \leq i \leq \tau \\ N(\mu_1, \sigma_1), & \tau < i \leq T \end{cases} \quad (1)$$

, , y μ_0, μ_1, σ_0 y σ_1 son desconocidas, y T es la primera muestra fuera de control detectada por el CUSUM de auto-inicio. X_i es un vector de varias observaciones de tamaño m donde cada elemento $X_{i,j}$ sigue la distribución presentada anteriormente. En la Sección 2.1 se revisa el CUSUM de auto-inicio, en la Sección 2.2 se describe el punto de cambio obtenido por el MLE, y en la Sección 2.3 se muestra la metodología propuesta.

3.3.1 CUSUM de auto-inicio

El CUSUM de auto-inicio desarrollado por Hawkins [5] es una carta de control para procesos con parámetros iniciales desconocidos. Una estrategia usada en el SPC para tratar con los parámetros desconocidos consiste en la aplicación de la Fase I, donde el proceso es analizado y sus parámetros son estimados. Sin embargo, cada estimación implica errores debido a la falta de certeza. Estas estimaciones erróneas se convierten en sesgo cuando se implementa una carta de control CUSUM o cualquier otra carta de control basada en parámetros estimados, y pueden conducir a problemas indeseables para mantener la estabilidad del proceso. Mediante la modificación de la carta de control CUSUM original para considerar los errores de estimación y la actualización de información, nuevas cartas, llamadas CUSUM de auto-inicio, fueron desarrolladas para detectar errores en ubicación y escala. Estas cartas no necesitan una Fase I y pueden ser implementadas de inmediato en cualquier proceso que pueda ser modelado como

una serie de observaciones normales independientes. Mediante el control de la distribución del error, el error Tipo I se disminuye y el poder para detectar cambios muestra un desempeño aceptable.

Para implementar la carta de control CUSUM de auto-inicio, las ecuaciones recursivas (2) y (3) son usadas para estimar los parámetros del proceso conforme la muestra i es observada. La primera ecuación corresponde a la media muestral, la segunda ecuación es la suma de cuadrados muestral actualizados al momento i . Las ecuaciones (4) y (5) son usadas para las desviaciones estudentizadas de las estimaciones previas de la media del proceso.

$$\bar{X}_i = \bar{X}_{i-1} + (X_i - \bar{X}_{i-1})/i \quad (2)$$

$$S_i = S_{i-1} + (i - 1)(X_i - \bar{X}_{i-1})^2/i \quad (3)$$

$$T_i = a_i(X_i - \bar{X}_{i-1}) / \sqrt{S_{i-1}/(i - 2)} \quad (4)$$

Donde,

$$a_i = \sqrt{(i - 1)/i} \quad (5)$$

Fue mostrado en Hawkins (1969) que los residuales estudentizados T_i 's son independientes y siguen a una función de distribución t con $i - 2$ grados de libertad. Para evitar recalcular constantemente los límites de control - conforme los grados de libertad se incrementan, la forma de la función de distribución t cambia - T_i es transformada en una variable aleatoria U_i cuya función de distribución es cercana a la de la Normal.

$$U_i = \left(\frac{8(i-2)+1}{8(i-2)+3} \right) (i-2) \ln \left(1 + \frac{T_i^2}{i-2} \right) \quad (6)$$

U_i se distribuye aproximadamente como una $N(0,1)$. Esta aproximación fue obtenida por Wallace [90] y la exactitud fue medida por Preizer y Partt [91]. U_i puede ser considerada como un residual estudiantizado, reestandarizada para distribuirse aproximadamente como una $N(0,1)$. La secuencia de U_i consiste en un conjunto de variables aleatorias independientes cuya función de distribución es aproximadamente $N(0,1)$, donde cada valor mide la desviación de la X correspondiente a el promedio de sus predecesores.

Para construir los intervalos de control, dos cartas CUSUM son creadas de la manera usual. La que muestra la ecuación (7) controla cambios positivos, conocida como CUSUM hacia arriba, la otra ecuación (8) controla los cambios negativos, conocida como CUSUM hacia abajo. Ambas inician en 0, como se muestra en la ecuación (9):

$$L_i^+ = \max(0, L_{i-1}^+ + U_i - k) \quad (7)$$

$$L_i^- = \max(0, L_{i-1}^- + U_i + k) \quad (8)$$

$$L_0^+ = L_0^- = 0 \quad (9)$$

En este caso, k es una cantidad permitida por observación, de tal forma que los movimientos de X_i menores a lo permitido no son registrados. El parámetro h es una decisión del límite de control. Los valores absolutos de L_i^+ o L_i^- mayores que h indican una señal fuera de control. Si una señal es dada en L_i^+ , entonces X_i se incrementa. Si la señal es dada en L_i^- , entonces X_i se decrementa. Observe que las primeras dos observaciones no pueden ser usadas para calcular T_j y U_j . Solo cuando la tercera observación está disponible una desviación estándar no trivial es posible por lo que $i - 2 > 0$. En consecuencia, la carta de control CUSUM de auto-inicio inicia desde la tercera observación de la serie de datos.

Los valores de h y k son obtenidos mediante la decisión de qué tamaño de cambios de la media se desean detectar para obtener el mejor desempeño. De acuerdo a Lucas y Crosier [92], al parámetro k usualmente se le asigna el valor correspondiente a la mitad del cambio en el proceso lo cual es importante para detectar el cambio rápidamente, mientras que el valor de h es escogida para obtener un ARL adecuado. La ARL (por sus siglas en inglés) o PLC (Promedio de la Longitud de la Corrida) en español, es el número promedio de puntos que deben graficarse antes de que un punto indique una condición fuera de control. De acuerdo a Hawkins [59], esta carta de control CUSUM es construida usando observaciones independientes $N(0,1)$ y tendrá el control ARL exacto de las propiedades de la información normal con media y varianza conocidas. Por ejemplo, Montgomery [6] indica que, en la práctica, un valor de $k = \frac{1}{2}$ es muy usado. Combinado con un valor $h = 4$, el ARL_0 es 168, y para $h = 5$, el ARL_0 se convierte en 465. Por otro lado, Hawkins [93] dijo que los valores de k y h deberían producir un $ARL_0 = 370$ para un CUSUM bilateral, y propone varias combinaciones de los valores h y k para este propósito. Ya que no hay un estándar absoluto, los autores de esta investigación siguieron las sugerencias de Montgomery [6] al momento de evaluar el desempeño de la carta de control con el punto de cambio MLE para series de observaciones normales independientes.

3.3.2 Estimación del Punto de Cambio

Usando la teoría desarrollada por Hinkley [30], considere un proceso normal donde un cambio ocurre en la media entre el tiempo τ y $\tau - 1$. Ahora el proceso presenta la misma función de distribución pero con parámetros diferentes a partir del momento τ , la primera con parámetros desconocidos μ_0 y σ_0^2 ; y la segunda con parámetros desconocidos μ_1 y σ_0^2 . La función de verosimilitud es

$$L(t, \mu_0, \mu_1, \sigma_0 | x) = \prod_{i=1}^{\tau} \prod_{j=1}^m f(t, \mu_0, \sigma_0 | x_{i,j}) \prod_{i=\tau+1}^T \prod_{j=1}^m f(t, \mu_1, \sigma_0 | x_{i,j}) \quad (10)$$

Con el fin de facilitar el trabajo de la obtención de las parciales y el álgebra involucrada se ha utilizado el logaritmo natural de la función de verosimilitud, el cual es

$$L(t, \mu_0, \mu_1, \sigma_0 | x) = \prod_{i=1}^{\tau} \prod_{j=1}^m f(t, \mu_0, \sigma_0 | x_{i,j}) \prod_{i=\tau+1}^T \prod_{j=1}^m f(t, \mu_1, \sigma_0 | x_{i,j}) \quad (11)$$

Con el fin de encontrar el MLE para el tiempo τ es necesario encontrar el MLE para los tres parámetros desconocidos μ_0 , μ_1 y σ^2 , y ; y después, encontrar el tiempo t que maximice las ecuaciones (10) u (11). Esto es hecho usando los siguientes estimadores en las ecuaciones (12-15).

$$\hat{t}_{MLE} = arg \min_{2 \leq t \leq T-2} \{\hat{\sigma}_p\} \quad (12)$$

Donde

$$\hat{\sigma}_p^2 = \frac{\sum_{i=1}^t \sum_{j=1}^m (x_{i,j} - \hat{\mu}_0)^2 + \sum_{i=t}^{T-t} \sum_{j=1}^m (x_{i,j} - \hat{\mu}_1)^2}{mT} \quad (13)$$

$$\hat{\mu}_0 = \frac{\sum_{i=1}^t \sum_{j=1}^m x_{i,j}}{m\tau} \quad (14)$$

$$\hat{\mu}_1 = \frac{\sum_{i=\tau+1}^T \sum_{j=1}^m x_{i,j}}{m(T-\tau)} \quad (15)$$

Puede ser visto en la ecuación (12) no es solo un MLE sino también un estimador de mínimos cuadrados.

Por otro lado, cuando una carta de control CUSUM es construida, un estimador natural del punto de cambio puede ser creado [6]. Un analista puede seguir el número de casos L^+ , $|L^-| > 0$ definiendo dos variables adicionales llamadas N_i^+ y N_i^- . Una estimación del punto de cambio

es construido como se muestra en la ecuación (16) por la resta N_i^+ y N_i^- del tiempo T de las muestras que salen de control. Estas variables son un contador de la cantidad de L_i^+ y L_i^- diferentes de cero que se encuentran previas al momento en que un punto sale de los límites de control (T) y posteriores al último cero obtenido en cada uno de los valores de L_i^+ y L_i^- . Al realizar este procedimiento y restar los valores al momento T lo que se obtiene es una estimación del punto de cambio, el cual resulta congruente con la idea de que el punto de cambio es previo al momento en que la carta de control manda una señal de alarma.

$$\hat{t}_{CUSUM} = \min \{T - N_T^+, T - N_T^-\} \quad (16)$$

La principal ventaja de los últimos estimadores es la facilidad para calcularlos. De cualquier forma, como se muestra en la Sección 3.4, este posee un sesgo negativo que debe ser considerado cuando se implementa.

3.3.3 Análisis del Modelo Secuencial del Punto de Cambio

Como se ve en la Figura 8, el CUSUM de auto-inicio y el \hat{t}_{MLE} son usados secuencialmente para detectar cuando un cambio sostenido ocurre y estimar el momento en que inició. El proceso de monitoreo empieza con una carta de control CUSUM. Tan pronto como la muestra fuera de control es señalada, esta muestra es llamada T, y la ecuación (12) puede ser usada para estimar el punto de cambio. Si la precisión necesita ser incrementada, muestras adicionales w deben ser obtenidas y la estadística $T' = T + w$ debe ser usada en la ecuación (12) en vez de solamente T

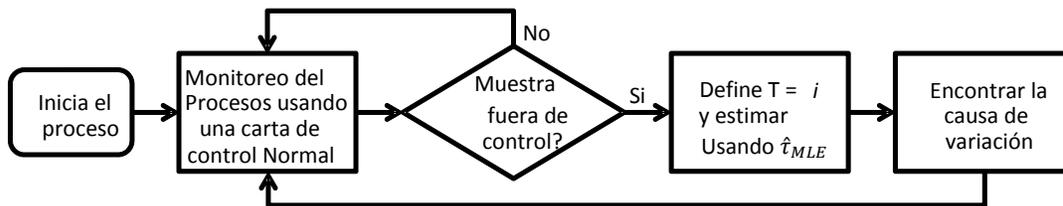


Figura 8 Detección secuencial del punto de cambio y estimación con la CUSUM de auto-inicio y el MLE del punto de cambio para series de observaciones normales independientes

Ambos, el CUSUM de auto-inicio y el \hat{t}_{MLE} , asumen una serie de observaciones normales independientes sin un conocimiento previo acerca de los parámetros y su homocedasticidad. En la siguiente sección se evalúa el desempeño del \hat{t}_{MLE} cuando es aplicado en este modelo y compara sus resultados con el \hat{t}_{CUSUM} y el Average Run Length (ARL) del CUSUM de auto-inicio en diferentes escenarios.

3.4 Desempeño de los Estimadores

3.4.1 Diseño de la Experimentación

Para evaluar el desempeño de la integración de la carta de control CUSUM de auto-inicio con el MLE para el punto de cambio, se han establecido diferentes escenarios los cuales son indicados en la Tabla 2. Es importante señalar que estos escenarios consideran que no existe una señal de alarma (falsa o no) previa al momento en que se ha establecido el punto de cambio (τ).

Tabla 2 Factores para medir el desempeño de la integración del CUSUM de auto-inicio y el MLE

Factores	Niveles
Estimadores	\hat{t}_{MLE} , \hat{t}_{CUSUM} , ARL
τ	25, 50, 75, 100
Cambio en la media (δ)	0, 0.25σ , 0.5σ , 0.75σ , 1σ , 1.5σ , 2.0σ , 3.0σ
Tamaño del subgrupo (m)	1, 3, 5
Datos extra después de que el cambio es detectado (w)	0, 5, 10, 15, 20

Usando la experimentación Monte Carlo para evaluar cada escenario presentado en la Tabla 2 los siguientes pasos fueron seguidos:

1. Marcar un punto de cambio τ .
2. Definir $i = 1$
3. Mientras ($i \leq \tau$) Y (la CUSUM no envíe una señal de alarma) generar X_i con vectores aleatorios utilizando con la ecuación (1)
4. Si el CUSUM de auto-inicio envía una señal de alarma parar e ir al paso número 2, si no detecta cambios, ir al paso número 5.
5. Generar X_i vectores aleatorios utilizando la ecuación (1) hasta que la CUSUM de auto-inicio envíe una señal de alarma, al momento en que se envía la señal de alarma se le llamará T.
6. Estimar el punto de cambio \hat{t}_{MLE} y \hat{t}_{CUSUM} para los conjuntos de datos.
7. Calcular el error estimado $\hat{t}_{MLE} - \tau$ y $\hat{t}_{CUSUM} - \tau$
8. Repetir del paso número 2 al 4, hacerlo 10,000 veces.
9. Estimar el sesgo, el error estándar de \hat{t}_{MLE} y \hat{t}_{CUSUM} , y el ARL de la carta de control CUSUM de auto-inicio.
10. Calcular la probabilidad $P(|\hat{t}_{MLE} - \tau| \leq D)$, donde $D = \{0, 1, 2, \dots, 24\}$
11. Seleccionar otro escenario y regresar al paso número 1.

3.4.2 Resultados de la Experimentación

Para cambios en la media, la carta de control CUSUM de auto-inicio integrada con el MLE es evaluada en los escenarios descritos en la sección previa. La Tabla 3 muestra el desempeño para diferentes puntos de cambio y magnitud del cambio para el subgrupo con tamaño $m = 1$. Se observa que el ARL mejoró y las desviaciones y el error estándar del punto de cambio MLE se volvieron menores cuando la magnitud del cambio se incrementa o cuando el cambio ocurre alejado al punto inicial. En la Tabla 4 el desempeño es analizado cuando el cambio ocurre en $\tau = 50$ sobre diferentes tamaños de subgrupos y diferentes magnitudes de cambio. Los estimadores de estimación mejoran cuando el tamaño de los subgrupos se vuelve mayor.

En las Tablas 5 y 6 se han considerado w datos extras después de que el CUSUM de auto-inicio detecta un cambio con el fin de analizar si el MLE mejora su desempeño al tener más información del conjunto de datos, para simular este escenario se han considerado cambios en $\tau = 50$ y $\tau = 100$. El resultado es que se observa que el sesgo y el error estándar mejoran conforme el número de muestras extras, w , se incrementa. Sin embargo, este crecimiento en la precisión es provocado por no detener la simulación cuando la alerta es detectada, este hecho debe ser evaluado considerando el costo, en la vida real, de tener esta nueva información.

En la Tabla 7 el desempeño del estimador del punto de cambio del CUSUM de auto-inicio es evaluado bajo los mismos escenarios considerados en la Tabla 2. Se observa que el $\hat{\tau}_{CUSUM}$ tiene menos sesgo y menos error estándar que el $\hat{\tau}_{MLE}$ cuando ocurren pequeños cambios ($\delta < 1$). También el sesgo y el error estándar del $\hat{\tau}_{CUSUM}$ parecen disminuir con respecto a un cambio en la media, todo esto en una menor tasa que el sesgo y el error estándar del $\hat{\tau}_{MLE}$.

Finalmente en las Tablas 8 y 9 la probabilidad $P(|\hat{\tau}_{MLE} - \tau| \leq D)$ es calculada para $\tau = 100$ con w datos extras ($w = 0$ y $w = 20$ respectivamente). La probabilidad de detectar un punto de cambio tempranamente se incrementa cuando w se incrementa.

Tabla 3 Los estimadores de desempeño sobre diferentes puntos de cambio τ de la carta de control CUSUM de auto-inicio para ARL, el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y el error estándar asociado del punto de cambio usando $\hat{\sigma}(\hat{\tau}_{MLE})$, tamaño del subgrupo de 1 ($m = 1$). Los valores fueron calculados usando 10,000 repeticiones.

		δ							
τ	Métrica	0.00	0.25	0.50	0.75	1.0	1.5	2.0	3.0
25	ARL	111.644	94.039	51.063	22.062	10.881	5.628	4.077	2.759
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	98.745	80.350	36.856	10.081	1.699	-0.345	-0.409	-0.265
	$\hat{\sigma}(\hat{\tau}_{MLE})$	111.188	109.300	75.311	36.898	13.473	3.767	2.552	1.595
50	ARL	111.617	78.908	33.849	14.168	8.632	5.104	3.750	2.532
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	97.145	63.078	18.122	2.846	0.139	-0.653	-0.566	-0.328
	$\hat{\sigma}(\hat{\tau}_{MLE})$	113.291	95.497	49.708	16.079	7.590	4.789	3.503	2.207
75	ARL	114.016	71.736	27.083	12.601	8.218	4.982	3.623	2.475
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	99.225	53.974	11.476	1.387	-0.490	-0.785	-0.584	-0.294
	$\hat{\sigma}(\hat{\tau}_{MLE})$	115.369	87.458	35.581	13.299	9.641	6.100	3.852	2.087
100	ARL	112.779	67.935	23.921	11.899	7.854	4.854	3.567	2.424
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	97.348	49.790	8.465	0.333	-0.699	-1.016	-0.645	-0.389
	$\hat{\sigma}(\hat{\tau}_{MLE})$	113.473	83.919	27.686	15.376	10.580	7.853	4.430	3.546

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros τ datos. En esta tabla se han considerado los casos en que el punto de cambio τ toma los valores de 25, 50, 75 y 100. Los cambios en la media corresponden a $\mu_1 = \mu_0 + \delta\sigma_0$.

Tabla 4 El desempeño sobre muestras de diferente tamaño de subgrupo (m) de la carta de control CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y la estimación del error estándar asociado para la estimación del punto de cambio $\hat{\sigma}(\hat{\tau}_{MLE})$ cuando $\tau = 50$. Los valores fueron calculados usando 10,000 repeticiones.

		δ							
m	Métrica	0.00	0.25	0.50	0.75	1.0	1.5	2.0	3.0
1	ARL	111.617	78.908	33.849	14.168	8.632	5.104	3.750	2.532
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	97.145	63.078	18.122	2.846	0.139	-0.653	-0.566	-0.328
	$\hat{\sigma}(\hat{\tau}_{MLE})$	113.291	95.497	49.708	16.079	7.590	4.789	3.503	2.207
3	ARL	111.814	42.462	10.825	6.059	4.374	2.888	2.247	1.700
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	97.785	26.058	1.053	-0.583	-0.585	-0.363	-0.188	-0.029
	$\hat{\sigma}(\hat{\tau}_{MLE})$	113.022	57.751	10.240	5.640	4.150	2.296	1.418	0.421
5	ARL	112.348	25.803	7.375	4.499	3.332	2.298	1.847	1.434
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	98.086	11.227	-0.322	-0.623	-0.494	-0.236	-0.084	-0.005
	$\hat{\sigma}(\hat{\tau}_{MLE})$	115.189	34.590	6.752	4.247	3.167	1.694	0.722	0.089

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros $\tau = 50$ datos. En esta tabla se han considerado los casos en que el punto de cambio τ toma los valores de 25, 50, 75 y 100. Los cambios en la media corresponden a $\mu_1 = \mu_0 + \delta\sigma_0$.

Tabla 5 El desempeño de la carta de control CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y el error estándar asociado de la estimación del punto de cambio $\hat{\sigma}(\hat{\tau}_{MLE})$ cuando $\tau = 50$ y $m = 1$ utilizando diferentes valores de w datos extras.

w	Métrica	δ							
		0.00	0.25	0.50	0.75	1.0	1.5	2.0	3.0
5	ARL	115.080	79.773	33.349	14.076	8.654	5.105	3.742	2.545
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	84.752	50.533	11.301	0.778	-0.075	-0.212	-0.113	-0.030
	$\hat{\sigma}(\hat{\tau}_{MLE})$	116.186	91.337	40.642	12.889	7.068	3.314	1.598	0.570
10	ARL	112.916	80.789	33.049	14.195	8.668	5.125	3.717	2.535
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	75.577	44.517	8.958	1.100	0.064	-0.084	-0.009	-0.005
	$\hat{\sigma}(\hat{\tau}_{MLE})$	111.842	88.981	37.946	13.115	6.645	2.755	1.356	0.543
15	ARL	114.027	79.834	33.267	14.205	8.656	5.118	3.764	2.532
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	72.141	41.822	9.270	1.333	0.151	-0.012	0.002	-0.002
	$\hat{\sigma}(\hat{\tau}_{MLE})$	112.866	87.750	39.193	13.048	6.367	2.471	1.331	0.523
20	ARL	112.511	82.059	31.984	14.002	8.675	5.080	3.751	2.527
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	69.091	41.164	8.284	1.386	0.267	0.019	0.003	-0.009
	$\hat{\sigma}(\hat{\tau}_{MLE})$	110.566	87.179	37.096	13.159	6.077	2.342	1.266	0.518

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros $\tau = 50$ datos. En esta tabla se han considerado $w = 5, 10, 15$ y 20 datos extras después de que la carta de control CUSUM manda la señal de un valor fuera de control estadístico. Los cambios en la media corresponden a $\mu_1 = \mu_0 + \delta\sigma_0$.

Tabla 6 El desempeño de la carta de control CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\hat{\sigma}(\hat{\tau}_{MLE})$ cuando $\tau = 100$ y $m = 1$ (tamaño del subgrupo) utilizando diferentes valores de w datos extras.

w	Métrica	δ							
		0.00	0.25	0.50	0.75	1.0	1.5	2.0	3.0
5	ARL	113.756	67.167	24.231	11.979	8.035	4.897	3.590	2.449
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	78.173	32.732	3.005	-0.794	-0.732	-0.248	-0.086	-0.029
	$\hat{\sigma}(\hat{\tau}_{MLE})$	121.544	79.691	30.719	15.912	9.987	3.609	1.446	0.575
10	ARL	113.051	68.596	23.863	11.777	7.971	4.882	3.609	2.450
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	64.405	27.172	1.590	-0.820	-0.413	-0.101	-0.031	-0.002
	$\hat{\sigma}(\hat{\tau}_{MLE})$	119.498	82.015	29.299	14.391	8.439	2.988	1.348	0.514
15	ARL	113.096	69.559	23.743	11.950	7.971	4.907	3.614	2.436
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	59.595	25.338	1.779	-0.075	0.054	0.042	-0.019	-0.003
	$\hat{\sigma}(\hat{\tau}_{MLE})$	123.382	83.362	27.874	13.272	6.287	2.271	1.252	0.501
20	ARL	113.936	65.916	24.084	11.876	7.939	4.909	3.591	2.438
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	58.353	21.911	1.981	0.327	0.125	0.060	-0.001	-0.008
	$\hat{\sigma}(\hat{\tau}_{MLE})$	127.451	79.075	28.525	11.668	5.657	2.265	1.274	0.499

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros $\tau = 100$ datos. En esta tabla se han considerado 5, 10, 15 y 20 datos extras después de que la carta de control CUSUM manda la señal de un valor fuera de control estadístico. Los cambios en la media corresponden a $\mu_1 = \mu_0 + \delta\sigma_0$.

Tabla 7 El sesgo para la estimación del punto de cambio mediante el uso de la carta CUSUM de auto-inicio y del MLE, $\hat{E}(\hat{\tau}_{CUSUM} - \tau)$ y el error estándar asociado p cambio $\hat{\sigma}(\hat{\tau}_{CUSUM})$ para $m = 1$ utilizando diferentes valores de τ . Los valores fueron calculados usando 10,000 simulaciones.

τ	Métrica	δ							
		0.00	0.25	0.50	0.75	1.0	1.5	2.0	3.0
25	$\hat{E}(\hat{\tau}_{CUSUM} - \tau)$	44.971	26.320	4.468	-3.187	-4.574	-4.485	-4.483	-4.637
	$\hat{\sigma}(\hat{\tau}_{CUSUM})$	92.699	84.533	47.875	19.441	7.304	5.830	5.788	5.919
50	$\hat{E}(\hat{\tau}_{CUSUM} - \tau)$	46.441	14.035	-2.638	-4.817	-4.860	-4.949	-4.841	-4.976
	$\hat{\sigma}(\hat{\tau}_{CUSUM})$	97.412	68.363	25.351	10.399	7.028	7.248	7.155	7.161
75	$\hat{E}(\hat{\tau}_{CUSUM} - \tau)$	46.543	7.874	-4.288	-4.973	-4.847	-4.966	-5.008	-4.948
	$\hat{\sigma}(\hat{\tau}_{CUSUM})$	97.412	55.589	16.839	7.395	7.021	7.373	7.331	7.289
100	$\hat{E}(\hat{\tau}_{CUSUM} - \tau)$	46.827	4.326	-4.619	-4.900	-4.921	-5.020	-5.027	-4.887
	$\hat{\sigma}(\hat{\tau}_{CUSUM})$	97.669	49.120	8.550	7.363	7.212	7.322	7.455	7.172

Esta tabla presenta los resultados para la estimación del punto de cambio mediante el uso de la carta de control CUSUM de auto-inicio para los valores del punto de cambio $\tau = 25, 50, 75$ y 100 . Se considera un subgrupo de 1 observación, tamaño de muestra $m = 1$, los cambios en la media corresponden a $\mu_1 = \mu_0 + \delta\sigma_0$. Se realizaron 10,000 simulaciones para la obtención de los datos en esta tabla.

Tabla 8 Probabilidades de los diferentes valores de que el sesgo del estimador y el valor real sea menor a un rango de valores de $\{0, 1, 2, \dots, 24\}$ con $\tau = 100$, tamaño de subgrupo $m = 1$ y $w = 0$ datos extras.

	δ						
	0.25	0.50	0.75	1.0	1.5	2.0	3.0
$P(\hat{\tau}_{MLE} = \tau)$	0.026	0.078	0.160	0.259	0.431	0.598	0.809
$P(\hat{\tau}_{MLE} - \tau \leq 1)$	0.063	0.185	0.327	0.467	0.672	0.813	0.932
$P(\hat{\tau}_{MLE} - \tau \leq 2)$	0.099	0.265	0.441	0.596	0.790	0.892	0.962
$P(\hat{\tau}_{MLE} - \tau \leq 3)$	0.124	0.326	0.523	0.682	0.855	0.927	0.973
$P(\hat{\tau}_{MLE} - \tau \leq 4)$	0.152	0.381	0.588	0.748	0.893	0.945	0.980
$P(\hat{\tau}_{MLE} - \tau \leq 5)$	0.180	0.428	0.643	0.798	0.919	0.958	0.985
$P(\hat{\tau}_{MLE} - \tau \leq 6)$	0.203	0.467	0.687	0.834	0.936	0.966	0.988
$P(\hat{\tau}_{MLE} - \tau \leq 7)$	0.224	0.503	0.725	0.863	0.948	0.971	0.991
$P(\hat{\tau}_{MLE} - \tau \leq 8)$	0.246	0.536	0.756	0.886	0.956	0.976	0.992
$P(\hat{\tau}_{MLE} - \tau \leq 9)$	0.266	0.567	0.781	0.904	0.963	0.980	0.993
$P(\hat{\tau}_{MLE} - \tau \leq 10)$	0.285	0.594	0.806	0.921	0.968	0.983	0.994
$P(\hat{\tau}_{MLE} - \tau \leq 12)$	0.317	0.642	0.846	0.941	0.976	0.987	0.995
$P(\hat{\tau}_{MLE} - \tau \leq 14)$	0.349	0.680	0.877	0.956	0.980	0.992	0.996
$P(\hat{\tau}_{MLE} - \tau \leq 16)$	0.378	0.713	0.901	0.966	0.983	0.993	0.997
$P(\hat{\tau}_{MLE} - \tau \leq 18)$	0.403	0.743	0.919	0.970	0.985	0.994	0.997
$P(\hat{\tau}_{MLE} - \tau \leq 20)$	0.430	0.767	0.934	0.975	0.987	0.995	0.997
$P(\hat{\tau}_{MLE} - \tau \leq 22)$	0.453	0.790	0.943	0.978	0.988	0.996	0.998
$P(\hat{\tau}_{MLE} - \tau \leq 24)$	0.474	0.811	0.951	0.981	0.989	0.996	0.998

Tabla 9 Probabilidades de que el sesgo del estimador y el valor real sea menor a un rango de valores de $\{0, 1, 2, \dots, 24\}$ con $\tau = 100$, tamaño de subgrupo $m = 1$ y $w = 20$ datos extras.

	δ						
	0.25	0.50	0.75	1.0	1.5	2.0	3.0
$P(\hat{\tau}_{MLE} = \tau)$	0.026	0.090	0.173	0.283	0.466	0.628	0.854
$P(\hat{\tau}_{MLE} - \tau \leq 1)$	0.066	0.199	0.353	0.507	0.724	0.855	0.976
$P(\hat{\tau}_{MLE} - \tau \leq 2)$	0.101	0.287	0.470	0.637	0.840	0.936	0.995
$P(\hat{\tau}_{MLE} - \tau \leq 3)$	0.133	0.356	0.557	0.727	0.904	0.969	0.999
$P(\hat{\tau}_{MLE} - \tau \leq 4)$	0.160	0.411	0.626	0.784	0.937	0.986	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 5)$	0.182	0.461	0.680	0.834	0.960	0.994	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 6)$	0.206	0.498	0.728	0.870	0.975	0.997	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 7)$	0.228	0.534	0.766	0.895	0.986	0.999	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 8)$	0.249	0.565	0.799	0.913	0.991	0.999	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 9)$	0.268	0.594	0.823	0.930	0.993	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 10)$	0.287	0.619	0.843	0.943	0.995	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 12)$	0.323	0.661	0.876	0.961	0.998	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 14)$	0.353	0.700	0.900	0.974	0.999	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 16)$	0.378	0.726	0.919	0.981	0.999	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 18)$	0.408	0.752	0.932	0.987	1.000	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 20)$	0.433	0.775	0.944	0.991	1.000	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 22)$	0.454	0.796	0.954	0.995	1.000	1.000	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 24)$	0.475	0.813	0.962	0.995	1.000	1.000	1.000

3.5 Conclusiones y Trabajos Futuros

La CUSUM de auto-inicio usa una estimación recursiva de los parámetros mediante la implementación de medias y desviaciones estándar calculadas repetidamente en base a todas las observaciones disponibles en cada momento del proceso. Esto evita sesgos por encima y debajo de la estimación creada por la Fase I del SPC. Las cartas de control para la media tipo Shewhart son efectivas si la magnitud del cambio se encuentra en un rango de 1.5σ a 2σ o mayores [6] en tal caso las cartas de control CUSUM (\hat{t}_{CUSUM}) son una buena alternativa y permiten estimar el punto cambio; sin embargo el sesgo de estas estimaciones puede ser reducido usando el correspondiente estimador de máxima verosimilitud (\hat{t}_{CUSUM}). Una vez que la señal es enviada por la carta de control se sugiere usar el MLE para punto de cambio, en la práctica, en la presencia de cambios mayores o iguales a 0.75σ unidades de distancia de la media considerando un periodo de calentamiento, de la carta de control CUSUM, de 50 muestras tal como lo recomienda Hawkins [5].

CAPÍTULO 4. INVESTIGACIÓN 2

Detección Secuencial y Estimación de Cambios Sostenidos en la Varianza de Series de Tiempo Usando una carta de Control con Auto-Inicio y el MLE para Observaciones con Distribución Normal

Esta investigación retoma la inspección de procesos que siguen una distribución Normal considerando que ahora el parámetro que cambia es la desviación estándar, esto es, las condiciones de normalidad se siguen suponiendo solo que ahora mientras la media se mantiene constante ($\mu_0 = \mu_1$), antes y después del tiempo desconocido τ , la desviación estándar es la que cambia ($\sigma_0 \neq \sigma_1$). Se propone el uso secuencial de cartas de control CUSUM de auto-inicio con estimadores de máxima verosimilitud para el monitoreo de este proceso. Se realizan simulaciones para comparar el desempeño del estimador de máxima verosimilitud del punto de cambio y el estimador de la carta de control de auto-inicio. Basado en el sesgo y la desviación estándar del estimador ($\hat{\tau}$) respecto al valor real τ .

Este artículo será enviado para su publicación a: *Quality and Reliability Engineering International*

Detección Secuencial y Estimación de Cambios Sostenidos en la Varianza a través de Series de Tiempo Usando la Carta de Control CUSUM de auto-inicio y MLE para Observaciones con Distribución Normal

Víctor G. Tercero-Gómez, Ph. D., Ph. D.

Álvaro E. Cordero-Franco, Ángel S. Pérez-Blanco, M.C.,

Universidad Autónoma de Nuevo León- Facultad de Ciencias Físico Matemáticas
San Nicolás de los Garza, Nuevo León, México

Cuando un proceso es monitoreado por cartas de control y un cambio ocurre se espera recibir una señal de alarma a la brevedad. Una vez que la señal es recibida se inicia un procedimiento de búsqueda de la causa que la provocó. Esta actividad implica analizar las condiciones del proceso en y antes del momento en que la señal fue emitida debido a que la causa de variación pudo haber empezado antes dicha emisión. Disponer de herramientas que provean una buena estimación del momento exacto en que la causa se manifestó iniciando el cambio ayudaría a las personas encargadas del proceso en la búsqueda de la causa asignable con lo que tendrían un ahorro en tiempo, esfuerzo y dinero. El éxito de esta búsqueda les auxiliaría en decidir y realizar las actividades preventivas y/o correctivas que regresen al proceso a estar en control estadístico. En esta investigación se analiza el desempeño de un estimador obtenido a partir del uso secuencial de una carta de control CUSUM de auto-inicio y el MLE para el punto de cambio para procesos que son caracterizados por una función de distribución Normal y cuya varianza cambia a partir de un momento específico desconocido.

Palabras clave: Estimación del punto de cambio; CUSUM de auto-inicio; MLE; varianza; distribución Normal.

4.1. Introducción

Las cartas de control en el Control Estadístico de Procesos (SPC) son herramientas usadas para monitorear el estado de un proceso. Mantener una vigilancia sobre la variabilidad en la media y/o en la varianza de los mismos son una de sus principales razones de ser. Cuando la carta de control manda una señal indicando que el proceso ha salido de control estadístico es conveniente iniciar una búsqueda de la(s) causa(s) que provoca(n) esta situación. Debido a la potencial demora en que una carta de control envía la señal de alerta la estimación del momento en que el cambio ocurrió es de ayuda para el personal encargado de controlar y monitorear el proceso, sin duda ellos desean tener la mayor cantidad de información posible. Conocer el momento en que el cambio ocurrió ayuda en la búsqueda para identificar la causa

de variación más rápidamente y a la toma de las acciones apropiadas para volver a tener al proceso en control estadístico. Esta situación ha sido una de las motivaciones para que se hayan desarrollado procedimientos para estimar el momento en que el cambio ocurre, este momento es conocido como Punto de Cambio (CP por sus siglas en inglés). Alrededor de 1950 empiezan a aparecer artículos con propuestas sobre como estimar el punto de cambio desde diferentes enfoques: bayesiano, clásico, no paramétrico y con diferentes suposiciones: solo un cambio en un parámetro, varios cambios, varios parámetros, los cuales serán comentados en la siguiente sección. Desde un enfoque clásico dos herramientas propuestas y utilizadas son las cartas CUSUM creada por Page [29] y las EWMA (Exponentially Weighted Moving Average Chart) por Roberts [7] las cuales son comparadas por Nishina en [31] donde concluye que tienen desempeños casi equivalentes basándose en el ARL, si los parámetros son adecuadamente elegidos. Él recomienda la CUSUM desde el punto de vista de la estimación del punto de cambio. Samuel, Pignatiello y Calvin [32] propusieron usar la función de máxima verosimilitud del punto de cambio para cambios en la media utilizando cartas de control para \bar{X} mientras que los mismos autores en [94] proponen un estimador de la desviación estándar usando la función de máxima verosimilitud para cartas de control para la S . Hawkins [5] desarrolla una CUSUM de auto-inicio en la cual cada observación sucesiva es estandarizada usando la media y la desviación estándar previas, esto permite empezar a monitorear el sistema sin necesidad de conocer los parámetros iniciales. Una de las características de la CUSUM es que conforme las observaciones son tomadas y utilizadas en los cálculos de la misma, éstos convergerán a los valores reales [95] de la media y la desviación. La CUSUM de auto-inicio fue creada para detectar pequeños cambios con la posibilidad de ser implementada sin una Fase I o conocimiento previo de los parámetros del proceso. En esta investigación se considera el uso secuencial de la carta de control CUSUM de auto-inicio y el estimador de máxima verosimilitud para el punto de cambio para estimar el CP para un proceso caracterizado por una distribución Normal para la cual el parámetro que cambiará será σ . Esto es, antes del momento τ la función de distribución que define al proceso es $N(\mu_0, \sigma_0)$ y posterior a él es $N(\mu_0, \sigma_1)$ donde $\sigma_0 \neq \sigma_1$.

Para la función de distribución Normal se han desarrollado diferentes estimadores, tanto para cambios en la media como en la varianza, en particular Cordero et al. [81] en 2012 desarrollaron los estimadores para cambios en la media, varianza y ambos, para el caso de parámetros iniciales desconocidos. Debido a que estos estimadores son obtenidos a partir de consideraciones similares a las que se tiene para las cartas de control CUSUM de auto-inicio parece natural combinar ambas herramientas para desarrollar un nuevo estimador.

La forma en que esta investigación ha considerado utilizar ambas herramientas es mediante un uso secuencial considerando primero una carta de control que señale el momento en que un proceso está fuera de control estadístico para inmediatamente proceder con la estimación del punto de cambio mediante el uso de un estimador de máxima verosimilitud correspondiente,

todo esto considerando un proceso caracterizado por una función de distribución Normal con parámetros desconocidos. Este uso secuencial de dos herramientas del SPC aprovecha la naturaleza de las cartas de control, que analizan dato por dato en tiempo real y se complementa con la naturaleza del MLE para el punto de cambio cuyo análisis es necesariamente fuera de línea, ya que utiliza un conjunto fijo de datos para determinar el estimador.

El desarrollo de la investigación incluye la Sección 4.2 de revisión de la literatura que refiere principalmente a los estimadores de punto de cambio obtenidos por una carta de control CUSUM y por el MLE para el punto de cambio; la Sección 4.3 presenta la forma en que se propone combinar ambas herramientas; la Sección 4.4 detalla los resultados obtenidos a partir de una simulación Monte Carlo donde se analizan los resultados de aplicar la metodología propuesta considerando diferentes escenarios; finalmente la Sección 4.5 presenta las conclusiones basadas en los resultados obtenidos de las simulaciones así como posibles áreas de aplicación y trabajos futuros dentro del contexto del SPC.

4.2. Revisión Literaria

Las cartas de control de Shewhart [3] fueron pioneras en monitorear la variabilidad en un proceso, con el tiempo fueron apareciendo nuevas cartas de control con características que las hacían más sensibles a los cambios en la media y/o la varianza, tal es el caso de las cartas CUSUM desarrolladas por Page [4]. Estas cartas permiten la detección de cambios en el proceso cuando estos se presentan de manera sostenida y pueden ser tan pequeños como 1.5 desviaciones estándar [6]. Estudios sobre el desempeño y propiedades de estas cartas pueden ser consultados en Page [56] y [21], Hawkins y Olwell [96], Acosta [60] y Acosta, Pignatiello y Rao [14].

Debido a que la carta de control CUSUM supone parámetros conocidos es útil para la Fase II del SPC, sin embargo para la Fase I se han desarrollado cartas CUSUM de auto-inicio que salvan el obstáculo de no tener información sobre los parámetros. Esta carta de control es desarrollada por Hawkins [5] en 1987 y posteriormente Quesenberry [53] desarrolla cartas de control para series de datos que se caracterizan por una distribución Normal para monitorear la media y la varianza y en donde no se cuenta con información a priori de los parámetros. Una aportación importante al monitoreo de la varianza, y la media, la hacen Li, Zhang y Wanga en 2010 [34] al presentar una carta de control de auto-inicio que permite monitorear la media y la varianza simultáneamente cuando los parámetros iniciales son desconocidos. Por su parte las cartas de control tipo EWMA, desarrolladas por Roberts [7] son también de gran utilidad en detectar la variación en un proceso al considerar la historia del proceso. Estas cartas tienen la característica de dar ponderación a los datos con la finalidad de que todos tengan un peso específico sobre la decisión a tomar sobre si las observaciones están dentro o fuera de control estadístico.

Las herramientas mencionadas anteriormente representan una parte de lo que compone el SPC, todas ellas tienen como finalidad determinar cuándo un proceso está en control estadístico o cuando no lo está. Sin embargo adolecen de un punto, no todas son capaces de estimar el instante exacto en que el cambio en los parámetros de la función que caracteriza a la muestra de observaciones ha iniciado, para ello es necesario recurrir a otras técnicas. El término Análisis de Punto de Cambio ha sido acuñado para definir a los métodos de detección y estimación de cambios sostenidos en series de tiempo mediante la creación de pruebas y estimadores para el punto de cambio.

El problema del punto de cambio ha sido analizado desde diferentes perspectivas, Girshick y Rubin [8] lo hicieron desde un enfoque bayesiano mientras que Page [21] utilizó la estadística clásica, una aportación importante fue el uso del MLE y las pruebas de cociente de verosimilitud para detectar el punto de cambio aportado por Hinkley [30]. Este hecho sentó las bases para crear estimadores para diferentes distribuciones con parámetros iniciales conocidos y/o desconocidos y en los últimos años algunos investigadores han direccionado sus esfuerzos en este sentido. Algunos de ellos son Nedumaran, Pignatiello y Calvin [63] y Pignatiello y Samuel [67]. En otro frente, Hawkins y Zamba [10] propusieron una carta de control basada en el cociente de verosimilitud general (GLR) basado en las pruebas estadísticas de Bartlett, Tercero et al. [82] crearon una carta de control, basada en los p -valores de una prueba de Chi Cuadrada para monitorear cambios en la varianza de un proceso caracterizado por una función de distribución Normal.

4.3 Modelo

El proceso es caracterizado por un conjunto de muestras independientes normales $X_1, \dots, X_\tau, X_{\tau+1}, \dots, X_T$ como se muestra en la ecuación (1)

$$X_i \sim \begin{cases} N(\mu_0, \sigma_0), & 1 \leq i \leq \tau \\ N(\mu_1, \sigma_1), & \tau < i \leq T \end{cases} \quad (1)$$

τ , μ_0 , σ_0 y σ_1 son desconocidos, y T es el primer valor fuera de control de la muestra detectado por la carta de control CUSUM de auto-inicio. X_i es un vector de observaciones de tamaño m donde cada elemento $X_{i,j}$ sigue una distribución como la indicada. La Sección 4.3.1 revisa la carta de control CUSUM de auto-inicio, la Sección 4.3.2 describe el enfoque de punto de cambio usando el MLE, la Sección 4.3.3 presenta un análisis del modelo de punto de cambio de la metodología propuesta y la Sección 4.4 detalla el diseño de la experimentación.

4.3.1 CUSUM de auto-inicio

Las cartas de control CUSUM de auto-inicio desarrolladas por Hawkins [5] son útiles cuando se sabe que el proceso es caracterizado por una distribución Normal y sus parámetros iniciales son

desconocidos, estas cartas no requieren de la Fase I del SPC y la única suposición es la Normalidad del proceso, como se indicó. Su implementación implica la estimación de los parámetros del proceso mediante el uso de ecuaciones recursivas las cuales son detalladas en (2) y (3), con la primera se va calculando el valor de la media muestral y la segunda es una forma de calcular la desviación estándar cuyos valores se actualizan con cada muestra que se toma en el momento i . Las ecuaciones (4) y (5) son usadas para calcular desviaciones estudentizadas de las estimaciones anteriores. Estas ecuaciones son formulaciones recursivas del estadístico de prueba usado en la prueba t para dos muestras independientes.

$$\bar{X}_i = \bar{X}_{i-1} + (X_i - \bar{X}_{i-1})/i \quad (2)$$

$$S_i = S_{i-1} + (i - 1)(X_i - \bar{X}_{i-1})^2/i \quad (3)$$

$$T_i = a_i(X_i - \bar{X}_{i-1}) / \sqrt{S_{i-1}/(i - 2)} \quad (4)$$

Donde,

$$a_i = \sqrt{(i - 1)/i} \quad (5)$$

Hawkins [97] demostró que los residuales estudentizados T_i 's son independientes y siguen una función de distribución t con $i - 2$ grados de libertad. Con el fin de evitar calcular los límites de control conforme los grados de libertad se incrementan la función de distribución t cambia por lo que T_i es transformada en una variable aleatoria U_i cuya función de distribución se aproxima a la Normal.

$$U_i = \left(\frac{8(i-2)+1}{8(i-2)+3} \right) (i-2) \ln \left(1 + \frac{T_i^2}{i-2} \right) \quad (6)$$

La U_i tiene una distribución aproximada a una $N(0,1)$. Esta aproximación fue obtenida por Wallace [90] y la exactitud fue medida por Preizer y Partt [91]. U_i puede ser considerada como un residual estudiantizado, reestandarizada para distribuirse aproximadamente como una $N(0,1)$, donde cada valor mide la desviación de la X correspondiente de sus predecesores.

La otra ecuación involucrada en los cálculos es:

$$V_i = (\sqrt{\text{abs}(U_i)} - 0.822) / 0.349 \quad (7)$$

Hawkins [98] sugiere que estos nuevos valores son más sensibles a los desplazamientos en la varianza que a los de la media. Montgomery [6] indica que los V_i son sensibles a los desplazamientos tanto de la media como de la varianza, debido a que la distribución de las V_i ,

bajo control es aproximadamente $N(0,1)$ [49], es posible establecer los límites para la CUSUM de la desviación estándar mediante las siguientes ecuaciones:

$$S_i^+ = \max(0, S_{i-1}^+ + U_i - k) \quad (8)$$

$$S_i^- = \max(0, S_{i-1}^- + U_i + k) \quad (9)$$

$$S_0^+ = S_0^- = 0 \quad (10)$$

La ecuación (8) monitorea los cambios positivos y es conocida como CUSUM hacia arriba y la otra ecuación (9) monitorea los cambios negativos y es conocida como CUSUM hacia abajo. . Ambas empiezan en 0, como lo indica la ecuación (10):

En este caso, k es una cantidad permitida por observación, de tal forma que los movimientos de X_i menores a lo permitido no son registrados. El parámetro h es una decisión del límite de control. Los valores absolutos de S_i^+ o S_i^- mayores que h indican una señal fuera de control. Si una señal es dada en S_i^+ , entonces X_i se incrementa. El contador se decrementa si la señal es dada por S_i^- . Observe que las primeras dos observaciones no pueden ser usadas para calcular T_j, U_j y V_j . Solo cuando la tercera observación está disponible se puede calcular una desviación estándar no trivial, esto implica que $i - 2 > 0$. En consecuencia, una carta de control CUSUM de auto-inicio empieza desde la tercera observación del conjunto de datos.

Los criterios para seleccionar los valores de h y k son semejantes a los que se siguen para seleccionar los valores para el caso del CUSUM para la media en donde se sugiere un h de 5 veces la σ del proceso mientras que la k es la mitad de la distancia ente las medias [6]. El hecho de estandarizar la CUSUM ofrece dos ventajas [6] :

- a) La primera ventaja es que muchas cartas CUSUM pueden tener ahora los mismos valores de k y h , y la elección de estos parámetros no depende de σ ;
- b) La segunda ventaja es que una CUSUM estandarizada lleva de manera natural a una CUSUM para controlar la variabilidad.

Bajo las consideraciones anteriores y siguiendo las recomendaciones de Lucas y Crosier [92], y a la estandarización a una $N(0,1)$ se usarán los valores de $k = \frac{1}{2}$ y $h = 5$ para realizar la simulación de casos.

4.3.2 Estimación del Punto de Cambio

Basados en la teoría desarrollada por Hinkley [30] se ha considerado un proceso Normal que presenta un cambio en la desviación estándar entre el tiempo $\tau - 1$ y τ . Esta suposición generará dos procesos, el primero con parámetros desconocidos μ_0 y σ_0 ; y el segundo con parámetros desconocidos μ_0 y σ_1 . La función de verosimilitud está definida por

$$L(t, \mu_0, \sigma_0, \sigma_1 | x) = \prod_{i=1}^t \prod_{j=1}^m f(t, \mu_0, \sigma_0 | x_{i,j}) \prod_{i=\tau+1}^T \prod_{j=1}^m f(t, \mu_0, \sigma_1 | x_{i,j}) \quad (11)$$

Es bien conocido que el utilizar la función logaritmo natural simplifica la obtención de derivadas y facilita el trabajo algebraico por lo que se utilizará en la función de máxima verosimilitud obteniendo

$$\ln L = \sum_{i=1}^t \sum_{j=1}^m \ln \left(\frac{1}{\sqrt{2\pi}\sigma_0} \exp \left(-\frac{(x_{i,j} - \mu_0)^2}{2\sigma_0^2} \right) \right) + \sum_{i=\tau+1}^T \sum_{j=1}^m \ln \left(\frac{1}{\sqrt{2\pi}\sigma_1} \exp \left(-\frac{(x_{i,j} - \mu_0)^2}{2\sigma_1^2} \right) \right) \quad (12)$$

Con el fin de encontrar el MLE para el tiempo τ es necesario encontrar el MLE para los tres parámetros desconocidos σ_0 , σ_1 y τ ; y después de ello encontrar el momento que maximiza a ecuación (11) o (12). Esto es hecho por el uso de las siguientes ecuaciones (13-15).

$$\hat{t}_{MLE} = \arg \min_{2 \leq t \leq T-2} \{ \hat{\sigma}_1^{m\tau} \hat{\sigma}_2^{m(T-\tau)} \} \quad (13)$$

donde

$$\sigma_0^2 = \frac{\sum_{i=1}^{\tau} \sum_{j=1}^m (x_{i,j} - \mu_0)^2}{m\tau} \quad (14)$$

$$\sigma_1^2 = \frac{\sum_{i=\tau+1}^T \sum_{j=1}^m (x_{i,j} - \mu_0)^2}{m(T-\tau)} \quad (15)$$

Siguiendo a Montgomery [6] al construir una carta de control CUSUM es posible obtener un estimador natural del punto de cambio. Un analista puede seguir el número de cambios S^+ , $|S^-| > 0$ definiendo dos variables adicionales llamadas N_i^+ and N_i^- . Una estimación del punto de cambio es construido como se muestra en la ecuación (16) por la resta N_i^+ y N_i^- del tiempo T de las muestras que salen de control. Estas variables son un contador de la cantidad de S_i^+ y S_i^- diferentes de cero que se encuentran previas al momento en que un punto sale de de los límites de control (T) y posteriores al último cero obtenido en cada uno de los valores de S_i^+ y S_i^- . Al realizar este procedimiento y restar los valores al momento T lo que se obtiene es una estimación del punto de cambio, el cual resulta congruente con la idea de que el punto de cambio es previo al momento en que la carta de control manda una señal de alarma.

$$\hat{t}_{CUSUM} = \min \{ T - N_T^+, T - N_T^- \} \quad (16)$$

La principal ventaja de los últimos estimadores es la facilidad para calcularlos.

4.3.3 Análisis del Modelo del Punto de Cambio Secuencial

En la Figura 9 se detalla el uso secuencial del CUSUM de auto-inicio y el \hat{t}_{MLE} para detectar cuando un cambio sostenido ocurre y estimar el momento en que inició. El monitoreo del proceso inicia con una carta de control CUSUM la cual envía una señal cuando una muestra fuera de control es encontrada, esta muestra es llamada T , y se utiliza la ecuación (12) para estimar el punto de cambio. Si se considera necesario incrementar la precisión se recomienda obtener muestras adicionales w y la estadística $T' = T + W$ deberá ser considerada en la ecuación (12) en lugar de únicamente T .

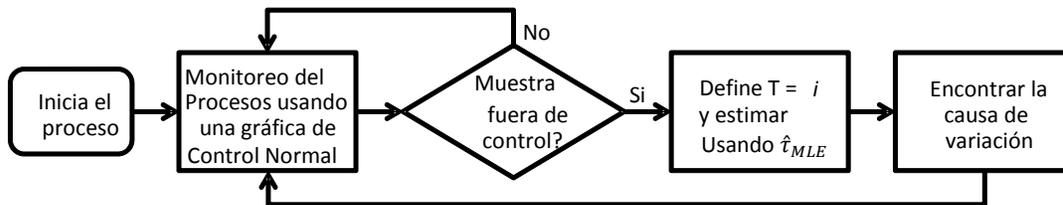


Figura 9 Detección secuencial del punto de cambio y su estimación con el CUSUM de auto-inicio y el MLE del punto de cambio para series de observaciones normales independientes.

Es importante señalar que se ha considerado una serie de observaciones normales independientes cuyos parámetros iniciales son desconocidos así como su homocedasticidad, lo cual aplica para el CUSUM de auto-inicio y para el \hat{t}_{MLE} . En la siguiente sección se consideran diferentes escenarios para evaluar el desempeño del \hat{t}_{MLE} obtenido por este modelo y compararlo con los resultados obtenidos por el \hat{t}_{CUSUM} y el Average Run Length (ARL) del CUSUM de auto-inicio.

4.3.4 Diseño de la Experimentación

Para evaluar el desempeño de la integración de la carta de control CUSUM de auto-inicio con el MLE, se han considerado diferentes escenarios los cuales son detallados en la Tabla 10.

Tabla 10 Factores para medir el desempeño de la integración del CUSUM de auto-inicio y el MLE.

Factores	Niveles
Estimadores	$\hat{\tau}_{MLE}$, $\hat{\tau}_{CUSUM}$, ARL
τ	25, 50, 75, 100
Cociente σ_1/σ_0	1, 1.10, 1.20, 1.30, 1.40, 1.50, 2.00, 3.00
Tamaño del subgrupo (m)	1, 3, 5
Datos extra después de que el cambio es detectado (w)	0, 5, 10, 15, 20

Usando la experimentación Monte Carlo para evaluar cada escenario presentado en la Tabla 10 los siguientes pasos fueron seguidos:

1. Marcar un punto de cambio τ .
2. Definir $i = 1$
3. Mientras ($i \leq \tau$) Y (la CUSUM no envíe una señal de alarma) generar X_i con vectores aleatorios utilizando la ecuación (1)
4. Si el CUSUM de auto-inicio envía una señal de alarma parar e ir al paso número 2, si no detecta cambios, ir al paso número 5.
5. Generar X_i vectores aleatorios utilizando la ecuación (1) hasta que la CUSUM de auto-inicio envíe una señal de alarma, al momento en que se envía la señal de alarma se le llamará T.
6. Estimar el punto de cambio $\hat{\tau}_{MLE}$ y $\hat{\tau}_{CUSUM}$ para los conjuntos de datos.
7. Calcular el error estimado $\hat{\tau}_{MLE} - \tau$ y $\hat{\tau}_{CUSUM} - \tau$
8. Repetir del paso número 2 al 4, hacerlo 10,000 veces.
9. Estimar el sesgo, el error estándar de $\hat{\tau}_{MLE}$ y $\hat{\tau}_{CUSUM}$, y el ARL de la carta de control CUSUM de auto-inicio.
10. Calcular la probabilidad $P(|\hat{\tau}_{MLE} - \tau| \leq D)$, donde $D = \{0, 1, 2, \dots, 24\}$
11. Seleccionar otro escenario y regresar al paso número 1.

4.4 Resultados de la Experimentación

Considerando los escenarios descritos en la sección previa se han realizado simulaciones Monte Carlo para evaluar el desempeño del estimador obtenido mediante el uso secuencial del CUSUM de auto-inicio y el MLE. Los resultados son presentados en tablas cuyos encabezados indican el escenario que se está considerando. La Tabla 11 muestra el desempeño para diferentes ubicaciones del punto de cambio y la magnitud del cambio para subgrupos de tamaño $m = 1$. De los datos en esta tabla se deduce que el ARL mejora el sesgo y el error estándar del punto de cambio calculado por el MLE del punto de cambio, éste se hace más pequeño cuando la magnitud del cambio se incrementa así como cuando el cambio ocurre lejos del punto inicial, en estos casos el CUSUM puede presentar falsas alarmas esto es, indicar que el proceso está fuera de control cuando en realidad no lo está. En la Tabla 12 se analiza el desempeño cuando la ubicación del cambio es en $\tau = 50$ considerando diferentes tamaños de subgrupo. La precisión del estimador también mejora cuando el tamaño del subgrupo se incrementa.

En las Tablas 13 y 14 se han incluido w datos extras correspondientes a observaciones después de que el CUSUM de auto-inicio ha detectado que un cambio ha ocurrido, esto se hizo con el fin de probar si el desempeño del MLE para el punto de cambio mejora cuando tiene más datos, se usó $\tau = 50$ and $\tau = 100$. De las mismas tablas se puede observar que el sesgo y el error estándar mejoran conforme w se incrementa. Aun cuando esta mejora en la precisión es significativa, su uso en la vida real deberá ser evaluado en base al costo de obtener nueva información.

La Tabla 15 muestra el desempeño de la carta de control CUSUSM con auto-inicio para estimar el punto de cambio, la cual es evaluada considerando los mismos escenarios usados en la Tabla 10. El $\hat{\tau}_{CUSUM}$ tiene menor sesgo y menor error estándar que el $\hat{\tau}_{MLE}$ cuando cambios pequeños tienen lugar ($\sigma_1/\sigma_0 < 1.3$). También el sesgo y el error estándar del $\hat{\tau}_{CUSUM}$ se decrementa cuando el cambio en el cociente de las varianzas tiene una tasa menor, todo esto respecto al sesgo y el error estándar del $\hat{\tau}_{MLE}$.

Finalmente en las Tablas 16 y 17 la probabilidad $P(|\hat{\tau}_{MLE} - \tau| \leq D)$ es calculada para $\tau = 100$ con $w=0$ y $w=20$ (valores extras) respetivamente. La probabilidad de detectar el punto de cambio de manera pronta se incrementa cuando w se hace más grande.

Tabla 11 Desempeño estimado considerando diferentes puntos de cambio τ para el CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\hat{\sigma}(\hat{\tau}_{MLE})$ usando $m = 1$. Los valores fueron obtenidos realizando 10,000 simulaciones.

τ	Métrica	σ_1/σ_0							
		1.00	1.10	1.20	1.30	1.40	1.50	2.0	3.0
25	ARL	113.245	98.979	82.461	65.591	49.925	39.105	13.815	5.692
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	90.595	77.553	60.279	43.470	28.633	18.531	2.438	0.263
	$\hat{\sigma}(\hat{\tau}_{MLE})$	114.224	104.131	94.055	81.661	64.257	49.645	13.186	3.823
50	ARL	114.486	94.995	70.058	49.692	34.682	25.254	9.420	4.772
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	89.964	70.060	44.568	25.931	13.157	6.830	0.537	-0.052
	$\hat{\sigma}(\hat{\tau}_{MLE})$	114.631	103.816	85.149	61.847	42.370	28.292	9.411	5.265
75	ARL	112.934	89.082	62.783	41.926	28.851	20.820	8.610	4.523
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	85.831	62.090	35.143	16.984	7.853	3.097	-0.353	-0.253
	$\hat{\sigma}(\hat{\tau}_{MLE})$	116.120	100.518	79.100	51.065	33.731	23.789	11.667	6.651
100	ARL	114.987	86.921	58.784	37.187	26.107	19.258	8.217	4.455
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	84.610	55.716	29.011	11.363	4.641	1.209	-1.045	-0.472
	$\hat{\sigma}(\hat{\tau}_{MLE})$	121.320	100.616	72.271	47.543	33.701	25.502	14.682	8.543

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros τ datos considerando los casos en que el punto de cambio se coloca en 25, 50, 75 y 100. El tamaño de subgrupo es $m = 1$. Se realizaron 10,000 simulaciones para la obtención de los datos en esta tabla.

Tabla 12 Desempeño considerando diferentes muestras de tamaño (n) para el CUSUM de auto-inicio para el ARL , el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\hat{\sigma}(\hat{\tau}_{MLE})$ $\sigma(\hat{\tau}_{MLE} - \tau)$ cuando $\tau = 50$. Los valores fueron calculados realizando 10,000 simulaciones.

		σ_1/σ_0							
m	Métrica	1.00	1.10	1.20	1.30	1.40	1.50	2.0	3.0
1	ARL	114.486	94.995	70.058	49.692	34.682	25.254	9.420	4.772
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	89.964	70.060	44.568	25.931	13.157	6.830	0.537	-0.052
	$\hat{\sigma}(\hat{\tau}_{MLE})$	114.631	103.816	85.149	61.847	42.370	28.292	9.411	5.265
3	ARL	112.994	94.414	70.954	49.869	34.624	24.706	9.445	4.772
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	39.719	37.085	27.635	18.123	11.028	7.211	0.857	0.079
	$\hat{\sigma}(\hat{\tau}_{MLE})$	113.069	100.677	79.236	62.733	46.653	35.250	15.070	8.618
5	ARL	115.314	86.112	56.793	37.567	26.030	19.043	8.368	4.487
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	13.306	15.570	7.774	4.455	2.431	1.012	0.129	-0.010
	$\hat{\sigma}(\hat{\tau}_{MLE})$	128.114	105.450	76.224	54.817	41.101	31.540	14.957	7.402

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros $\tau = 50$ datos. El tamaño de subgrupo en las observaciones es $m = 1, 3$ y 5 . Se realizaron 10,000 simulaciones para la obtención de los datos en esta tabla.

Tabla 13 Desempeño del CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\hat{\sigma}(\hat{\tau}_{MLE})$ cuando $\tau = 50$ y $m = 1$ considerando diferentes valores de w datos extras. Los resultados fueron obtenidos realizando 10,000 simulaciones.

		σ_1/σ_0							
w	Métrica	1.00	1.10	1.20	1.30	1.40	1.50	2.0	3.0
5	ARL	115.087	94.248	69.737	49.401	34.421	25.714	9.453	4.789
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	72.536	58.266	37.153	20.262	10.880	5.727	0.876	0.556
	$\hat{\sigma}(\hat{\tau}_{MLE})$	112.237	98.721	79.355	55.341	38.011	27.965	8.681	3.288
10	ARL	114.144	87.273	58.474	38.931	25.892	19.228	8.209	4.471
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	53.903	39.478	19.800	7.692	2.934	1.394	0.798	0.703
	$\hat{\sigma}(\hat{\tau}_{MLE})$	120.774	100.360	69.988	47.984	31.563	24.200	9.148	2.520
15	ARL	113.307	93.638	68.636	49.821	34.172	24.272	9.491	4.758
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	67.487	54.233	34.575	20.343	10.810	5.695	1.489	0.801
	$\hat{\sigma}(\hat{\tau}_{MLE})$	111.292	97.209	75.415	56.019	37.501	23.582	7.185	2.521
20	ARL	114.188	95.449	68.720	49.304	34.952	25.424	9.479	4.777
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	69.453	55.493	34.249	19.547	10.838	6.038	1.793	0.754
	$\hat{\sigma}(\hat{\tau}_{MLE})$	111.558	100.678	74.994	56.331	37.398	24.124	6.987	2.503

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros $\tau = 50$ datos. El tamaño de subgrupo en las observaciones es $m = 1$. Se han tomado $w = 5, 10, 15$ y 20 datos extras. Se realizaron 10,000 simulaciones para la obtención de los datos en esta tabla.

Tabla 14 El desempeño del CUSUM de auto-inicio para el ARL, el sesgo de la estimación del punto de cambio $\hat{E}(\hat{\tau}_{MLE} - \tau)$ y el error estándar asociado para la estimación del punto de cambio $\hat{\sigma}(\hat{\tau}_{MLE})$ cuando $\tau = 100$ y $m = 1$ considerando diferentes valores de w datos extras. Los resultados fueron obtenidos mediante la realización de 10,000 simulaciones.

w	Métrica	σ_1/σ_0							
		1.00	1.10	1.20	1.30	1.40	1.50	2.0	3.0
5	ARL	113.177	87.111	59.017	37.625	26.279	18.947	8.292	4.496
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	61.598	41.227	21.327	7.364	3.216	0.844	0.137	0.465
	$\hat{\sigma}(\hat{\tau}_{MLE})$	120.458	100.670	72.058	45.752	32.832	24.803	11.171	4.281
10	ARL	114.144	87.273	58.474	38.931	25.892	19.228	8.209	4.471
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	53.903	39.478	19.800	7.692	2.934	1.394	0.798	0.703
	$\hat{\sigma}(\hat{\tau}_{MLE})$	120.774	100.360	69.988	47.984	31.563	24.200	9.148	2.520
15	ARL	115.824	86.210	57.775	38.007	25.825	18.943	8.302	4.470
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	54.214	36.246	18.045	7.985	3.116	2.184	1.252	0.776
	$\hat{\sigma}(\hat{\tau}_{MLE})$	123.980	98.952	71.037	46.309	30.894	22.845	7.651	2.653
20	ARL	115.632	87.736	57.538	39.093	26.227	19.323	8.369	4.486
	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	53.696	37.149	18.216	8.803	4.444	2.747	1.537	0.844
	$\hat{\sigma}(\hat{\tau}_{MLE})$	124.934	102.692	70.134	47.067	30.518	21.819	6.692	2.461

De acuerdo al algoritmo el ARL es calculado considerando que no hubo ninguna falsa alarma previa a la obtención de los primeros $\tau = 100$ datos. El tamaño de subgrupo en las observaciones es $m = 1$. Se han tomado $w = 5, 10, 15$ y 20 datos extras. Se realizaron 10,000 simulaciones para la obtención de los datos en esta tabla.

Tabla 15 Comparación de las desviaciones estándar de las estimaciones obtenidas por el uso secuencial de la carta CUSUM de auto-inicio y el MLE para el punto de cambio y el equivalente en la CUSUM de auto-inicio. para $m = 1$ considerando diferentes valores de τ . Los resultados fueron obtenidos utilizando 10,000 simulaciones.

τ	Métrica	σ_1/σ_0							
		1.00	1.10	1.20	1.30	1.40	1.50	2.0	3.0
25	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	97.865	83.840	67.277	50.207	34.756	24.216	0.996	-4.134
	$\hat{\sigma}(\hat{\tau}_{MLE})$	116.565	106.128	96.734	86.612	70.851	59.901	22.886	7.833
50	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	98.974	79.630	54.550	34.330	19.474	10.496	-2.293	-4.414
	$\hat{\sigma}(\hat{\tau}_{MLE})$	113.829	103.605	86.761	67.819	49.305	34.932	10.087	7.651
75	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	97.379	73.735	47.190	26.405	13.688	6.426	-2.828	-4.330
	$\hat{\sigma}(\hat{\tau}_{MLE})$	113.171	97.636	79.680	54.760	36.223	23.396	8.986	7.558
100	$\hat{E}(\hat{\tau}_{MLE} - \tau)$	99.388	71.612	43.231	21.738	11.080	5.000	-3.027	-4.362
	$\hat{\sigma}(\hat{\tau}_{MLE})$	115.853	94.938	71.994	44.169	30.956	20.380	8.761	7.659

En esta tabla se calcula el sesgo y la desviación estándar considerando un punto de cambio τ colocado en 25, 50, 75 y 100 para un tamaño de subgrupo $m = 1$. Se realizaron 10,000 simulaciones para la obtención de los datos en esta tabla.

Tabla 16 Probabilidades de que el sesgo del estimador del parámetro este dentro de un rango de valores pre-establecidos $\{0, 1, 2, \dots, 24\}$ para diferentes valores de $\hat{\tau}_{MLE}$ con $\tau = 100$, $m = 1$ (tamaño del subgrupo) y $w = 0$ datos extras.

	σ_1/σ_0						
	1.10	1.20	1.30	1.40	1.50	2.0	3.0
$P(\hat{\tau}_{MLE} = \tau)$	0.015	0.029	0.048	0.077	0.102	0.238	0.406
$P(\hat{\tau}_{MLE} - \tau \leq 1)$	0.038	0.071	0.113	0.164	0.218	0.437	0.652
$P(\hat{\tau}_{MLE} - \tau \leq 2)$	0.061	0.104	0.170	0.237	0.306	0.558	0.776
$P(\hat{\tau}_{MLE} - \tau \leq 3)$	0.082	0.138	0.220	0.299	0.377	0.649	0.850
$P(\hat{\tau}_{MLE} - \tau \leq 4)$	0.102	0.170	0.263	0.354	0.435	0.715	0.895
$P(\hat{\tau}_{MLE} - \tau \leq 5)$	0.119	0.197	0.305	0.402	0.490	0.766	0.920
$P(\hat{\tau}_{MLE} - \tau \leq 6)$	0.136	0.226	0.338	0.440	0.539	0.810	0.939
$P(\hat{\tau}_{MLE} - \tau \leq 7)$	0.150	0.248	0.369	0.477	0.575	0.839	0.951
$P(\hat{\tau}_{MLE} - \tau \leq 8)$	0.166	0.273	0.397	0.508	0.610	0.860	0.959
$P(\hat{\tau}_{MLE} - \tau \leq 9)$	0.180	0.293	0.426	0.536	0.636	0.879	0.967
$P(\hat{\tau}_{MLE} - \tau \leq 10)$	0.195	0.313	0.449	0.562	0.661	0.896	0.972
$P(\hat{\tau}_{MLE} - \tau \leq 12)$	0.221	0.350	0.490	0.609	0.705	0.919	0.977
$P(\hat{\tau}_{MLE} - \tau \leq 14)$	0.244	0.384	0.526	0.647	0.742	0.933	0.981
$P(\hat{\tau}_{MLE} - \tau \leq 16)$	0.265	0.415	0.557	0.679	0.772	0.944	0.983
$P(\hat{\tau}_{MLE} - \tau \leq 18)$	0.283	0.437	0.586	0.708	0.796	0.950	0.985
$P(\hat{\tau}_{MLE} - \tau \leq 20)$	0.306	0.461	0.615	0.732	0.817	0.956	0.987
$P(\hat{\tau}_{MLE} - \tau \leq 22)$	0.324	0.484	0.638	0.753	0.834	0.960	0.988
$P(\hat{\tau}_{MLE} - \tau \leq 24)$	0.341	0.505	0.658	0.773	0.850	0.963	0.989

Tabla 17 Probabilidades del sesgo del estimador para diferentes valores predefinidos $\{0, 1, 2, \dots, 24\}$ de $\hat{\tau}_{MLE}$ con $\tau = 100$, $m = 1$ (tamaño del subgrupo) y $w = 20$ datos extras.

	δ						
	1.10	1.20	1.30	1.40	1.50	2.0	3.0
$P(\hat{\tau}_{MLE} = \tau)$	0.011	0.024	0.048	0.076	0.105	0.240	0.419
$P(\hat{\tau}_{MLE} - \tau \leq 1)$	0.026	0.062	0.112	0.167	0.231	0.448	0.679
$P(\hat{\tau}_{MLE} - \tau \leq 2)$	0.040	0.098	0.166	0.241	0.322	0.580	0.803
$P(\hat{\tau}_{MLE} - \tau \leq 3)$	0.056	0.128	0.214	0.300	0.400	0.675	0.876
$P(\hat{\tau}_{MLE} - \tau \leq 4)$	0.068	0.155	0.254	0.353	0.460	0.743	0.920
$P(\hat{\tau}_{MLE} - \tau \leq 5)$	0.081	0.177	0.290	0.397	0.510	0.792	0.946
$P(\hat{\tau}_{MLE} - \tau \leq 6)$	0.095	0.200	0.324	0.437	0.555	0.830	0.964
$P(\hat{\tau}_{MLE} - \tau \leq 7)$	0.104	0.223	0.350	0.475	0.590	0.860	0.975
$P(\hat{\tau}_{MLE} - \tau \leq 8)$	0.118	0.242	0.376	0.510	0.628	0.885	0.985
$P(\hat{\tau}_{MLE} - \tau \leq 9)$	0.127	0.259	0.401	0.538	0.659	0.905	0.989
$P(\hat{\tau}_{MLE} - \tau \leq 10)$	0.139	0.278	0.424	0.566	0.683	0.920	0.993
$P(\hat{\tau}_{MLE} - \tau \leq 12)$	0.159	0.310	0.463	0.610	0.727	0.943	0.996
$P(\hat{\tau}_{MLE} - \tau \leq 14)$	0.178	0.340	0.501	0.648	0.761	0.959	0.998
$P(\hat{\tau}_{MLE} - \tau \leq 16)$	0.195	0.367	0.533	0.681	0.790	0.969	0.999
$P(\hat{\tau}_{MLE} - \tau \leq 18)$	0.214	0.394	0.563	0.709	0.813	0.976	0.999
$P(\hat{\tau}_{MLE} - \tau \leq 20)$	0.234	0.417	0.587	0.733	0.833	0.982	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 22)$	0.252	0.442	0.614	0.756	0.851	0.986	1.000
$P(\hat{\tau}_{MLE} - \tau \leq 24)$	0.270	0.466	0.639	0.778	0.866	0.989	1.000

4.5 Conclusiones y Trabajos Futuros

Debido a la forma recursiva en que la carta de control CUSUM de auto-inicio realiza la estimación de los parámetros, haciendo un cálculo continuo de las medias y la desviación estándar considerando todas las observaciones, se evita sesgos por encima y debajo de la estimación creada en la Fase I del SPC. La estimación del punto de cambio usando un enfoque común a las cartas de control CUSUM (\hat{t}_{CUSUM}) se sesga aún para cambios mayores a $\sigma_1/\sigma_0 > 1.3$. Es en este momento, después de la detección por parte de la carta de control CUSUM, cuando el uso del estimador de máxima verosimilitud (MLE) muestra su utilidad al obtener un estimador con menor sesgo, y se sugiere sea usado en la práctica cuando se tengan cocientes de las varianzas mayores o iguales a $\sigma_1/\sigma_0 > 1.3$, complementando esta sugerencia es conveniente considerar un periodo de calentamiento para la carta de control CUSUM de auto-inicio de 50 muestras, tal como lo recomienda Hawkins [5].

CAPÍTULO 5. INVESTIGACIÓN 3

Estimación del Punto de Cambio Después de la Señal de una Carta de Control para la Gamma

Esta investigación es la tercera parte del presente trabajo. En esta investigación se considera la inspección de procesos que siguen una distribución Gamma cuyos parámetros de forma y escala (α y β respectivamente) cambian en el momento desconocido (τ), esto es $\alpha_0 \neq \alpha_1$ y $\beta_0 \neq \beta_1$. Se propone el uso secuencial de una carta de control con estimadores de máxima verosimilitud para el monitoreo de este proceso. Se han realizado simulaciones para comprar el desempeño del estimador de máxima verosimilitud del punto de cambio y el estimador de la carta de control, basado en el sesgo y la desviación estándar del estimador ($\hat{\tau}$) respecto al valor real τ .

Este artículo está basado en el artículo in extenso publicado en el Congreso Internacional ISERC 2013 llevado a cabo en San Juan Puerto Rico en las fechas del 18 al 22 de Mayo. Se presentaron resultados preliminares, pero la investigación completa se planifica enviar a Quality and Reliability Engineering International.

Estimación del Punto de Cambio Después de la Señal de una Carta de Control para la Gamma

Ángel S. Pérez-Blanco, M.C.,
Alvaro E. Cordero-Franco, Ph.D., and Víctor G. Tercero-Gómez, Ph.D.,
Universidad Autónoma de Nuevo León- Facultad de Ciencias Físico Matemáticas
San Nicolás de los Garza, Nuevo León, México

Resumen

Para manejar procesos transaccionales y de manufactura, frecuentemente es necesario monitorear el tiempo entre eventos. Tal es el caso de las fallas en un lapso de tiempo o las llegadas en intervalos de tiempo, los cuales pueden ser modelados usando una función de distribución Gamma. Es sabido que las cartas de control son usadas para detectar cambios en los procesos que pueden llevar a la identificación de causas asignables de variación; lo cual puede ser utilizado para mantener el sistema bajo control. Particularmente, la carta de control Gamma es capaz de detectar cambios sostenidos, sin embargo no indica el momento inicial del cambio, este problema de estimar el momento de un cambio estructural es llamado el problema del punto de cambio. El presente trabajo propone el uso secuencial de un estimador de máxima verosimilitud para un conjunto de datos independientes provenientes de una función de distribución Gamma después de que la carta de control para la Gamma ha indicado que el proceso está fuera de control estadístico. La derivación de este estimador de punto de cambio es presentada así como un modelo conceptual de su integración utilizando una carta de control para la Gamma. También se presentan un ejemplo numérico y el comportamiento de los resultados mediante la estimación del sesgo y el error estándar del estimador. Los ingenieros en calidad pueden encontrar esta herramienta útil cuando tratan con eventos en intervalos de tiempo, en los cuales el estimador propuesto puede ser implementado, sin hacerle grandes cambios, en las aplicaciones actuales que esté realizando de acuerdo a los planes de control de procesos que esté llevando a cabo.

Palabras clave: Punto de cambio, procesos Gamma, control estadístico de procesos, estimador de máxima verosimilitud

5.1 Introducción

Las cartas de control son usadas en el Control Estadístico de Procesos (SPC) para monitorear sistemas. Mantener un proceso en un estado de control estadístico implica que sea predecible, haciendo con ello que sea posible manejarlo. Trabajando en la idea que las muestras de una característica de calidad de un producto siguen algún tipo de función de distribución de probabilidad; es posible predecir mediante un intervalo de confianza, los resultados de un sistema dentro de un rango delimitado por un Límite de Control Superior y un Límite de Control Inferior (UCL, LCL respectivamente). Cuando el valor medido de una característica de calidad del producto está fuera de los límites de control establecidos, se considera que algo perturbó al proceso y, en consecuencia, los parámetros del modelo han cambiado. Los cambios pueden ser aislados o sostenidos. Cuando tratamos con cambios sostenidos, el momento en que el cambio ocurre es llamado el punto de cambio y el análisis de punto de cambio es la herramienta que se usará para analizarlo. Mientras más cercana sea la estimación del punto de cambio al valor real del mismo, más útil será para encontrar la causa asignable que provocó la perturbación del sistema, facilitando de esta forma la mejora del proceso. Esta investigación analiza los conjuntos de datos obtenidos como una serie en el tiempo que pueden ser modelados mediante una función de distribución Gamma cuando los parámetros iniciales son conocidos. Esta función de distribución es usada para modelar el tiempo entre llegadas, y es reportada en diferentes áreas como la industria química, los estudios biológicos y control de inventarios. En particular en [99] se puede encontrar una aplicación de la Gamma al control de inventarios y en [100] una aplicación sobre los efectos sobre el flujo sanguíneo a ciertos estímulos. Esta investigación propone estimar el punto de cambio y los parámetros de una función de distribución Gamma usando un estimador de máxima verosimilitud (MLE) después que una carta de control EWMA, para observaciones provenientes de una función de distribución Gamma, indica que un punto ha salido de control. El correspondiente punto de cambio usando MLE para el punto de cambio es obtenido asumiendo que un cambio ha ocurrido en ambos parámetros. Se ha descrito un modelo de cómo realizar la estimación del punto de cambio usando de manera secuencial una carta de control y el MLE, también se presenta un ejemplo numérico y simulaciones realizadas con el fin de evaluar el desempeño del modelo propuesto.

5.2 Trabajos Previos

Los primeros trabajos sobre CPA fueron realizados por Girshisck y Rubin utilizando un enfoque Bayesiano para su uso en el SPC [8]. Posteriormente, Page propuso la técnica de CUSUM en 1954 para detectar cambios sostenidos en series de tiempo [4] [36] [21]. Respecto a procedimientos paramétricos, Hinkley [30] consolidó la teoría correspondiente para encontrar el punto de cambio utilizando MLE's y pruebas de cociente de verosimilitud (LTR, por sus siglas en inglés) considerando los parámetros iniciales conocidos y desconocidos. Samuel [32] propuso el uso de un estimador después de que una carta de control de Shewhart indicara un cambio en el proceso. Este mismo enfoque fue analizado por Nedumaran, Pignatiello y Calvin [63] y Samuel

et al. [61]. Una herramienta diferente que crea cartas de control basadas en un enfoque de CPA usando una analogía del cociente de verosimilitud generalizado fue desarrollada por Hawkins y Zamba [71]. Usando una combinación de MLEs y cartas de control para determinar una estimación del punto de cambio Shao y Hou [35] mostraron que una carta de control EWMA puede ser una forma relativamente buena para estimar el punto de cambio en un proceso Gamma cuando solo el parámetro de escala cambia. Ellos han combinado el MLE para el punto de cambio con una carta de control Gamma para estimar y detectar puntos de cambio, respectivamente. Sus resultados muestran que el MLE fue un mejor estimador del punto de cambio que el indicado por la carta de control EWMA. La presente investigación continúa ese trabajo derivando el estimador correspondiente del punto de cambio mediante el uso del MLE en el caso en que ambos parámetros cambian para un conjunto de datos que pueden ser modelados mediante una función de distribución Gamma.

5.3 Gráfica de Control EWMA

La carta de control EWMA (Exponentially Weighted Moving Average) pondera las observaciones históricas del proceso con las recientes. Esto hace que sea útil en la detección de cambios pequeños y sostenidos en la media. Su ecuación es

$$z_i = \lambda \bar{x}_i + (1 - \lambda)z_{i-1} \quad (1)$$

El parámetro λ es el que da el “peso” a las observaciones, en particular mayor peso a las observaciones más recientes. Su valor debe estar entre (0, 1]. La carta que se obtendrá será una transformación de los datos originales. Las ecuaciones que definen los límites de control (3) incluyen un segundo parámetro L el cual se utiliza para definir un múltiplo de la desviación estándar.

El modelo propuesto usa una carta de control EWMA para detectar los cambios durante el monitoreo del proceso. Una vez que el cambio es detectado, el punto de cambio τ es estimado usando el MLE para series de observaciones independientes provenientes de una función de distribución Gamma $X_1, \dots, X_\tau, X_{\tau+1}, \dots, X_T$, como se muestra en la ecuación (2)

$$X_i \sim \begin{cases} G(\alpha_0, \beta_0), & 1 \leq i \leq \tau \\ G(\alpha_1, \beta_1), & \tau < i \leq T \end{cases} \quad (2)$$

α_0 y β_0 , son parámetros conocidos, mientras que α_1 y β_1 son desconocidos, y T es a última observación de la serie de datos. α es identificado como el parámetro de la forma y β es identificado como el parámetro de la escala. Los límites de control para la carta de control EWMA son los siguientes:

$$UCL = \mu_0 + L\sigma_0 \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2i}]}, \quad LCL = \mu_0 - L\sigma_0 \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2i}]} \quad (3)$$

Donde $\mu_0 = \alpha_0\beta_0$ y $\sigma_0 = \sqrt{\alpha_0\beta_0^2}$ se refieren a la media y la desviación estándar del proceso respectivamente. Es conveniente comentar que debido a estas ecuaciones un cambio en los parámetros α_0 y/o β_0 no necesariamente implica un cambio en μ_0 y/o σ_0 . El $\hat{\tau}_{EWMA}$ es definido como el primer punto fuera de control en una carta de control EWMA. Si ya no son tomadas más muestras después de que la carta de control EWMA indica que se ha dado el primer punto fuera de los límites de control entonces se tiene $T = \hat{\tau}_{EWMA}$.

En las secciones siguientes se presenta un ejemplo de cómo se crea, utiliza y define una carta EWMA.

5.4 Estimación del Punto de cambio

Para realizar la estimación del punto de cambio se usará la teoría desarrollada por Hinkley [30] y se considerará un proceso Gamma donde los cambios en el parámetro ocurren entre el tiempo τ y $\tau + 1$ en esta investigación, además, se asumirán conocidos los parámetros α_0 y β_0 y se considerará que las observaciones son individuales. Bajo estas consideraciones la función de verosimilitud usada para obtener los estimadores de máxima verosimilitud es (4)

$$L(\tau, \alpha_1, \beta_1 | \bar{x}) = \prod_{i=1}^{\tau} f(\tau, \alpha_0, \beta_0 | x_i) \cdot \prod_{i=\tau+1}^T f(\tau, \alpha_1, \beta_1 | x_i) \quad (4)$$

Sin embargo, utilizar el logaritmo de la función de verosimilitud facilita el manejo algebraico por lo que se utilizará (5)

$$\begin{aligned} \ln L = & -\tau \ln \Gamma(\alpha_0) - \tau \alpha_0 \ln(\beta_0) - (T - \tau) \ln \Gamma(\alpha_1) - (T - \tau) \alpha_1 \ln(\beta_1) \\ & + (\alpha_0 - 1) \sum_{i=1}^{\tau} \ln(x_i) - \sum_{i=1}^{\tau} \frac{x_i}{\beta_0} + (\alpha_1 - 1) \sum_{i=\tau+1}^T \ln(x_i) - \sum_{i=\tau+1}^T \frac{x_i}{\beta_1} \end{aligned} \quad (5)$$

Para encontrar el valor del estimador de máxima verosimilitud correspondiente a α_1 y β_1 es necesario resolver el sistema de ecuaciones dado por la primera derivada parcial de la ecuación (5) con respecto a α_1 y β_1 e igualarlas a cero, (6) y (7) representan este sistema de ecuaciones, y su solución es usada para encontrar el punto de cambio asociado al MLE, ver ecuación (8).

Las ecuaciones obtenidas son:

$$(T - \tau) \frac{\Gamma'(\hat{\alpha}_1)}{\Gamma(\hat{\alpha}_1)} = \frac{\sum_{i=\tau+1}^T \ln(x_i)}{T - \tau} - (T - \tau) \ln(\hat{\beta}_1) \quad (6)$$

$$\hat{\beta}_1 = \frac{\sum_{i=\tau+1}^T x_i}{(T - \tau)\hat{\alpha}_1} \quad (7)$$

$$\hat{\tau}_{MLE} = \arg \max \ln L(\alpha_1, \beta_1, \tau) \quad (8)$$

Donde el cociente $\frac{\Gamma'(\hat{\alpha}_1)}{\Gamma(\hat{\alpha}_1)}$ es conocido como la función DiGamma.

5.5 Integración del Modelo

La Figura 10 muestra como la carta de control EWMA y el $\hat{\tau}_{MLE}$ pueden ser integrados para detectar si un cambio sostenido ocurre y estimar cuando fue que inició. El monitoreo del proceso inicia con la carta de control Gamma y tan pronto como una muestra es señalada como fuera de control, esta muestra será conocida como T, la ecuación (8) puede ser usada para estimar el punto de cambio. Si es necesario incrementar la precisión, es posible permitir que el proceso siga funcionando por un número extra w y el estadístico cambiará a $T' = T + w$ para realizar la estimación del punto de cambio.

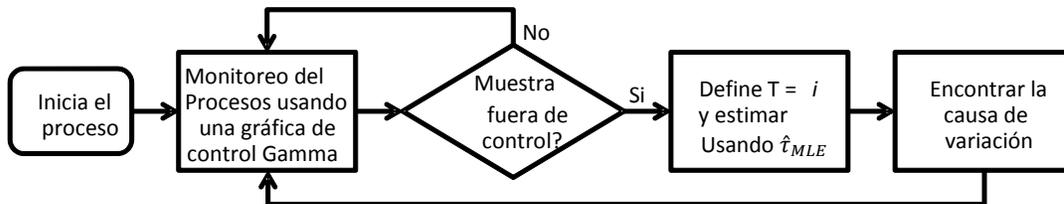


Figura 10 Detección secuencial del punto de cambio y estimación del mismo usando una gráfica de control para la Gamma y el MLE para el punto de cambio para una serie de observaciones independientes.

Ambos, la carta de control para la Gamma y el $\hat{\tau}_{MLE}$ asumen que los parámetros iniciales son conocidos (antes del cambio) en el conjunto de observaciones independientes que son modeladas mediante la función de distribución Gamma. En las siguientes secciones se detalla la ejecución de simulaciones para evaluar el desempeño del $\hat{\tau}_{MLE}$ cuando este esquema es presentado y compara sus resultados con el $\hat{\tau}_{EWMA}$ en diferentes escenarios.

Antes de iniciar con el ejemplo numérico es conveniente recordar que las ecuaciones para calcular la μ y σ para una función de distribución Gamma son:

$$\mu = \alpha\beta \quad (9)$$

$$\sigma = \sqrt{\alpha\beta^2} \quad (10)$$

Estas ecuaciones implican que para diferentes valores de α y β se pueden tener iguales valores de μ y/o σ , es por ello que en la sección 5.7 se analizara esta situación y se detallará en los resultados reportados en la Tabla 22.

5.6 Ejemplo Numérico

Un ejemplo del uso secuencial de la carta de control EWMA y el MLE para el punto de cambio propuesto para el punto de cambio es detallado. Se han considerado 15 datos provenientes de una función de distribución Gamma ($\alpha_0=1, \beta_0=1, \mu_0=1, \sigma_0=1$), obtenidos mediante simulación, y a partir de este punto se ha generado un cambio en el proceso generando nuevos datos a partir de una función de distribución Gamma ($\alpha_1=2, \beta_1=3, \mu_1=6, \sigma_1=3\sqrt{2}$) hasta que la carta de control EWMA detecta una observación fuera de los límites de control. Los valores de estos parámetros han sido seleccionados con el fin de explicar brevemente el efecto de los cambios de α_1 y β_1 en la estimación del punto de cambio mediante la carta de control EWMA solamente y el uso secuencial de la carta EWMA y el MLE para el punto de cambio. Los mismos valores de los parámetros son utilizados en la simulación. Los límites de control fueron calculados considerando $\lambda = 0.20$ y $L = 2.962$ los cuales fueron seleccionados a partir de las recomendaciones de Montgomery [6] y Lucas y Saccucci [101] en donde indican cinco parejas de valores que son útiles en desviaciones de la media que van de 0.00 a 4.00σ , la elección de estos valores determina el tamaño del ARL. En esta investigación los valores seleccionados corresponden a la pareja intermedia entre las cinco recomendadas. Las observaciones y cálculos para la carta de control EWMA y el MLE para el punto de cambio son presentadas en la Tabla 18. Como se ve en esta Tabla, la carta de control EWMA detecta un valor fuera de los límites de control en la observación 21. El punto de cambio estimado por el MLE para el punto de cambio está en la observación 17, cuatro observaciones antes de la detección de la carta de control EWMA.

Tabla 18 Cálculos requeridos para el EWMA y el MLE. Las primeras 15 observaciones provienen de una Gamma ($\alpha_1 = 1, \beta_1 = 1$) y las siguientes 6 de una Gamma ($\alpha_1 = 2, \beta_1 = 3$), el tamaño de subgrupo es 1, $\lambda = 0.20$ y $L = 2.962$. Las col. 3-5 corresponden a los cálculos para EWMA y las col. 6-8 corresponden a los cálculos para MLE.

	Observación	EWMA			MLE		
i	X_i	Z_i	UCL	LCL	$\hat{\alpha}_1$	$\hat{\beta}_1$	$Ln(\hat{a}_1, \hat{b}_1, \hat{\tau})$
0		1					
1	0.7393	0.9479	1.5924	0.4076	1.5565	1.5624	-56.6714
2	0.2257	0.8034	1.7586	0.2414	1.5276	1.6436	-55.2213
3	5.0164	1.6460	1.8481	0.1519	1.5314	1.5789	-57.3813
4	1.4830	1.6134	1.9007	0.0993	1.5409	1.5925	-57.5861
5	1.6873	1.6282	1.9328	0.0672	1.5524	1.6005	-57.9435
6	0.0521	1.3130	1.9528	0.0472	1.4894	1.7362	-55.5954
7	1.1500	1.2804	1.9654	0.0346	1.4938	1.7729	-55.4472
8	0.5191	1.1281	1.9733	0.0267	1.4824	1.8518	-54.5835
9	0.9342	1.0893	1.9784	0.0216	1.4819	1.9106	-54.1729
10	2.9848	1.4684	1.9816	0.0184	1.4946	1.8892	-55.2997
11	1.4126	1.4573	1.9837	0.0163	1.5023	1.9290	-55.3071
12	0.4868	1.2632	1.9850	0.0150	1.4842	2.0428	-54.2707
13	0.0959	1.0297	1.9858	0.0142	1.4105	2.2720	-52.3195
14	0.7560	0.9750	1.9864	0.0136	1.3959	2.4054	-51.5957
15	0.8316	0.9463	1.9867	0.0133	1.3810	2.5532	-50.9130
16	0.6284	0.8827	1.9869	0.0131	1.3537	2.7576	-49.9923
17	0.7726	0.8607	1.9871	0.0129	1.3280	2.9825	-49.1952
18	2.2890	1.1464	1.9872	0.0128	1.3285	3.0862	-49.6384
19	3.1737	1.5518	1.9872	0.0128	1.3341	3.1364	-50.7592
20	3.0167	1.8448	1.9873	0.0127	1.3399	3.2099	-51.7517
21	8.4450	3.1648	1.9873	0.0127	1.3292	2.8894	-56.7940

En este ejemplo la observación 16 es el primer valor generado después de que el cambio ocurrió. Las Figuras 11 y 12 muestran el comportamiento de los datos originales y transformados para ser usados en la carta de control EWMA respectivamente.

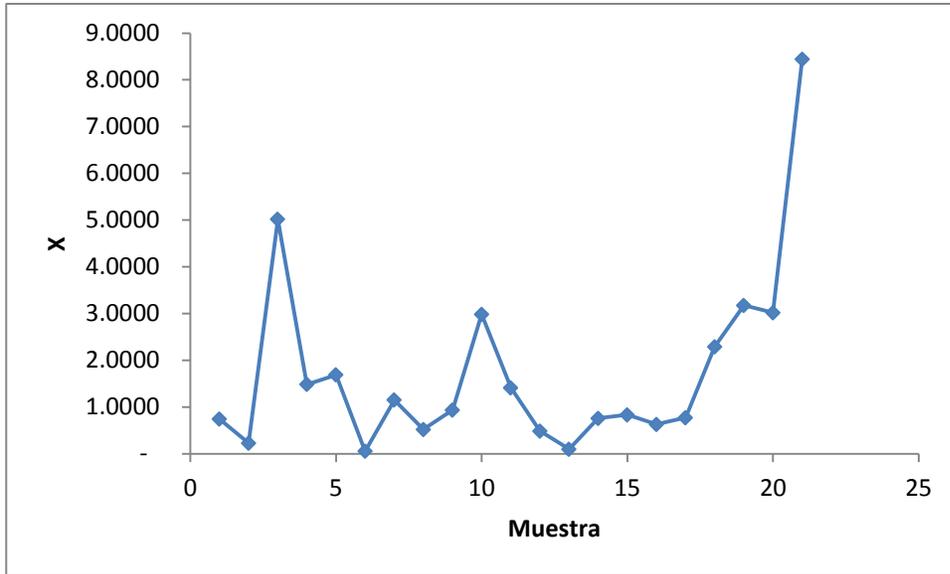


Figura 11 Gráfica de dispersión de los datos del ejemplo numérico.

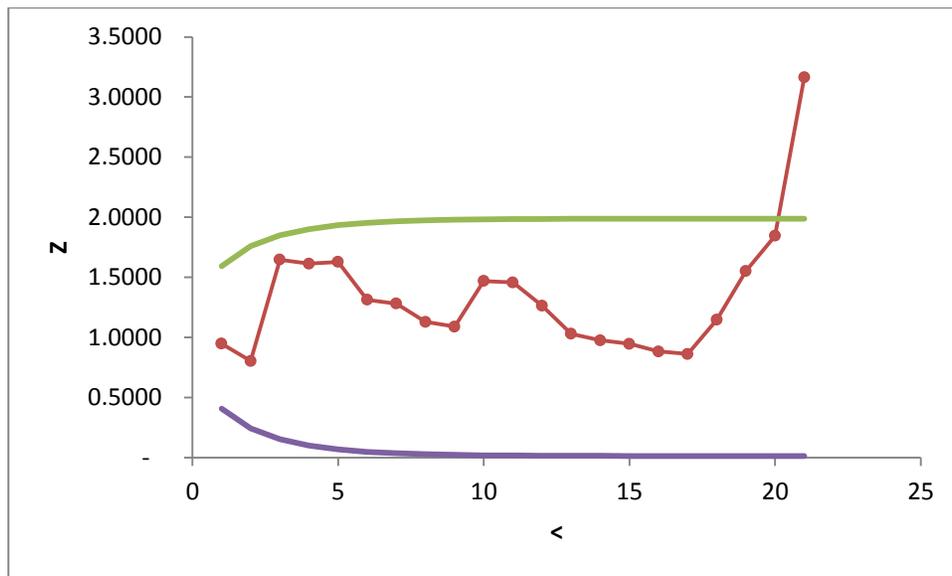


Figura 12 Carta de control EWMA para los datos del ejemplo numérico.

En esta investigación se ha considerado un modelo que sigue una función de distribución Gamma debido a sus aplicaciones en diferentes areas del conocimiento, principalmente en procesos industriales. Durante el análisis de la Fase II los parámetros iniciales son considerados como conocidos, sin embargo, cuando un cambio ocurre, es difícil asumir que solo un parámetro ha cambiado. Es por ello que se ha considerado importante evaluar la situación cuando ambos parámetros, de forma y de escala, de la función de distribución Gamma cambian al mismo tiempo, además de la relación directa y simultánea que existe entre éstos μ y σ de esta función y que permitiría tener un cambio en los parámetros α y β sin presentar

cambios en la μ . Los parámetros de forma α y escala β definen la función de distribución Gamma y su estimación requiere de métodos numéricos para resolver las ecuaciones (5) y (6), debido a que este sistema de ecuaciones no tiene una solución analítica. La ecuación (5) incluye la función DiGamma lo que hace la derivación de los estimadores de máxima verosimilitud aún más compleja. Una discusión completa sobre la función DiGamma puede ser consultada en [102].

5.7 Resultados de la simulación y la experimentación

La Tabla 19 indica los factores considerados para probar la técnica propuesta, el uso secuencial de una carta de control EWMA y la estimación del punto de cambio utilizando el MLE. Las simulaciones consideran un cambio en la observación $\tau = 100$ en el modelo (1) con parámetros $\alpha_0 = 1$, $\beta = 1$ para la función de distribución Gamma y los valores de $\lambda = 0.20$ y de $L = 2.962$ para la carta de control EWMA.

Tabla 19 Factores usados para evaluar el desempeño de $\hat{\tau}_{EWMA}$ y $\hat{\tau}_{MLE}$ para $\tau = 100$, $\alpha_0 = 1$, $\beta_0 = 1$, $\lambda = 0.20$ y $L = 2.962$.

Factores Evaluados	Niveles
Cociente α_1/α_0	1.0, 1.3, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0
Cociente β_1/β_0	1.0, 1.3, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0

La simulación tiene los siguientes pasos:

1. Seleccionar un escenario de la Tabla 19.
2. Generar τ observaciones en control de acuerdo al escenario seleccionado y hacer $i = \tau$.
3. Generar otra observación fuera de control bajo el escenario seleccionado y hacer $i = i + 1$
4. Si la última observación está fuera de control entonces hacer $T = i, \hat{\tau}_{EWMA} = T$ e ir al paso 5. Si no es así ir al paso 3.
5. Calcular $\hat{\tau}_{MLE}$.
6. Repetir del paso 2 al paso 5. Hacerlo 10,000 veces.
7. Calcular la media y la desviación estándar de le error (usando $\hat{\tau}_{EWMA}$ y $\hat{\tau}_{MLE}$).
8. Regresar al paso 1 y seleccionar otro escenario.

El desempeño ha sido evaluado considerando los escenarios de la Tabla 19. La Tabla 20 muestra el desempeño de la carta de control de EWMA que se está considerando. El ARL y la desviación estándar de $\hat{\tau}_{EWMA}$ fueron medidos cuando el punto de cambio fue establecido en 100. La Tabla 21 muestra el desempeño del $\hat{\tau}_{MLE}$, su sesgo y su desviación estándar. Es posible observar que en la mayoría de los casos $\hat{\tau}_{MLE}$ presenta un sesgo menor que $\hat{\tau}_{EWMA}$, mientras que los últimos presentan una desviación estándar menor. Sin embargo la desviación estándar de los $\hat{\tau}_{MLE}$ se

espera decrezca conforme el tamaño de muestra crezca, esto debido a la eficiencia asintótica de los MLEs (esto significa que ningún otro estimador insesgado tiene un error cuadrado medio menor cuando el tamaño de la muestra tiende a infinito).

Tabla 20 Desempeño considerando diferentes valores de α_1 y β_1 para las muestras de la Gamma, utilizando una carta de control EWMA para estimar el punto de cambio. Los parámetros de la simulación son: $\tau = 100$; 1,000 simulaciones; $\alpha_0 = 1$, $\beta_0 = 1$, $\lambda = 0.20$, $L = 2.962$; el ARL y desviación estándar S_{EWMA} , son calculados.

α_1	β_1	1.00	1.30	1.50	2.00	2.50	3.00	3.50	4.00
1.00	ARL	178.8900	37.5860	20.3440	9.3810	6.0710	4.5290	3.8170	3.2660
	S_{EWMA}	180.2196	36.2459	18.5335	7.7639	4.9788	3.3726	2.8117	2.2254
1.30	ARL	48.0460	15.7590	9.9290	5.5540	3.9770	3.1290	2.7670	2.3490
	S_{EWMA}	46.5036	12.9945	7.8245	3.9266	2.8015	2.1159	1.7297	1.5157
1.50	ARL	25.7850	10.2690	7.3010	4.2440	3.2120	2.6560	2.4160	2.0730
	S_{EWMA}	22.7360	7.6592	5.4525	2.8097	2.0740	1.6584	1.4809	1.1649
2.00	ARL	9.8560	5.4880	4.2280	2.8730	2.3310	1.9300	1.7700	1.6550
	S_{EWMA}	7.7938	3.4278	2.6512	1.6600	1.3180	1.0355	0.9571	0.8711
2.50	ARL	5.6970	3.5640	2.9100	2.1300	1.8360	1.5320	1.4510	1.3650
	S_{EWMA}	3.5864	2.0455	1.5499	1.1173	0.9518	0.7235	0.6783	0.5866
3.00	ARL	4.0410	2.7950	2.4030	1.8330	1.5530	1.3910	1.2860	1.2020
	S_{EWMA}	2.2073	1.4018	1.2329	0.8636	0.6839	0.6087	0.5220	0.4306
3.50	ARL	3.1450	2.2910	2.0290	1.5890	1.3960	1.2470	1.1700	1.1280
	S_{EWMA}	1.6747	1.0989	0.9513	0.7312	0.5979	0.4798	0.3966	0.3574
4.00	ARL	2.7080	1.9150	1.8030	1.4050	1.2440	1.1430	1.1000	1.0780
	S_{EWMA}	1.2878	0.8824	0.8043	0.5704	0.4782	0.3587	0.3132	0.2720

Tabla 21 Desempeño considerando diferentes valores de α_1 y β_1 para muestras de la Gamma utilizando el uso secuencial del EWMA y el MLE para estimar el punto de cambio. Los parámetros son $\tau = 100$; 1,000 simulaciones; $\alpha_0 = 1, \beta_0 = 1, \lambda = 0.20$ y $L = 2.962$; el sesgo ($\hat{t}_{MLE} - \tau$) y desviación estándar, S_{MLE} , son calculados.

α_1	β_1	1.00	1.30	1.50	2.00	2.50	3.00	3.50	4.00
1.00	<i>Sesgo</i>	171.37	27.84	10.60	1.80	-0.27	0.24	-0.36	-0.52
	S_{MLE}	181.65	37.03	21.23	13.82	12.13	6.80	9.25	7.77
1.30	<i>Sesgo</i>	11.89	2.18	0.07	-0.46	-1.05	-0.83	-1.16	-0.89
	S_{MLE}	32.45	15.59	12.96	9.30	8.82	7.24	8.56	7.03
1.50	<i>Sesgo</i>	-0.12	-1.89	-1.69	-1.27	-1.40	-0.80	-1.48	-0.96
	S_{MLE}	19.20	14.41	12.65	9.16	8.65	5.74	9.11	7.64
2.00	<i>Sesgo</i>	-3.76	-2.10	-2.06	-1.45	-1.49	-1.24	-0.90	-0.78
	S_{MLE}	14.01	10.05	9.59	8.36	8.11	7.43	4.91	4.73
2.50	<i>Sesgo</i>	-3.49	-2.23	-1.76	-1.55	-0.85	-0.95	-0.55	-0.75
	S_{MLE}	12.58	8.83	7.53	7.05	5.49	5.45	3.02	4.83
3.00	<i>Sesgo</i>	-2.76	-1.56	-1.71	-1.28	-0.69	-0.65	-0.40	-0.33
	S_{MLE}	9.53	6.77	8.34	6.96	4.45	4.28	2.03	2.06
3.50	<i>Sesgo</i>	-2.15	-1.59	-1.55	-0.85	-0.58	-0.57	-0.34	-0.41
	S_{MLE}	8.85	7.10	7.31	4.68	3.71	4.25	1.70	2.85
4.00	<i>Sesgo</i>	-2.29	-1.25	-1.33	-0.82	-0.49	-0.55	-0.34	-0.28
	S_{MLE}	9.71	5.48	6.29	4.52	2.45	3.76	2.36	2.64

Tabla 22a Comparativa del desempeño de la carta de control EWMA y el uso secuencial de la EWMA y el MLE para el punto de cambio. Presenta variaciones en β de 1.00 a 2.00. Se han incluido el valor de μ y σ . Esta tabla está basada en las tablas 20 y 21 y permite comparar la detección del punto de cambio de ambos estimadores para diferentes parámetros α y β pero con misma μ o σ .

β	1.00	1.30	1.50	2.00
α	(μ, σ)	(μ, σ)	(μ, σ)	(μ, σ)
1.00	(1.00, 1.00)	(1.00, 1.30)	(1.00, 1.50)	(1.00, 2.00)
(ARL, Sesgo)	(178.89, 171.37)	(37.59, 27.84)	(20.34, 10.60)	(9.38, 1.80)
1.30	(1.3, 1.14)	(1.69, 1.48)	(1.95, 1.71)	(2.60, 2.28)
(ARL, Sesgo)	(48.05, 11.89)	(15.76, 2.18)	(9.93, 0.07)	(5.55, -0.46)
1.50	(1.50, 1.22)	(1.95, 1.59)	(2.25, 1.83)	(3.00, 2.44)
(ARL, Sesgo)	(25.79, -0.12)	(10.27, -1.89)	(7.30, -1.69)	(4.24, -1.27)
2.00	(2.00, 1.41)	(2.60, 1.83)	(3.00, 2.12)	(4.00, 2.82)
(ARL, Sesgo)	(9.86, -3.76)	(5.49, -2.10)	(4.23, -2.06)	(2.87, -1.45)
2.50	(2.50, 1.58)	(3.25, 2.05)	(3.75, 2.37)	(5.00, 3.16)
(ARL, Sesgo)	(5.70, -3.49)	(3.56, -2.23)	(2.91, -1.76)	(2.13, -1.55)
3.00	(3.00, 1.73)	(3.90, 2.25)	(4.50, 2.58)	(6.00, 3.46)
(ARL, Sesgo)	(4.04, -2.76)	(2.80, -1.56)	(2.40, -1.71)	(1.83, -1.28)
3.50	(3.50, 1.87)	(4.55, 2.43)	(5.25, 2.80)	(7.00, 3.74)
(ARL, Sesgo)	(3.15, -2.15)	(2.29, -1.59)	(2.03, -1.55)	(1.59, -0.85)
4.00	(4.00, 2.00)	(5.20, 2.60)	(6.00, 3.00)	(8.00, 4.00)
(ARL, Sesgo)	(2.71, -2.29)	(1.92, -1.25)	(1.80, -1.33)	(1.41, -0.82)

Tabla 23b Comparativa del desempeño de la carta de control EWMA y el uso secuencial de la EWMA y el MLE para el punto de cambio. Presenta variaciones en β de 1.00 a 2.00. Se han incluido el valor de μ y σ . Esta tabla está basada en las tablas 20 y 21 y permite comparar la detección del punto de cambio de ambos estimadores para diferentes parámetros α y β pero con misma μ o σ .

β	2.50	3.00	3.50	4.00
α	(μ, σ)	(μ, σ)	(μ, σ)	(μ, σ)
1.00	(1.00, 2.50)	(1.00, 3.00)	(1.00, 3.50)	(1.00, 4.00)
(ARL, Sesgo)	(6.07, -0.27)	(4.53, 0.24)	(3.82, 0.36)	(3.27, -0.52)
1.30	(3.25, 2.85)	(3.90, 3.42)	(4.55, 3.99)	(5.20, 4.56)
(ARL, Sesgo)	(3.98, -1.05)	(3.13, -0.83)	(2.77, -1.16)	(2.35, -0.89)
1.50	(3.75, 3.06)	(4.50, 3.67)	(5.25, 4.28)	(6.00, 4.89)
(ARL, Sesgo)	(3.21, -1.40)	(2.66, -0.80)	(2.42, -1.48)	(2.07, -0.96)
2.00	(5.00, 3.53)	(6.00, 4.24)	(7.00, 4.94)	(8.00, 5.65)
(ARL, Sesgo)	(2.33, -1.49)	(1.93, -1.24)	(1.77, -0.90)	(1.66, -0.78)
2.50	(6.25, 3.95)	(7.50, 4.74)	(8.75, 5.53)	(10.00, 6.32)
(ARL, Sesgo)	(1.84, -0.58)	(1.53, -0.95)	(1.45, -0.55)	(1.37, -0.75)
3.00	(7.50, 4.33)	(9.00, 5.19)	(10.50, 6.06)	(12.00, 6.92)
(ARL, Sesgo)	(1.55, -0.69)	(1.39, -0.65)	(1.29, -0.40)	(1.20, -0.33)
3.50	(8.75, 4.67)	(10.50, 5.61)	(12.25, 6.54)	(14.00, 7.48)
(ARL, Sesgo)	(1.40, -0.58)	(1.25, -0.57)	(1.17, -0.34)	(1.13, -0.41)
4.00	(10.00, 5.00)	(12.00, 6.00)	(14.00, 7.00)	(16.00, 8)
(ARL, Sesgo)	(1.24, -0.49)	(1.14, -0.55)	(1.10, -0.34)	(1.08, -0.28)

En las Tablas 22a y 22b es posible observar como las estimaciones hechas por el uso secuencial de la carta de control EWMA y el MLE para el tipo de cambio son más cercanas a las estimaciones que las hechas por la carta EWMA. De esta tabla se observa que para cambios mayores a $\beta_1/\beta_0 \geq 2.0$ el sesgo esta en el intervalo de (-1.55, 1.8) mientras que para $\alpha_1/\alpha_0 \geq 2.0$ en (-3.76, -0.28). Las estimaciones obtenidas de la carta EWMA definitivamente tienen un ancho mayor y presentan una distancia al punto de cambio superior a una unidad. Un resultado de esta tabla es que en los casos extremos se puede observar que la última columna, cuando el cociente de los β 's es igual a 4, se mantiene a menos de una unidad del τ , sin importar el valor del cociente de los α 's, mientras que en la última fila donde el cociente de los α 's es igual a 4 la distancia de las estimaciones al valor τ sí alcanza valores superiores a una unidad. Esto indica, hasta donde los valores de las simulaciones muestran, que el uso secuencial de la EWMA y del MLE para el punto de cambio es más sensible a cambios en el parámetro de escala β .

5.8 Conclusiones y Trabajos Futuros

Esta investigación ha propuesto una metodología para ser usada en un conjunto de observaciones que pueden ser modeladas mediante una función de distribución Gama con el fin de detectar el punto de cambio. La propuesta consiste en usar una carta de control para la Gamma seguida por una estimación del punto de cambio utilizando el estimador de máxima verosimilitud para detectar el momento en que el cambio ocurre cuando ambos parámetros de la función de distribución Gamma cambian, tanto la forma como la escala, al mismo tiempo. Los resultados experimentales muestran como la integración de ambas herramientas mejoran la estimación del punto de cambio. La estimación del sesgo es reducida cuando se usa el MLE propuestosin embargo el error estándar es mayor. Para reducir este valor es necesario explorar varias opciones, tales como evaluar el desempeño utilizando diferentes tamaños de muestra, o bien utilizando observaciones extras después de que el cambio ha sido detectado. Los trabajos futuros considerarán estas opciones. Como trabajo futuro también se propone analizar series de datos que puedan ser modeladas mediante una función de distribución Gamma, y estimar los puntos de cambio cuando existen varios de ellos y los parámetros iniciales son desconocidos.

CAPITULO 6. CONCLUSIONES Y TRABAJO FUTURO

6.1 Conclusiones

En la constante búsqueda de la mejora continua de la calidad la gente involucrada en el SPC, e investigadores de diferentes áreas, se han esforzado en crear nuevas herramientas que auxilien en la consecución de este objetivo, desde luego esto implica una paradoja ya que una vez que se ha logrado hacer una mejora en la calidad, inmediatamente habrá alguien que desarrolle una nueva herramienta que mejore lo que hasta el momento se tiene. La mejora continua de la calidad es una filosofía que se aplica en cualesquier proceso, incluso en los procesos de vida y que por su misma naturaleza no tiene fin. La calidad ha evolucionado desde estados pírricos donde la satisfacción del bien producido era exclusivamente definida por el productor hasta estados actuales que incluyen a todos los actores en el proceso de diseño-producción-utilización-satisfacción-retroalimentación de un bien y/o servicio. En el estadio actual de la calidad el ciclo anterior es un ciclo sin fin, la competencia por conseguir el favor del cliente y su lealtad ha hecho que la calidad sea un factor inherente en cualquier producto y/o servicio. Gracias a ello y a la conciencia que las organizaciones han ido tomando de esto han provocado que más gente esté interesada en aportar nuevas ideas y, en consecuencia, nuevas herramientas para conseguir que los productos tengan una ventaja competitiva ante sus competidores por el gusto del cliente al ofrecer productos altamente confiables, siendo uno de los estándares de esa confiabilidad presentar productos que no tengan una variabilidad en su desempeño y/o uso. Aquí es donde este trabajo toma valor ya que el tener una herramienta adecuada para monitorear un proceso garantiza que el mismo cumpla con las especificaciones para las que fue creado, también lo es que en el momento en que el proceso no cumple con las expectativas es necesario determinar las causas que lo llevaron a no dar lo que se espera de él. El identificar la causa que provoca la variabilidad del proceso implica una serie de gastos en tiempo, dinero y esfuerzo que puede ser disminuido si se tiene una idea de en qué momento buscar esa causa. El problema de estimar ese momento atañe al control estadístico de procesos en lo correspondiente al análisis del punto de cambio. La solución de este problema conlleva una serie de ventajas pertinentes a los costos, de toda índole, en que se incurre al tratar de identificar la causa que provoca que el proceso salga de control estadístico. El contar con una herramienta confiable que indique el momento preciso en que el cambio ocurre será de gran ayuda a los involucrados en un departamento de calidad que necesita identificar la causa

asignable de variación. Si bien, identificar el momento exacto en que el cambio en el proceso ocurrió es el objetivo esto requiere ser hecho mediante análisis de las observaciones y es posible que no se determine exactamente, sin embargo una buena aproximación de ese momento ayudará a detectar la causa y ello es el objetivo final previo a la realización de una acción correctiva y/o preventiva para volver al proceso a estar en control estadístico.

La presente tesis ha tomado el problema de estimación del punto de cambio y ha presentado un modelo que permite a los involucrados con el departamento de calidad contar con una herramienta más para la generación de estimadores de punto de cambio de acuerdo a la serie de datos que están recolectando de sus procesos. La propuesta de utilizar de manera secuencial una carta de control estadístico y, a la señal de alarma de ésta, obtener una estimación mediante el uso del MLE para el punto de cambio ha sido probada en una de las funciones de distribución más utilizada en los procesos de todo tipo. La distribución normal aparece en gran cantidad de eventos y/o procesos de la vida real y ha sido ampliamente estudiada. El CPA ha tomado a la Normal como una de las principales receptoras de sus aportaciones y esta tesis contribuye al hacer su aportación al respecto. Los resultados obtenidos con esta herramienta de uso secuencial aplicado a la Normal son altamente satisfactorios y han sido detallados de forma individual en cada una de las primeras dos investigaciones (cap. 3 y 4) en donde se han considerado cambios en los parámetros que la definen. La otra función de distribución considerada fue la Gamma, que aunque menos estudiada que la Normal, es un referente en aquellos procesos donde el tiempo entre llegadas es determinante. Los investigaciones en el CPA que tienen como motivo central de interés las observaciones independientes y Gamma no son tan numerosas como en el caso de la Normal pero presentan aplicaciones y retos interesantes para los involucrados con el SPC ya que una de las características de esta función de distribución es la forma en que la media es calculada; al depender del producto de sus dos parámetros se presenta el caso en que para una misma media los parámetros pueden ser diferentes, esto implica tener la posibilidad de detectar un cambio en los mismos a pesar que la media siga teniendo el mismo valor. La misma situación se presenta para la varianza, también está en función de sus parámetros y se puede dar el caso de valores iguales de varianza para parámetros distintos.

En los casos estudiados para la función de distribución Normal cuando el parámetro que cambia es la μ se han considerado desplazamientos en la media que van de 0σ a 3σ . Como variante se simularon los casos en que se tomaron observaciones extras a cuando la señal de alerta es emitida por la carta de control, los valores considerados fueron 5, 10, 15 y hasta 20 observaciones extras, además de variar el tamaño del subgrupo en 1, 3 y 5 observaciones obteniendo resultados satisfactorios ya que en cada una de las variaciones que se hizo el estimador arrojó resultados iguales o mejores que el caso donde se tuvo una observación por subgrupo y cero datos extras. Si bien las estimaciones obtenidas por el estimador propuestos

siempre fueron más cercanas al punto de cambio que las ofrecidas por el estimador de la carta CUSUM el resultado de las desviaciones estándar muestra que a partir de 1.5σ la desviación estándar de las desviaciones del estimador propuesta son menores a las desviaciones estándar de las desviaciones del estimador de la CUSUM.

En los casos estudiados para la función de distribución Normal cuando el parámetro que cambia es la σ se han considerado proporciones de 1.0 a 3.0 entre la desviación antes y después del cambio. Como variante se simularon los casos en que se tomaron observaciones extras a cuando la señal de alerta es emitida por la carta de control, los valores considerados fueron 5, 10, 15 y 20 observaciones extras, además de variar el tamaño del subgrupo en 1, 3 y 5 observaciones obteniendo resultados satisfactorios ya que en cada una de las variaciones que se hizo el estimador arrojó resultados por debajo de la caso de una observación por subgrupo y cero datos extras. Si bien las estimaciones obtenidas por el estimador propuestos siempre fueron más cercanas al punto de cambio que las ofrecidas por el estimador de la carta CUSUM el resultado de las desviaciones estándar muestra que a partir de una proporción de 1.2 la desviación estándar de las desviaciones del estimador propuesto son menores a las desviaciones estándar de las desviaciones del estimador de la CUSUM.

En el caso de la función de distribución Gamma se consideraron proporciones para α 's y β 's que van de 1.0 a 4.0. Las simulaciones muestran que las estimaciones obtenidas del estimador propuesto son más sensibles que el correspondiente estimador de la carta EWMA ya que en todos los casos el valor obtenido por el uso secuencial de la carta EWMA y el MLE para el punto de cambio presento mayor cercanía que su equivalente. Sin embargo de las Tablas 20 y 21 se puede observar que la desviación estándar de las estimaciones del estimador obtenido de la carta EWMA es menor que su contraparte, salvo en dos casos, por lo que este método debe ser tomado con reservas mientras no se hagan investigaciones sobre como disminuir esta desviación estándar. En la sección siguiente "Trabajo Futuro" se proponen algunas líneas de investigación orientadas a atender esta situación.

Resumiendo, esta investigación ha propuesto el uso secuencial de una carta de control y el MLE para el punto de cambio para mejorar el desempeño de los estimadores del mismo. La propuesta de utilizar una carta de control y, cuando ésta emita la señal de proceso fuera de control, proceder a utilizar el estimador obtenido por el MLE para punto de cambio ha sido respaldada por simulaciones particularizadas en la distribución Normal y la distribución Gamma. Las simulaciones han mostrado que los estadísticos obtenidos como resultado de las estimaciones generadas por el estimador propuesto son menores a las mismas para las estimaciones generadas por el estimador obtenido mediante cartas de control CUSUM de auto-inicio y EWMA lo cual permite ofrecer una herramienta al personal responsable de la calidad de producto que le auxilie en la detección de las causas asignables que provocan que el proceso se salga de control mediante una estimación del momento en que inicia el cambio. En base a las Hipótesis planteadas en cada investigación y que han sido el eje de esta tesis se presentan los siguientes resultados:

- Investigación 1 (Capítulo 3):
 - Hipótesis 1: El estimador obtenido mediante la aplicación secuencial de una carta de control CUSUM de auto-inicio y un MLE para el punto de cambio en un conjunto de datos independientes caracterizado por una función de distribución Normal con parámetros iniciales μ_0 y σ_0 y que a partir del instante desconocido τ cambian a μ_1 y σ_1 con la característica que $\mu_0 \neq \mu_1$ y $\sigma_0 = \sigma_1$, todos ellos desconocidos, tiene un mejor desempeño que el estimador del punto de cambio obtenido por de la carta CUSUM de auto-inicio.
 - Respuesta: Se acepta la hipótesis 1 ya que siguiendo los resultados de las simulaciones realizadas en la investigación 1 de la Tabla 3 se observa que el sesgo, el promedio de las desviaciones entre las estimaciones generadas por el estimador propuesto y el valor real del cambio, es menor que el ARL, el equivalente para las estimaciones generadas por el estimador de la CUSUM de auto-inicio, en todos los casos. Esta situación prevalece sin importar el tamaño del subgrupo (Tabla 4), ni por la cantidad de valores extras tomados después de que la señal de alarma es provista por la carta de control CUSUM (Tabla 5 y Tabla 6) en donde se ha considerado que τ toma los valores de 50 y 100 respectivamente.
 - Respecto a la desviación estándar de ambos estimadores se observa de la tabla 3 tercer fila y la equivalente en la Tabla 7 que la desviación estándar de las estimaciones del $\hat{\tau}_{MLE}$ empiezan siendo mayores al $\hat{\tau}_{CUSUM}$ pero a partir de una desviación de la media original de $1.5\sigma_0$ empiezan a ser menores.

- Las respuestas anteriores permiten recomendar el \hat{t}_{MLE} sobre el \hat{t}_{CUSUM} en general, y en particular para cuando la desviación de la media es de al menos $1.5\sigma_0$, ya que el desempeño del \hat{t}_{MLE} ha superado al del \hat{t}_{CUSUM} .
- Investigación 2 (Capítulo 4):
 - Hipótesis 2: El estimador obtenido mediante la aplicación secuencial de una carta de control CUSUM de auto-inicio y un MLE para el punto de cambio en un conjunto de datos independientes caracterizado por una función de distribución Normal con parámetros iniciales μ_0 y σ_0 que a partir del instante desconocido τ cambian a μ_1 y σ_1 con la característica que $\mu_0 = \mu_1$ y $\sigma_0 \neq \sigma_1$, todos ellos desconocidos, tiene un mejor desempeño que el estimador del punto de cambio obtenido por la carta CUSUM de auto-inicio.
 - Respuesta: Se acepta la Hipótesis 2 ya que siguiendo los resultados de las simulaciones realizadas en la investigación 2 de la Tabla 11 se observa que el sesgo, el promedio de las desviaciones entre las estimaciones generadas por el estimador propuesto y el valor real del cambio, es menor que el ARL, el equivalente para las estimaciones generadas por el estimador de la CUSUM de auto-inicio, en todos los casos. Esta situación prevalece sin importar el tamaño del subgrupo (Tabla 12), ni por la cantidad de valores extras tomados después de que la señal de alarma es provista por la carta de control CUSUM (Tabla 13 y Tabla 14) en donde se ha considerado que τ toma los valores de 50 y 100 respectivamente.
 - Respecto a la desviación estándar de ambos estimadores se observa de la tabla 11 tercer fila y la equivalente en la Tabla 15 que la desviación estándar de las estimaciones del \hat{t}_{MLE} son menores a las del \hat{t}_{CUSUM} a partir de una proporción de la desviación original de 1.1.
 - Las respuestas anteriores permiten recomendar el \hat{t}_{MLE} sobre el \hat{t}_{CUSUM} en general, y en particular para cuando la proporción de la desviación es 1.1.
- Investigación 3 (Capítulo 5):
 - Hipótesis 3: El estimador obtenido mediante la aplicación secuencial de una carta de control EWMA y un MLE para el punto de cambio en un conjunto de datos caracterizado por una función de distribución Gamma con parámetros iniciales

conocidos α_0 y β_0 que a partir del instante desconocido τ cambian a α_1 y β_1 , ambos desconocidos, con la característica que $\alpha_0 \neq \alpha_1$ y/o $\beta_0 \neq \beta_1$, tiene un mejor desempeño que el estimador del punto de cambio obtenido por la carta EWMA.

- Respuesta: Se acepta la Hipótesis 3 con reservas ya que siguiendo los resultados de las simulaciones realizadas en la investigación 3 de la Tabla 20 y 21 se observa que el sesgo, el promedio de las desviaciones entre las estimaciones generadas por el estimador propuesto y el valor real del cambio, es menor que el ARL, el equivalente para las estimaciones generadas por el estimador de la EWMA, en todos los casos, salvo en el caso de falsa alarma. Sin embargo la desviación estándar de las estimaciones obtenidas del estimador propuesto son mayores que su contraparte de la carta EWMA.
- La Tabla 22 es obtenida de las Tablas 20 y 21 y se le agrega el valor de la media para la Gamma, en ella se puede observar que el estimador propuesto es más sensible a los cambio es el parámetro β que al parámetro α . En particular para los casos en que se tiene la misma μ , calculada a partir de diferentes valores de α y β , un mayor cambio en β genera un promedio de estimaciones menor los mismos promedios para valores de α mayores.
- Las respuestas anteriores permiten recomendar, con limitaciones en cuanto a la desviación estándar, el \hat{t}_{MLE} sobre el \hat{t}_{EWMA} ya que el desempeño para el Sesgo del MLE es menor al del ARL. En particular cuando $\alpha_0 \neq \alpha_1$ y $\beta_0 = \beta_1$.

6.2 Trabajo Futuro

En el desarrollo del presente trabajo se confirmó que el análisis de punto de cambio es un área del control estadístico de procesos que se mantiene en constante evolución. A pesar de haberse iniciado a principios de los 50's, los avances que se han logrado son abundantes y han sido enriquecidos por diferentes áreas del conocimiento, desde el análisis del problema del punto de cambio basado en la estadística clásica hasta los que se ocupan del mismo desde un enfoque de redes neuronales; los tratamientos son diversos, con ventajas y desventajas según el uso o enfoque que se le quiera dar a la herramienta que proponen. Es por ello que hablar de trabajos futuros en el CPA, eje central de la presente investigación, es tocar un tema muy amplio, con muchas líneas de aplicación, investigación y cada una con sus propias variantes, razón por la cual el tema será centrado en las líneas de investigación que siguió esta tesis y que podrían acotarse a cuatro frentes de trabajo, que definitivamente no son únicos ni exclusivos:

1.- Disminuir la variabilidad del MLE. . Un problema latente es que algunos estimadores obtenidos presentaron una desviación estándar mayor a los ya conocidos, si bien el valor

promedio de sus estimaciones fue mejor que el equivalente en la carta de control, al tener una variabilidad mayor se presenta un área de oportunidad que es conveniente no pasar por alto. Es por ello que disminuir el tamaño de la desviación estándar de las estimaciones generadas por estimadores obtenidos a partir del modelo propuesto es importante con el fin de mejorar su desempeño. Una posible solución es considerar eliminar los k -primeros y m -últimos datos de la serie de observaciones esperando con ello evitar proponer estimaciones que de entrada se sabe no ocurrirán, como es el caso de tener un cambio en el proceso antes de los primeros 10 datos, sobre todo si el proceso tuvo una Fase I. El poder determinar los valores k y m que permitan disminuir la variabilidad en las estimaciones es una alternativa a considerar.

2.- Uso de diferentes estimadores del punto de cambio. . Una alternativa al uso del MLE para el punto de cambio es el uso del “Estimador del Centroide de Verosimilitud” [103](CLE, por sus siglas en inglés) cuyo objetivo es que ofrezca un estimador del punto de cambio con un sesgo y desviación estándar menor que otros estimadores. En el caso particular de esta investigación la propuesta es desarrollar los CLE del punto de cambio y utilizarlos siguiendo el modelo propuesto, esto es; utilizar una carta de control hasta que mande una señal que indique que el proceso ha salido de control, tomar el conjunto de las observaciones hasta ese momento y proceder a obtener el CLE. Sería interesante conocer el rendimiento del estimador así obtenido con el estimador obtenido por el uso secuencial de una carta de control y el MLE del punto de cambio. Las comparaciones serían entre estos dos bajo los mismos esquemas.

3.- Estudiar las funciones de distribución que no han sido tan atendidas como la Normal. Las funciones de distribución son muchas y variadas y a la fecha la que más ha sido estudiada por el SPC y el CPA es la Normal, tal vez seguida por la Gamma, la Binomial y la Poisson, dejando abierta la puerta a otras que apenas han empezado a ser consideradas. Es por ello que otro frente de trabajo sería abordar o re-abordar las funciones de distribución que no han sido ampliamente atendidas por el SPC y el CPA. La Erlang y la Exponencial serían la extensión natural de este trabajo dada la relación que existe entre ellas y la Normal y la Gamma. Desde luego, esto estará sujeto a las aplicaciones en la vida diaria que tengan las mismas.

4.- Considerar más de un cambio en los parámetros y en más de un parámetro. Este trabajo se centró en considerar solo un cambio en una serie de observaciones lo cual tiene un sentido lógico al considerar que un proceso que salga de control estadístico debe ser atendido para tomar las acciones correctivas y/o preventivas que lo vuelvan a un estado de variabilidad controlada, sin embargo esto está sujeto a que la señal que indica que el proceso se encuentra fuera de control sea detectada justo al primer cambio. Lo anterior implica considerar la posibilidad de que un proceso tenga varios cambios que lo lleven a un estado fuera de control antes de que alguna señal sea detectada. En la Fase I del control estadístico siempre existe el riesgo de tener más de un cambio, y desarrollar técnicas de CPA ayudaría a solventar esta

necesidad. Esta área de oportunidad presenta un gran reto que está empezando a ser considerado por los investigadores del SPC y CPA.

Deseo cerrar insistiendo en que el CPA presenta muchas oportunidades de desarrollo y que las que señalo están basadas en la orientación que siguió este trabajo, en ningún momento puedo afirmar que sean las únicas. El cuestionarnos la razón de las cosas genera conocimiento al tratar de explicarlas lo cual genera nuevas preguntas que a su vez genera nuevo conocimiento, todo ello buscando mejorar nuestros niveles de vida; el ciclo se repite cada vez y la filosofía de la mejora continua de la calidad se hace presente de nuevo.

Bibliografía

-] G. Bounds, L. Yorks y M. Adams, *Beyond Total Quality Management: Toward the Emerging Paradigm*, McGraw-Hill, 1994.
- [2] H. Cantú-Delgado, *Desarrollo de una Cultura de Calidad*, McGraw-Hill Interamericana, 2006.
- [3] W. Shewhart, *Economic Control of Quality of Manufactured Product 50th Anniversary*, Asq Press, 1931.
- [4] E. S. Page, «Continuos inspection schemes,» *Biometrika*, vol. vol. 41, nº no. 1/2, pp. 100-115, Jun 1954.
- [5] D. M. Hawkins, «Self-starting cusum charts for location and scale,» *The Statistician*, vol. 36, pp. 299-315, 1987.
- [6] D. C. Montgomery, *Control Estadístico de la Calidad*, México: John Wiley & Sons, Ltd, 2007.
- [7] S. W. Roberts, «Control Chart Tests Based on Geometric Moving Averages,» *Technometrics*, vol. 1, nº 3, pp. 239-250, 1959.
- [8] M. A. Girshick y H. Rubin, «A Bayes Approach to a Quality Control Model,» *The Annals of Mathematical Statistics*, vol. 23, nº 1, pp. 114-125, 1952.
- [9] D. V. Hinkely, «Inference about the change-point from cumulative sum tests,» *Biometrika*, vol. 58, nº 3, pp. 509-523, December 1971.
- [10] D. M. Hawkins y K. D. Zamba, «Statistical Process Control for Shifts in Mean or Variance Using a Changepoint Formulation,» *Technometrics*, vol. 47, nº 2, pp. 164-173, 2005.
- [11] R. Amin y M. Reynoldds, «Nonparametric Quality Control Charts Based on the Sign Statistics,» *Communications in Statistcs. Part A: Theory and Methods*, vol. 24, nº 6, pp. 1597-1623, 1996.
- [12] N. Das y A. Bhattacharya, «A New Non-Parametric Control Chart for Controlling Variability,» *Quality Technology of Quantitative Management*, vol. 5, nº 4, pp. 351-361, 2008.
- [13] V. Tercero, M. d. C. Temblador-Pérez, M. Beruvides y A. A. Hernández-Luna, «Nonparametric Estimator for the Time of a Step Change in the Trend of Random Walk Models with Drift,» *Quality and Reliability Engineering International*, vol. 28, nº 3, 08 Feb 2012.
- [14] C. Acosta-Mejia, J. Pignatiello y B. Rao, «A comparasion of control charting procedures for monitoring process dispersion,» *IIE Transactions*, vol. 31, nº 6, pp. 569-579, 1999.

- [15] J. Garza-Venegas, A. Cordero-Franco, V. Tercero-Gómez y J. Camacho-Vallejo, «Estimación de Múltiples Puntos de Cambio en Series de Tiempo con Distribución Normal Utilizando un Algoritmo Heurístico de Construcción,» de *Congreso Latino-Iberoamericano de Investigación Operativa*, Río de Janeiro, Brazil, 2012.
- [16] D. A. Hsu, «Detecting Shifts of Parameter in Gamma Sequences with Applications to Stock Price and Air Traffic Flow Analysis,» *Journal of the American Statistical Association*, vol. 74, nº 365, pp. 31-40, 1979.
- [17] R. Jiang, «Determining B_X Life of a Normal Component Based on Change Point of Failure Rate,» de *System Science, Engineering Design and Manufacturing Informatization*, 2010.
- [18] K. Kyungsook y S. Young, «Bayesian change point analysis for a sequence of normal observations: application to the winter average temperature in Seoul.,» *Korean J. Appl. Statist.*, vol. 17, nº 2, pp. 281-289, 2004.
- [19] G. Box y A. Luceño, *Statistical Control by Monitoring and Feedback Adjustment*, New York: Wiley.
- [20] M. Koutras, S. Bersimis y P. Maravelakis, «Statistical Process Control using Shewhart Control Charts with Supplementaru Runs Rules,» *Methodology and Computing in Applied Probability*, vol. 9, nº 2, pp. 207-224, June 2007.
- [21] E. Page, «On problems in wich a change in a parameter occurs at an unknown point,» *Biometrika*, vol. 44, nº 1-2, pp. 248-252, 1957.
- [22] A. Shiryaev, «On Optimum Methods in Quickest Detection Problems,» *Theory of Probability and Its Applications*, vol. 8, nº 1, pp. 26-51, 1963.
- [23] A. Smith, «A bayesian approach to inference about a change-point in a sequence of random variables,» *Biometrika*, nº 62, pp. 407-416, 1975.
- [24] A. Smith y D. Cook, «Straight Lines with a Change Point: a Bayesian Analysis of Some Renal Transplant Data,» *Applied Statistics*, nº 29, pp. 180-189, 1980.
- [25] E. Gombay y L. Horvarth, «Asymptotic Distributions of Maximum Likelihood Tests for Change in the Mean,» *Biometrika*, vol. 77, pp. 411-414, 1990.
- [26] E. Gombay y L. Horvath, «An Application of the Maximum Likelihood Test to the Change-Point Problem,» *Stochastic Processes and their Applications*, vol. 904, nº 50, pp. 161-171, 1994.
- [27] A. Raferty y V. Akman, «Bayesian Analysis of a Poisson Process with a Change-Point,» *Biometrika*, nº 73, pp. 85-89, 1986.

- [28] T. Lai y H. Xing, «Sequential Change-point detection when the Pre- and Post- Change Parameters are Unknown,» *Sequential Analysis: Design Methods and Applications*, vol. 29, nº 2, pp. 162-175, 2010.
- [29] E. Page, «Cumulative Sum Charts,» *Technometrics*, vol. 3, nº 1, pp. 1-9, February 1961.
- [30] D. Hinkley, «Inference about the change-point in a sequence of random variables,» *Biometrika*, vol. 57, nº 1, pp. 1-17, 1970.
- [31] K. Nishina, «A Comparison of Control Charts from the Viewpoint of Change Point Estimation,» *Quality and Reliability Engineering International*, vol. 8, nº 6, pp. 537-541, 1992.
- [32] T. Samuel, J. Pignatiello y J. Calvin, «Identifying the Time of a Step Change with \bar{x} Control Charts,» *Quality Engineering*, vol. 10, nº 3, pp. 521-527, 1998a.
- [33] Y. Shao, C. Hou y H. Wang, «Estimation of the Change Point of a Gamma Process by Using the S control Chart and MLE,» *Journal of the Chinese Institute of Industrial Engineers*, vol. 23, nº 3, pp. 207-214, 2006.
- [34] Z. Li, J. Zhang y Z. Wang, «Self-starting control chart for simultaneously monitoring process mean and variance,» *International Journal of Production research*, vol. 48, nº 15, pp. 4537-4553, 2010.
- [35] Y. Shao y C. Hou, «A combined MLE and EWMA chart approach to estimate the change point of a Gamma process with individual observations,» *International Journal of Innovate Computing, Information and Control*, vol. 7, nº 5, pp. 2109-2122, 2011.
- [36] E. Page, «A test for a Change in a Parameter Occurring at an Unknown Point,» *Biometrika*, nº 42, pp. 523-527, 1955.
- [37] G. Bhattacharyya y R. Johnson, «Nonparametric Tests for Shift at an Unknown Time Point,» *The Annals of Mathematical Statistics*, vol. 5, nº 39, pp. 1731-1743, 1968.
- [38] A. Pettit, «A Non-Parametric Approach to the Change-Point Problem,» *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 2, nº 28, pp. 126-135, 1979.
- [39] D. Hinkley y E. Schechtman, «Conditional bootstrap methods in the mean shift model,» *Biometrika*, vol. 74, nº 1, pp. 85-93, 1987.
- [40] G. W. Cobb, «The Problem of the Nile: Conditional Solution to a Change Point Problem,» *Biometrika*, vol. 2, nº 65, pp. 243-251, 1978.
- [41] J. Koziol, «A Note on Non-Parametric tests for the Change-Point Problem,» *Biometrical Journal*, vol. 29, nº 7, pp. 791-794, 1987.

- [42] G. Guverich y A. Vexler, «Retrospective Change Point Detection: from Parametric to distribution Free Policies,» *Communications in Statistics-Simulation and Computation*, vol. 39, nº 5, pp. 899-920, 2010.
- [43] L. Hovart y P. Koskoszka, «Change Point Detection with Non-Parametric regression,» *Statistics*, vol. 36, nº 1, pp. 9-31, 2002.
- [44] N. Neumeyer y I. Van Keilegom, «Change-Point Tests for the Error Distribution in Non-parametric Regression,» *Scandinavian Journal of Statistics*, vol. 36, nº 3, pp. 518-5541, 2009.
- [45] A. S. Perez-Blanco, A. E. Cordero-Gómez y V. Tercero-Gómez, «Punto de Cambio: Estado Actual y Tendencias,» de *2do Foro de Divulgación Científica y Tecnológica*, Monterrey, 2012.
- [46] H. Chernoff y S. Zacks, «Estimating the Current Mean of a Normal Distribution wich is Subjected to Changes in Time,» *The Annals of Mathematical Statistics*, vol. 35, nº 3, pp. 999-1018, 1964.
- [47] Z. Kander y S. Zacks, «Test Procedures for Possible Changes in parameters of Statistical Distributions Occurring at Unknown Time Points,» *The Annals of MAtheMatical Statistics*, vol. 37, nº 5, pp. 1196-1210, 1966.
- [48] A. Smith, «A BAYesian Approach to Inference About a Change Point in a Sequence of Random Variables,» *Biometrika*, vol. 62, nº 2, pp. 407-416, 1975.
- [49] D. Hawkins, «A CUSUM for a Scale Parameter,» *Journal of Quality and Technology*, vol. 13, nº 4, pp. 228-231, 1981.
- [50] L. Nelson, «The Shewhart control chart- tests for special causes,» *Journal of Quality and Technology* , vol. 26, pp. 274-287, 1994.
- [51] C. P. Quesenberry, «SPC Q charts for Binomial parameter p: short or long runs,» *Journal of Quality Technology*, vol. 23, nº 3, pp. 239-246, 1991.
- [52] C. P. Quesenberry, «SPC Q charts for a Poisson parameter lambda: short or long runs,» *Journal of Quality Technology*, vol. 23, nº 4, pp. 296-303, 1191.
- [53] C. Quesenberry, «SPC Q charts for start-up processes and short or ong runs.,» *Journal of Quality Technology*, vol. 23, nº 3, pp. 213-224, 1991.
- [54] J. MacGregor y T. Harris, «The Exponentially Weighted Moving Variance,» *Journal of Quality and Technology*, vol. 25, nº 2, pp. 106-118, 1993.
- [55] T. Lai, «Sequential Change-Point Detection in Qaulity Control and Dynamical Systems,» *Journal of Royal Statistical Society*, vol. 57, nº 4, pp. 613-658, 1995.

- [56] I. Nikiforov, «A Generalized Change Point Detection Problem,» *IEEE Transactions on Information Theory*, vol. 41, nº 1, pp. 171-187, 1995.
- [57] E. Gombay y L. Horvarth, «Approximations for the Time of Change and the Power functions in Change Point Models,» *Journal of Statistical Planning and Inference*, vol. 52, pp. 43-66, 1996.
- [58] J. Sullivan y W. Woodall, «A Control Chart for Preliminary Analysis of Individual Observations,» *Journal of Quality and Technology*, vol. 28, nº 3, pp. 265-278, 1996.
- [59] D. a. O. D. Hawkins, *Cumulative sum charts and charting for quaiity improvement*, Springer, 1998.
- [60] C. Acosta-Mejia, «Monitoring reduction in variability with the range,» *IIE Transactions*, vol. 30, nº 6, pp. 515-523, 1998.
- [61] T. Samuel, J. Pignatiello y J. Calvin, «Identifying the Time of a Step Change in a Normal Process Variance,» *Quality Engineering*, vol. 10, nº 3, pp. 529-538, 1998b.
- [62] D. H. Timmer, J. Pignatiello y M. Longnecker, «The development and evaluation of CUSUM-based control charts for AR(1) process,» *IIE Transactions*, vol. 30, nº 6, pp. 525-534, 1998.
- [63] G. Nedumaran, J. Pignatello y J. A. Calvin, «Identifying the time of a step-change with χ^2 control charts,» *Quality Control and applied statistics*, vol. vo. 47, nº no. 2, pp. 125-126, 2002.
- [64] T. Lai, «Sequential Multiple Hypotesis Testing and Efficient Fault Detection-Isolation in Stochastic Systems,» *IEEE Transactions on Information Theory*, vol. 46, nº 2, pp. 595-608, 2000.
- [65] A. S. Dabye y Y. Kutoyants, «Misspecified Change-Point Estimaton Problem for a Poisson Process,» *Journal of Applied Probability*, vol. 38, pp. 122-130, 2001.
- [66] K. W. B. Kumar, «Detection of Change Points in Time Series Analysis with Fuzzy Statistics,» *International Journal of Systems Science*, vol. 32, nº 9, pp. 1185-1192, 2001.
- [67] J. Pignatiello y T. Samuel, «Estimation of the Change Point of a Normal Process mean in SPC Applications,» *Journal of Quality Technology*, vol. 33, pp. 82-95, 2001.
- [68] J. Sullivan y L. Jones, «A self-starting control chart for multivariate individual observations,» *Technometrics*, vol. 44, nº 1, pp. 24-33, 2002.
- [69] E. Gombay, «Sequential Change Point Detection and Estimation,» *Sequential Analysis: Design Methods and Applications*, vol. 22, nº 3, pp. 203-222, 2003.
- [70] J. Park y S. Park, «Estimation of the Change Point in X y S Control Charts,» *Communications in Sttistics - Simulation and Computation*, vol. 33, nº 4, pp. 1115-1132, 2004.

- [71] D. M. Hawkins y K. D. Zamba, «Statistical process control for shifts in mean or variance using a changepoint formulation,» *Technometrics*, vol. 47, nº 2, pp. 164-173, 2005.
- [72] A. Ramanayake, «Tests for a Change Point in the Shape Parameter of Gamma Random Variables,» *Communications in Statistics - Theory and Methods*, vol. 33, nº 4, pp. 821-833, 2005.
- [73] A. Costa y M. Machado, «A New Chart Based on Sample Variances for Monitoring the Covariance Matrix of Multivariate Processes,» *International Journal of Advanced Manufacturing Technology*, nº 41, pp. 770-779, 2009.
- [74] D. Hawkins y E. Maboidou-Tchao, «Self-starting multivariate exponentially weighted moving average control charting,» *Technometrics*, vol. 49, nº 2, pp. 199-209, 2007.
- [75] E. Schechtman, G. Bandner y S. Meginy, «Detecting a change in a scale parameter- a combination of SPC and change point procedures,» *International Journal of Production Research*, vol. 45, nº 23, pp. 5535-5545, 2007.
- [76] C. Zhou, C. Zou, Y. Zhang y Z. Wang, «Nonparametric control chart based on change-point model,» *Statistical Papers*, vol. 50, nº 1, pp. 13-28, 2009.
- [77] C. Zou, Y. Zhang y Z. Wang, «A control chart based on a change-point model for monitoring linear profiles,» *IIE Transactions*, vol. 38, nº 12, pp. 1093-1103, 2006.
- [78] B. Brodsky, «Sequential Detection and Estimation of Change-Points,» *Sequential Analysis: Design Methods and Applications*, vol. 29, nº 2, pp. 217-233, 2010.
- [79] J. Chen y A. Gupta, *Parametric Statistical Change Point Analysis*, 1era ed., New York: Birkhäuser Boston, 2010.
- [80] E. Maboudou-Tchao y D. Hawkins, «Self-Starting Multivariate Control Charts for Location and Scale,» *Journal of Quality Technology*, vol. 43, nº 2, pp. 113-126, 2011.
- [81] A. Cordero-Franco, V. Tercero-Gómez, J. Garza-Venegas, M. d. C. Temblador-Pérez y M. Beruvides, «Change Point Estimators for Unkown Parameters in Normally Distributed Series of Independent Observations,» *Quality and Reliability Engineering International*.
- [82] V. Tercero-Gómez, A. Cordero-Franco, I. Martínez-Salazar y J. Ramirez-Galindo, «Change-point Estimation and Control Chart for Variance Based on the P-Value Function of the F Statistics,» de *Proceedings of teh 2012 Industrial and Systems Engeineering Research Conference*, Orlando, Florida, 2012.
- [83] M. Esquivel-Fraustro, J. Guerrero y J. Macías-Díaz, «Activity pattern detection in electroneurographic and electromyogram signals through a heteroscedastic change-point method,» *Mathematical*

Biosciences, vol. 224, nº 2, pp. 109-117, 2010.

- [84] C. Pao-Chin y Z. Xin, «Bayesian change-point analysis of tropical cyclone activity: The central North Pacific case,» *Journal of Climate*, vol. 17, nº 24, pp. 4893-4901, 2004.
- [85] C. Zou, C. Zhou, Z. Wang y F. Tsung, «A Self-Starting Control Chart for Linear Profiles,» *Journal of Quality Technology*, nº 39, pp. 364-375, 2007.
- [86] J. Pignatiello y T. Samuel, «Identifying the Time of a Step Change in the Process Fraction Nonconforming,» *Quality Engineering*, nº 13, pp. 357-365, 2001.
- [87] K. Zamba y D. Hawkins, «A Multivariate Change-Point Model for Statistical Process Control,» *Technometrics*, nº 48, pp. 539-549, 2006.
- [88] A. Batsidis, «Robustness of the Likelihood Ratio Test for Detection and Estimation of a Mean Change Point in a Sequence of Elliptically Contoured Observations,» *Statistics: A Journal of Theoretical and Applied Statistics*, nº 44, pp. 17-24, 2010.
- [89] M. Reynolds y L. Jianying, «An Evaluation of a GLR Control Chart for Monitoring the Process Mean,» *Journal of Quality Technology*, nº 42, pp. 287-310, 2010.
- [90] D. Wallace, «Bounds on Normal Approximations to Student's and the Chi-Square Distributions,» *Annals of Mathematical Statistics*, nº 30, pp. 1121-1130, 1959.
- [91] D. Peizer y J. Pratt, «A Normal Approximation for Binomial, F, Beta, and Other Common, Related Tail Probabilities,» *Journal of the American Statistical Association*, nº 63, pp. 1416-1456, 1968.
- [92] J. Lucas y R. Crosier, «Robust Cusum: a Robustness Study for Cusum Quality Control Schemes,» *Communications in Statistics-Theory and Methods*, nº 11, pp. 26-69, 1982.
- [93] D. Hawkins, «Cumulative Sum Control Charting: an Underutilized SPC tool,» *Quality Engineering*, nº 5, pp. 463-477, 1993.
- [94] T. R. Samuel, J. J. Pignatiello y J. A. Calvin, «Identifying the time of a step change in Normal process variance,» *Quality Engineering*, vol. 10, nº no. 3, p. 529, 1998.
- [95] D. Hawkins y D. H. Olwell, *Statistics for Engineering and Physical Science*, New York: Springer, 1998.
- [96] D. Hawkins y D. Olwell, *Cumulative sum chart and charting for quality improvement*, Springer, 1998.
- [97] D. Hawkins, «On the distribution and power of a test for a single outlier,» *South African Statistical Journal*, vol. 3, pp. 9-15, 1969.

- [98] D. Hawkins, «Cumulative Sum Control Charting: An Underutilized SPC Tool,» *Quality Engineering*, vol. 5, nº 3, pp. 463-477, 1993.
- [99] T. Burgin, «The Gamma Distribution and Inventory Control,» *Operational Research Quarterly*, vol. 2, nº 142, pp. 507-525, 1975.
- [100] I. Feinberg, «Gamma Distribution Model Describes Maturation Curves for Delta Wave Amplitude, Cortical Metabolic Rate and Synaptic Density,» *Journal of Theoretical Biology*, vol. 2, nº 142, pp. 149-161, 1990.
- [101] J. Lucas y M. Saccucci, «Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements,» *Technometrics*, vol. 32, nº 1, pp. 1-12, 1990.
- [102] M. Cristinel, «New Approximations of the Gamma Function in Terms of the DiGamma Function,» *Applied Mathematics Letters*, vol. 1, nº 23, pp. 97-100, 2010.
- [103] E. Pitman, «The "closest" Estimates of Statistical Parameters,» *Mathematical Proceedings of the Cambridge Philosophical Society*, nº 33, pp. 212-222, 1937.
- [104] S. W. Roberts, «A comparison of some control chart procedures,» *Technometrics*, vol. vol. 8, nº no. 3, pp. 411-430, August 1966.
- [105] D. Timmer y J. Pignatiello, «Cange Point Estimates for the Parameters of an AR(1) Process,» *Quality and Realiability Engineering International*, nº 19, pp. 335-369, 2003.