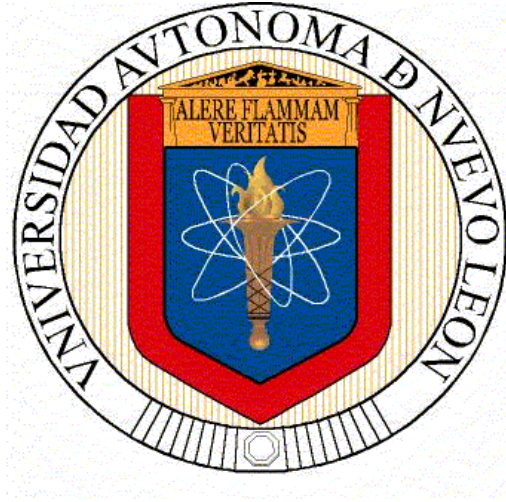


UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN.
FACULTAD DE CIENCIAS FÍSICO-MATEMÁTICAS.



Carta de Control No-Paramétrica de Punto de Cambio para Varianza.

Por

ELENA CRISTINA VILLANUEVA GUERRA

Como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS CON ORIENTACIÓN EN
MATEMÁTICAS.

DICIEMBRE 2013



UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN.
FACULTAD DE CIENCIAS FÍSICO-MATEMÁTICAS.



Carta de Control No-Paramétrica de Punto de Cambio para Varianza.

Por

ELENA CRISTINA VILLANUEVA GUERRA

Como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS CON ORIENTACIÓN EN MATEMÁTICAS.

Carta de Control No-Paramétrica de Punto de Cambio para Varianza.

Por

ELENA CRISTINA VILLANUEVA GUERRA

Como requisito parcial para obtener el grado:

MAESTRO EN CIENCIAS CON ORIENTACIÓN EN MATEMÁTICAS.

Miembros del comité:

Dr. Víctor Gustavo Tercero Gómez

Asesor

Dr. Alvaro Eduardo Cordero Franco

Co-asesor

Dr. José Guadalupe Ríos Alejandro.

Sinodal

Contenido

Lista de figuras	6
Lista de tablas	7
Resumen	8
CAPÍTULO 1. INTRODUCCIÓN	9
1.1. Introducción	9
1.2. Historia y Antecedentes	11
1.3. Declaración del problema	14
1.4. Preguntas de Investigación	14
1.5. Hipótesis de Investigación	14
1.6. Objetivo	15
1.7. Justificación	15
1.7.1. Justificación científica	15
1.7.2. Justificación práctica	16
1.8. Alcances y limitaciones	16
CAPÍTULO 2. MARCO TEÓRICO	18
2.1. Las cartas de control	18
2.1.1. Desempeño de una carta de control en términos del ARL	20
2.1.2. Implementación de cartas de control.	22

2.2.	Cartas de control paramétricas para Varianzas.....	22
2.2.1.	Carta de control R	22
2.2.2.	Cartas de control S	24
2.2.3.	Cartas de control S^2	25
2.2.4.	CUSUM para varianzas	26
2.2.5.	CUSUM Auto-inicial (SSCUSUM).....	27
2.2.6.	EWMA para varianzas.....	28
2.3.	Cartas No paramétricas para la varianza	29
2.3.1.	Carta de Control No-Paramétrica para controlar la variabilidad basada en la prueba de rangos al cuadrado.....	30
2.4.	Cartas de control basadas en el enfoque de punto de cambio	31
2.4.1.	Secuencia para validar la carta de control:	32
2.5.	Modelo de punto de cambio para cambios en varianza.....	33
2.6.	Prueba de igualdad de varianzas de Conover & Iman (1978).....	36
CAPÍTULO 3.	MODELO.....	40
3.1.	Propuesta.....	40
3.2.	Modelo.....	40
CAPÍTULO 4.	DESEMPEÑO	46
4.1.	Diseño de la experimentación Monte Carlo	46
4.2.	Desempeño de la carta de control propuesta	47
4.3.	Carta de control Auto-inicial CUSUM para varianzas.....	51
4.4.	Ejemplo Numérico.....	54
CAPÍTULO 5.	CONCLUSIONES	59

5.1. Conclusiones.....	59
5.2. Trabajo futuro.....	62
5.2.1. Carta de control de una cola	62
5.2.2. Usar diferentes estadísticos para el cálculo de los límites de control	62
5.2.3. Considerando cambios en media y varianza simultáneos.....	62
5.2.4. Analizar el desempeño de τ	62
5.2.5. Nota para los trabajos de Zhou et al. (2009) y Hawkins & Deng (2010), para implementarlos en monitoreo de varianzas.....	63
REFERENCIAS	64
CÓDIGOS DE MATLAB.....	67

Lista de figuras

Figura 1. Cartas de control tipo Shewhart: a) sin cambio, b) con cambio en media, c) con cambio en varianza y d) ambas.....	19
Figura 2. Datos comprendidos en el periodo de Julio 2004 a Julio 2009 del índice S&P 500.	54
Figura 3. Carta de control SSCUSUM para el índice bursátil S&P 500.	56
Figura 4. Desempeño de la carta propuesta con los diferentes niveles α	57

Lista de tablas

Tabla 1. Representación tabular de los posibles escenarios respecto a la H_0 . Errores tipo I y II.	21
Tabla 2. Desempeño en términos de ARL de la carta de	36
Tabla 3. Límites de control hn, α a partir de la muestra número 10. Los primeros 40 límites se calcularon mediante simulación y del 50 en adelante se usó la ecuación de regresión. .	45
Tabla 4. Desempeño en términos de ARL de la carta propuesta. Se marca con el símbolo	48
Tabla 5. Desempeño en términos de ARL de la carta propuesta para la distribución Gamma.	50
Tabla 6. Desempeño en términos de ARL del SSCUSUM para la distribución Normal.	52
Tabla 7. Desempeño en términos de ARL del SSCUSUM para la distribución Gamma.....	53
Tabla 8. Datos del S&P 500 fechados para la SSCUSUM	55
Tabla 9. Experimentación de la tabla de control propuesta.	58

Resumen

Una estrategia para mejorar la calidad es evitar el aumento de la variabilidad, para esto contamos con el Control estadístico de procesos, y una de sus principales herramientas son las Cartas de control (CC). Estas son utilizadas para el monitoreo y mejora de la producción y los servicios. En la llamada Fase I de las CC. Se desconoce la información referente a la distribución de probabilidad de las observaciones del proceso. Sin embargo, en ocasiones desde etapas tempranas se necesita dar seguimiento a los mismos. Para asistir en la solución de este problema, se propone una carta de control no paramétrica para monitorear la varianza que detecte cambios sostenidos en un proceso. Esta carta utiliza el enfoque del punto de cambio en conjunto con una prueba no paramétrica para varianzas basada en la transformación de rangos al cuadrado. Por la naturaleza no paramétrica del estimador no es necesario conocer ni los parámetros ni la distribución que estamos controlando. El desempeño de la carta se compara con su homólogo paramétrico basado en la distribución normal. En este trabajo se describe la construcción de la nueva carta de control, se muestran resultados experimentales sobre el desempeño, y se muestra un ejemplo práctico para ilustrar la aplicación del método en procesos reales.

CAPÍTULO 1. INTRODUCCIÓN

1.1. Introducción

Administración es predicción. Para administrar el rumbo de una organización, sistema, o proceso, es necesario ser capaces de predecir su comportamiento. Cuando este comportamiento es predecible, se dice que el proceso se encuentre dentro de control, y una de las herramientas más utilizadas para este fin son las cartas de control. A través de las cartas de control se evalúan ciertas características de calidad de procesos de productos o servicios. Entre estas características de calidad se encuentran medidas llamadas de localización y dispersión que han sido tratadas utilizando herramientas estadísticas paramétricas y no paramétricas. Las medidas paramétricas, tradicionalmente asumen normalidad, una situación que no siempre se puede sostener. Para solucionar este problema se desarrollaron estrategias no paramétricas, que no asumen o no requieren del conocimiento de la distribución de las observaciones. Sin embargo, las cartas de control del tipo no paramétricas para la dispersión, requieren de la utilización de “grandes” tamaños de muestra para lograr desempeños aceptables. Por otro lado, la aparición de una técnica de diseño de cartas de control basada en el análisis de puntos de cambio ha mostrado que es posible crear cartas no paramétricas con desempeños aceptables que compiten incluso con las opciones paramétricas. Así, tomando el análisis del problema de controlar la dispersión y una relativamente nueva técnica de diseño basado en puntos de cambio, en esta investigación se propone, desarrollo y evalúa una carta de control no paramétrica para la dispersión capaz de lidiar con observaciones individuales independientes entre sí.

Según Shewhart (1931) *“...un fenómeno se dice que está en control cuándo a través del uso de experiencia pasada podemos predecir por lo menos dentro de límites cómo el fenómeno se espera que varíe en un futuro...”* (p.6, traducción propia).

Para saber si un fenómeno está bajo control, es necesario medirlo. La medición se realiza a través de características especiales que pueden considerarse como variables o atributos.

Las primeras vienen de mediciones cuantitativas y las segundas de mediciones cualitativa. Los datos manejados en este trabajo son muestras de observaciones de estas variables.

Analizando estas variables, las series de datos reales presentan una variabilidad intrínseca del proceso que se puede presentar como error aleatorio debido a causas naturales de variación. En un proceso es realmente difícil encontrar dos productos exactamente iguales y más aún esperar que la línea de producción sea perfecta. También se puede dar el caso que la variabilidad del proceso sea debida a causas asignables de variación. Distinguir las causas de variabilidad ayuda a resolver el problema principal del Control estadístico de procesos (SPC) que es saber si existe una diferencia entre la distribución que debería seguir el proceso y la distribución real del proceso. El Dr. Shewhart (1931) creó las bases para las Cartas de control intentando resolver este problema y dando así inicio a lo que se conoce hoy como SPC. Las acciones de Shewhart inspiraron la creación de muchos tipos de cartas de control y otros métodos para estimar puntos de cambio en las distribuciones del proceso.

En un proceso siempre existirá variabilidad, pero ¿por qué es necesario el monitoreo de la variabilidad? Algunos autores definen la mejora de la calidad como la reducción de la variabilidad y ese es el objetivo último de la investigación.

En este trabajo desarrollamos una carta de control no paramétrica de punto de cambio para la varianza utilizando un enfoque conocido como de punto de cambio. En la literatura se pueden encontrar cartas de control no paramétricas para varianzas, así como cartas de control paramétricas basadas en el estadístico de punto de cambio, pero no existe una carta como la que se presenta en este documento cuya base de desarrollo es el análisis de puntos de cambio.

A lo largo del **CAPÍTULO 1** veremos el contexto del problema, preguntas de investigación, hipótesis y objetivos, entre otras cosas, para describir el problema a resolver en la investigación. En el **CAPÍTULO 2** se hace una revisión de literatura, investigaciones anteriores, metodologías relacionadas, definiciones esenciales, y demás datos de relevantes a la investigación. La propuesta, donde se desarrolla el modelo, los supuestos, y metodologías empleadas en detalle se muestran en el **CAPÍTULO 3**. También se muestra la forma de calcular los límites de control y la posible estimación del punto de cambio. El desempeño en términos de *ARL* (promedio de longitud de corrida, mejor definido en la sección **2.1.1**) de la carta de control propuesta se muestra en el **CAPÍTULO 4**, donde se compara con su contraparte paramétricas que utiliza la estrategia de sumatorias acumuladas, la carta de control CUSUM de auto inicio. Finalmente en el **CAPÍTULO 5** se hace una recapitulación del trabajo realizado y se concluye la investigación. Al final del documento, en la sección de Apéndices, se agregan los códigos en Matlab® utilizado para el desarrollo de la investigación

1.2. Historia y Antecedentes

A principios de la Segunda Guerra Mundial el Dr. Shewhart (1931) dio inicio a una metodología de medición y control de la calidad en procesos de manufactura, que actualmente conocemos como Control estadístico de procesos (Statistical Process Control SPC), con el desarrollo de las Cartas de control (CC también conocidas como CC tradicionales o CC tipo Shewhart).

En los procesos de manufactura y servicio, el cliente es una de las piezas clave, es quien adquiere (compra) el producto final. Para que el cliente siga adquiriendo el producto se debe mantener un estándar de calidad que “mida” las características que necesita para estar lo más próximo a lo que el cliente busca sin dejar de lado lo que la empresa necesita. Una de las medidas que se han utilizado para mantener estos estándares de calidad es la inspección, que busca encontrar productos defectuosos antes que puedan causar problemas. Sin embargo, en ocasiones la industria pierde más de lo que gana realizando

inspecciones prolongadas o destructivas. También, para atender esta problemática, existen herramientas estadísticas que nos permiten “evaluar” si un producto cumple o no cumple lo especificado en líneas de producción, se monitorean uno a uno los productos y detectan productos que no cumplen las especificaciones.

En el análisis de la línea de producción es dónde entran las Cartas de control, esta herramienta monitorea muestras para ayudar a analizar la distribución de los datos que arroja el proceso. Si el proceso siempre sigue la misma distribución, existe la posibilidad que la carta de control arroje una falsa alarma (error tipo I). Esto se determina buscando causas asignables de variación. Si no se encuentra, se toma como un punto atípico. Mientras que si por el contrario la distribución del proceso cambió en realidad y se recibe una alarma, se buscan las causas de variación y se toman acciones correctivas (de acuerdo al 3er postulado de control de Shewhart (1931) para así evitar seguir con la producción de este tipo de producto “defectuoso”. También existe la posibilidad de que no se detecte un cambio que ha ocurrido esto es un error tipo II.

Mantener en control el proceso de manufactura ayuda a que tanto el cliente como la empresa se sientan satisfechos. Por lo general no es rentable diseñar y crear productos en grandes cantidades para que al pasar por el control de calidad sean desechados y se elaboren nuevos productos que sí cumplan con las especificaciones. Esto se traduce en pérdidas, las cuales podrían causar un aumento del costo del producto y otras repercusiones negativas. Las cartas de control pueden servir en investigaciones médicas, biológicas, sociales, económicas, etc. dónde lo que nos interesa es saber el desempeño de un proceso.

Las cartas de control tradicionales asumen una distribución conocida y en la mayoría de los casos Normal. Sin embargo, en la práctica este requisito no siempre se puede sostener dado que el conocimiento de la distribución de los datos se tiene sólo para un número limitado de procesos. Así, si queremos tomar una decisión “inmediata”, la alternativa es estudiar los datos para obtener una distribución que se ajuste en buena medida a estos

datos, pero esto requiere una amplia recolección de información, lo que puede llegar a ser problemático y costoso. En los sistemas complejos, cuando se conoce la distribución que sigue el proceso, no suele ser de comportamiento Normal, y es aquí donde estas cartas paramétricas no son adecuadas. Se requiere conocimiento a priori de la distribución y en ocasiones hasta el conocimiento de los parámetros de los datos, lo que puede requerir de una Fase I de estimación muy larga para una buena estimación de estos.

1.3. Declaración del problema

Dada una serie de observaciones X_1, \dots, X_n de variables aleatorias independientes e idénticamente distribuidas, se desea diseñar una carta de control para varianzas sin asumir conocimiento previo sobre la distribución de los datos. Esto, siguiendo los lineamientos del enfoque de punto de cambio, donde el tipo de cambio a detectar es un cambio sostenido. Para ello se utilizaría un estadístico basado en una prueba de igualdad de varianzas.

1.4. Preguntas de Investigación

- Pregunta 1. ¿Es posible encontrar una función matemática que ayude a predecir los límites de control como los mostrados en Hawkins y Zamba (2005)?
- Pregunta 2. ¿La carta de control propuesta tiene una potencia (medida en términos de **ARL**) similar para detectar cambios en la varianza que su equivalente paramétrico cuando se cumplen los supuestos de normalidad?
- Pregunta 3. ¿La carta de control propuesta tiene mejor potencia (en términos de **ARL**) para detectar cambios en la varianza en el caso donde su equivalente paramétrico no es apropiado (cuando la distribución no es normal)?

Más información sobre el proceso de evaluación basado en **ARL** se encuentra en la página 20 sección 2.1.1. Desempeño de una carta de control en términos del **ARL**.

1.5. Hipótesis de Investigación

- Hipótesis 1. El modelo propuesto por Hawkins & Zamba (2005) se puede utilizar para crear una carta de control no paramétrica donde se ajuste una función que prediga los límites de control para distintos niveles de Error Tipo I.
- Hipótesis 2. La carta de control propuesta tiene un desempeño similar al de una carta de control paramétrica que asume normalidad.

- Hipótesis 3. La carta de control propuesta tiene un buen desempeño con los procesos que no cumplen con los supuestos de normalidad en comparación con las actuales cartas de control que si asumen normalidad.

1.6. Objetivo

- Objetivo 1. Usar la metodología de Hawkins & Zamba (2005) con un estadístico No-Paramétrico para crear una carta de control para varianzas y encontrar límites de control correspondientes. Similar al creado por Hawkins & Deng, 2010 para medias.
- Objetivo 2. Evaluar el desempeño de la carta de control propuesta y su equivalente paramétrico que asume normalidad cuando las observaciones son de hecho normales.
- Objetivo 3. Evaluar el desempeño de la carta de control propuesta y su equivalente paramétrico que asume normalidad cuando las observaciones no son normales.

1.7. Justificación

1.7.1. Justificación científica

Después de revisar la literatura correspondiente al tema se identificó que no existe desarrollada ninguna metodología ni aplicación para este tipo de problema, nuestra adaptación es la primera.

Existen en la actualidad cartas de control que detectan cuándo una serie de datos está fuera de control, las estimaciones que aquí se encuentran distan del verdadero punto de cambio. También existen cartas de control de punto de cambio que necesitan de muchos supuestos para funcionar, la carta de control de punto de cambio propuesta es holgada con las suposiciones por su naturaleza no paramétrica y permite estimar el momento en que inicia el cambio al mismo tiempo que lo detecta.

1.7.2. Justificación práctica

En específico, tenemos el del desgaste de herramientas. El desgaste se da entre los materiales y las herramientas que interactúan en el proceso de desarrollo. Se busca encontrar límites de control para detectar estos desgastes en líneas de producción antes que el producto final resulte defectuoso. El desgaste de las herramientas se monitorea a través de cartas de control para varianzas.

Otro problema específico es el de líneas de producción con más de un producto que se fabrican secuencialmente. Aquí se necesita tener monitoreada solamente la varianza de los datos, que se procura sea constante.

Montgomery (2008), *“La calidad es inversamente proporcional a la variabilidad”* (p.6 traducción propia). Esta definición nos dice el porqué es útil controlar la variabilidad.

1.8. Alcances y limitaciones

El origen no paramétrico del estadístico usado permite eliminar el supuesto de normalidad y el conocimiento de los parámetros. La base teórica del estadístico es fuertemente sustentada con artículos de investigación.

Como alcance tenemos que puede detectar cambios de la varianza en los procesos con distribución Normal tanto como con distribuciones con muchos puntos atípicos como la distribución Gamma.

Esta carta de control está limitada a usar como mínimo 10 datos estables iniciales usados como referencia a partir de los cuales se determina si el proceso está bajo control. Esto podría ser una ventaja al tomar en cuenta que para algunos procesos es necesario el uso de grandes cantidades de datos, pero aunque sea menor la necesidad se siguen ocupando y eso podría generar algún costo.

La carta de control está limitada a detectar cambios sostenidos en series con observaciones independientes e idénticamente distribuidas. No se detectan cambios puntuales, y el comportamiento en observaciones auto-correlacionadas es incierto.

La carta de control propuesta requiere de la evaluación de las particiones cuya complejidad numérica crece de manera no polinomial, por lo que cuando los datos en el proceso son demasiados, el análisis se puede volver lento en ejecución.

Se asumen medias iguales, por lo que si existe un cambio en la media, se tiene que tener el conocimiento del parámetro de cambio para hacer una transformación antes de aplicar la carta de control. Además se asume que sólo existe un cambio en la varianza. Múltiples cambios podrían ser estimados uno a la vez.

CAPÍTULO 2. MARCO TEÓRICO

Haciendo una revisión de literatura nos damos cuenta que existe muy poca investigación en trabajos relacionados con el estudio de la variabilidad, ya que se le da más importancia al monitoreo de la media. Se observa que hasta que se tienen trabajos con la media se extienden trabajos para controlar la varianza. En este capítulo se presentarán conceptos, definiciones y desarrollos claves para el análisis de la varianza, exponiendo las primeras metodologías para el monitoreo de la varianza, clasificándolas por paramétricos y no paramétricos, seguido de nuevas investigaciones, sobre todo en trabajos no paramétricos. Se detallará la metodología usada para la creación de la carta de control. Por último, se detalla un poco de historia y desarrollos con el estadístico de prueba seleccionado para el diseño de la carta de control propuesta.

2.1. Las cartas de control

Los requerimientos del cliente son parte fundamental en los procesos de producción y servicios, y para que los productos cumplan con dichos requerimientos, el proceso de producción debe tener poca variación entre las dimensiones o características esenciales del producto, es decir la producción sea estable. El SPC provee herramientas para monitorear, detección y mejora de la calidad. La herramienta que trabajamos en este documento son las cartas de control iniciadas por Shewhart (1931), esta técnica es la más compleja del SPC.

En todo proceso siempre existirá variabilidad en las características (esenciales) de los productos terminados, esta variabilidad comúnmente llamada “ruido de fondo” es a causa de un gran número de pequeñas causas que hacen variar de manera natural al proceso.

Se define a un proceso **bajo control estadístico** cuándo:

- Este proceso cuente sólo con su variabilidad natural o causas naturales de variación
Montgomery (2008)

Se dice que un proceso está fuera de control cuándo a un proceso le afectan causas asignables (no naturales) de variación (por ejemplo: errores humanos de operación, errores de medición, materia prima defectuosa, etc.). En dado caso es difícil predecir el comportamiento de la línea de producción y generalmente sale de los límites de control natural del proceso, esto debido a que la variabilidad debida por causas naturales y la variabilidad debida a causas asignables no es comparable y es inaceptable en la mayoría de los casos. La siguiente figura muestra cuatro cartas de control tipo Shewhart en la primera no se registra ningún cambio a), en la segunda se presenta un cambio en la media del proceso b), en la tercer imagen se distingue un cambio en la varianza c), y para el último caso se aprecia un cambio en ambos parámetros de la distribución d), todos los cambios localizados exactamente a la mitad del proceso.

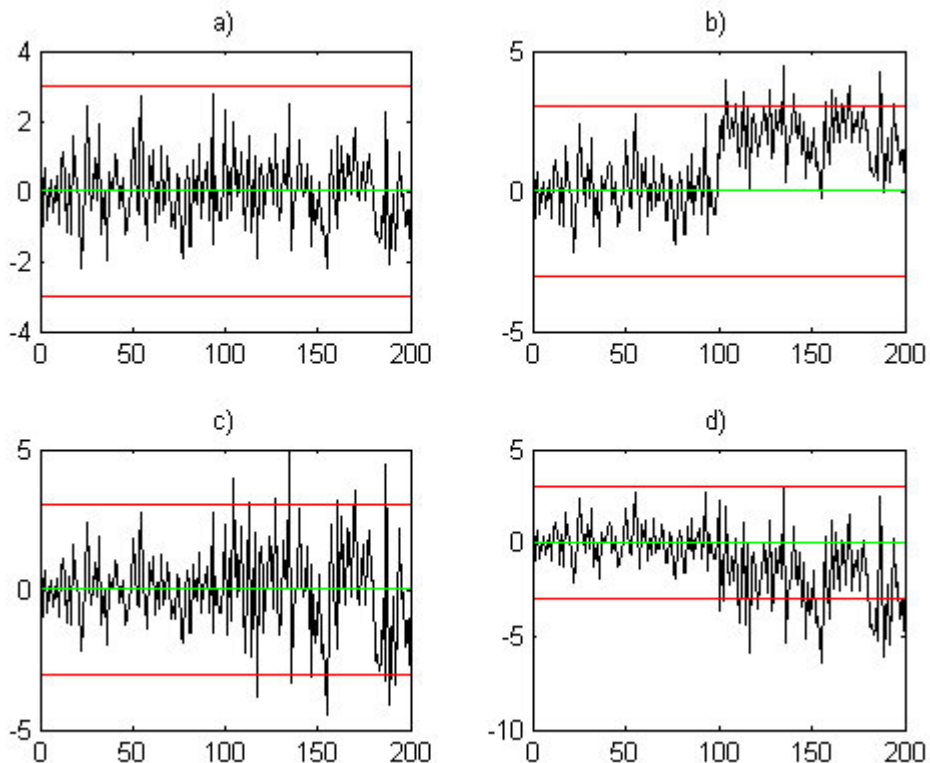


Figura 1. Cartas de control tipo Shewhart: a) sin cambio, b) con cambio en media, c) con cambio en varianza y d) ambas.

Las cartas tipo Shewhart cuentan con 3 características esenciales, una línea central que es la media (del proceso), y los límites de control inferior y superior (del proceso) que distan de tres veces la desviación estándar en control (del proceso) hacia arriba y hacia abajo respectivamente. Gráficamente se detecta un punto fuera de control cuando este sobrepasa los límites de control.

Por otra parte existen los límites de especificación del cliente, estos no están relacionados de forma matemática a los límites de control del proceso.

2.1.1. Desempeño de una carta de control en términos del ARL.

Una carta de control usualmente mide su desempeño en términos de su promedio de longitud de corrida o **ARL** (por sus siglas en inglés), que es el promedio de muestras recolectados antes de la primera señal de cambio. La longitud de la corrida **RL** (por sus siglas en inglés) es el tiempo que tarda en ocurrir una señal fuera de control esté o no fuera de control. Es fácil ver que el RL sigue una distribución Geométrica, al ser una serie de N ensayos repetidos con la misma probabilidad de que cada ensayo se obtenga un resultado fuera de los límites de control, asumiendo independencia, $RL \sim g(x; \theta) = \theta(1 - \theta)^{x-1}$, y la media de esta distribución es $\mu = \frac{1}{\theta}$. Por lo tanto en teoría el ARL en control es $ARL_0 = \frac{1}{\alpha}$, siendo α (error tipo I) la probabilidad de éxito o en este caso de una señal fuera de control, siendo que no se encuentra fuera de control.

Para el cálculo del ARL_0 es necesario entender los tipos de errores o falsas alarmas que se pueden presentar. A continuación en la Tabla 1 se ejemplifica. Se calcula el promedio de los momentos en lo que ocurre la primera señal fuera de control sin contar las falsas alarmas a este desempeño se le conoce como ARL fuera de control $ARL_1 = \frac{1}{1-\beta}$, y siendo β (error

tipo II) la probabilidad de una señal aceptar la hipótesis (no encontrar señal de alarma), siendo que no se encuentra en control.

Las falsas alarmas ocurren cuando la carta de control da una señal de cambio cuándo este no ha ocurrido, es decir un ARL de gran tamaño. Por otra parte cuando hay un cambio en el proceso y este está fuera de control una buena carta debería detectarlo lo más pronto posible es decir tener un ARL pequeño.

Tabla 1. Representación tabular de los posibles escenarios respecto a la H_0 . Errores tipo I y II.

		Realidad	
		H_0	Falsa
Decisión	Aceptar	✓ $(1 - \alpha)$	Error tipo II (β)
	Rechazamos	Error tipo I (α)	✓ $(1 - \beta)$

Decisión correcta: aceptar H_0 dado que es cierta la H_0 :

$$P(\text{Aceptar } H_0 | H_0 \text{ es cierta}) = 1 - \alpha \quad (1)$$

Decisión incorrecta: rechazar H_1 dado que la H_0 es cierta (Error tipo I):

$$P(\text{Rechazar } H_0 | H_0 \text{ es cierta}) = \alpha \quad (2)$$

Decisión incorrecta: aceptar H_0 dado que es cierta la H_1 (Error tipo II):

$$P(\text{Aceptar } H_0 | H_1 \text{ es cierta}) = \beta \quad (3)$$

Decisión correcta: rechazar H_0 dado que es cierta la H_1 :

$$P(\text{Rechazar } H_0 | H_1 \text{ es cierta}) = 1 - \beta \quad (4)$$

2.1.2. Implementación de cartas de control.

La manera tradicional de implementarlas es mediante la metodología de la fase I y fase II. La fase I es para la estimación de los parámetros, la cual es costosa en tiempo y en términos monetarios. La fase II es para el monitoreo de las nuevas observaciones, aquí es donde se implementan los límites de control para determinar si la serie está o no bajo control.

2.2. Cartas de control paramétricas para Varianzas

En el caso específico de las cartas de control para varianza, desde de un punto de vista paramétrico se ha trabajado en:

2.2.1. Carta de control R

La variabilidad del proceso puede monitorearse con una carta de control para el rango, esta carta es llamada carta R por lo general se usa en paralelo con una carta \bar{x} . Para el desarrollo de esta carta es necesaria una estimación de la desviación estándar σ , estimación que se puede realizar por medio de rangos de las m muestras, y cada muestra de tamaño n , el rango es la diferencia entre el dato más grande y el más chico de cada muestra:

$$R_j = x_{(n),j} - x_{(1),j}; j = 1:m \quad (5)$$

Donde $x_{(n)}$ es el estadístico de orden $n \forall i$, [o sea es el más grande y $x_{(1)}$ es el más chico.]

$$\bar{R} = \frac{R_1 + R_2 + \dots + R_m}{m} \quad (6)$$

Límites de control:

$$\begin{aligned} UCL &= D_4 \bar{R} \\ LC &= \bar{R} \\ LCL &= D_3 \bar{R} \end{aligned} \quad (7)$$

Los datos faltantes se mostrarán en los apéndices.

Estos límites de control nos sirven para verificar si las primeras m muestras (**Fase I**) están bajo control, si es así se siguen utilizando los límites actuales, si resulta estar fuera de control se deben buscar los datos atípicos y encontrar una causa asignable para ellos si no se encuentran causas asignables se descartan los puntos atípicos y se recalculan los límites de control y se repite el proceso. Teniendo la Fase I bajo control se siguen monitoreando los restantes valores a esto se le conoce como la **Fase II**.

Los patrones no aleatorios de las cartas son una señal de alarma. Al aplicar en paralelo ambas cartas \bar{x} y R y si ambas cartas cuentan con este patrón, se debe tomar en cuenta primero la carta R . Se buscan causas asignables para que se mejore el comportamiento de los datos en esta carta y así automáticamente se arreglará el patrón de la carta \bar{x} . Nunca debemos comenzar a analizar la carta \bar{x} .

En el caso donde se dude de la normalidad de los datos es necesario conocer la distribución de los datos para poder derivar distribuciones de muestreo de \bar{x} y R . Si al contrario no se conoce la distribución de los datos la opción es tratar a los datos como normales gracias a la utilización del teorema del límite central, al menos en el caso de la carta de medias.

2.2.2. Cartas de control S

Por otra parte, ocasionalmente lo que se busca es monitorear la desviación estándar del proceso directamente en lugar de hacer estimaciones con el rango. En este caso usaremos la carta S , estimaremos la desviación estándar del proceso y la denotaremos como S . Esta carta también se usará en conjunto con la carta \bar{x} . Esta carta es más recomendable que las anteriores cuando:

- El tamaño de la muestra (n) es poco más grande de 10 o 12.
- n es variable.
- Para el uso de esta carta es casi el mismo que la anterior con diferencias en la prueba \bar{x} que no serán vistos en este documento. Para mayor referencia ver Montgomery (2008).

$$UCL = B_6\sigma$$

$$LC = c_4\sigma \quad (8)$$

$$LCL = B_5\sigma$$

En caso de conocer la real desviación estándar, los datos faltantes se mostrarán en los apéndices.

Por otra parte:

$$UCL = B_4\bar{S}$$

$$LC = \bar{S} \quad (9)$$

$$LCL = B_3\bar{S}$$

En caso de conocer la real desviación estándar, los datos faltantes se mostrarán en los apéndices.

Cuando el tamaño de las muestras es variable se usan el enfoque de promedio ponderado para calcular tanto $\bar{\bar{x}}$ como \bar{S} , sea n_i el número de observaciones para la muestra i , se sigue que:

$$\bar{S} = \left[\frac{\sum_{i=1}^m (n_i - 1) S_i^2}{\sum_{i=1}^m n_i - m} \right] \quad (10)$$

Se usarán los mismos límites de control mostrados recientemente.

En la actualidad implementar este tipo de cartas S de control no resulta muy difícil, y esta resulta ser más usada cuando los tamaños de moderados a grandes ($n \gg 10$) y aun así por simplicidad en a la hora de hacer los cálculos se suele usarse la carta R . El desempeño de estas cartas...

2.2.3. Cartas de control S^2

Se presenta otra carta de control para el monitoreo directamente de la varianza muestral

$$\begin{aligned} UCL &= \frac{\bar{S}^2}{n-1} \chi_{\alpha/2, n-1}^2 \\ LC &= \bar{S}^2 \\ LCL &= \frac{\bar{S}^2}{n-1} \chi_{1-(\alpha/2), n-1}^2 \end{aligned} \quad (11)$$

Donde $\chi_{\alpha/2, n-1}^2$ y $\chi_{1-(\alpha/2), n-1}^2$ denotan los puntos porcentuales $\alpha/2$ superior e inferior de la distribución ji-cuadrada con $n - 1$ grados de libertad, y \bar{S}^2 es la varianza muestral promedio obtenida del análisis de los datos preliminares. Podría usarse un valor estándar de σ^2 en lugar de \bar{S}^2 en los límites anteriores.

Las cartas de control hasta aquí mencionadas se catalogan como cartas de control tipo Shewhart ya que estas cuentan con las características de las cartas tradicionales y además que evalúan la última observación del proceso, sin tomar en cuenta las anteriores. Las siguientes cartas por lo contrario usan las observaciones anteriores, esto con la intención de ganar terreno en los aspectos donde estas cartas son pobres en desempeño.

2.2.4. CUSUM para varianzas

La carta de control de suma acumulada es una carta de control para detectar pequeños corrimientos en el parámetro de localización, incorpora toda la información contenida en la secuencia de los valores muestrales graficando las sumas acumuladas C_i de las desviaciones que presentan los valores muestrales \dots, x_i, \dots respecto al valor objetivo \bar{x}_j .

$$C_i = \sum_{j=1}^i (\bar{x}_j - \mu_0) \quad (12)$$

Y para monitorear la variabilidad (Hawkins, 1981) sugiere crear una nueva cantidad para monitorear la variabilidad.

$$v_i = \frac{\sqrt{|y_i|} - 0.822}{.349} \quad (13)$$

Dónde y_i es el valor estandarizado de x_i :

$$y_i = \frac{(x_i - \mu)}{\sigma} \quad (14)$$

Hasta aquí describimos la CUSUM para varianzas, pero ahora para evitarnos el conocimiento a priori de los parámetros del proceso usaremos el Self-Starting CUSUM usaremos la transformación para varianzas v_i como datos del SSCUSUM.

2.2.5. CUSUM Auto-inicial (SSCUSUM)

Las cartas CUSUM son comúnmente usadas como monitoreo cuando ya se tienen una buena estimación de los parámetros iniciales, esto se garantiza con una Fase I larga. La Fase II, cuando no se cumple la suposición de una buena estimación de los parámetros, puede resultar contraproducente para la carta. Eso sucede mucho en los nuevos procesos, Hawkins (1987) propone una nueva técnica de monitoreo que no necesita la estimación de los parámetros para comenzar con el monitoreo, pero sigue con la necesidad del supuesto de la distribución Normal.

Con los datos transformados de la metodología anterior ahora usaremos el desarrollo de Hawkins (1987). Sea \bar{v}_n la media de los primeros n observaciones y sea $w_n = \sum_{i=1}^n (v_i - \bar{v}_n)^2$ la suma de desviaciones de las medias a las observaciones cuadradas. Por conveniencia de cálculos se formularon nuevas ecuaciones.

Para la media:

$$\bar{v}_n = \bar{v}_{n-1} + \frac{v_n - \bar{v}_{n-1}}{n} \quad (15)$$

Para las desviaciones:

$$w_n = w_{n-1} + \frac{(n-1)(v_n - \bar{v}_{n-1})^2}{n} \quad (16)$$

La varianza muestral de las primeras n observaciones es $s_n^2 = \frac{w_n}{n-1}$. Y la forma estandarizada de cada nueva observación se obtiene usando:

$$T_n = \frac{v_n - \bar{v}_{n-1}}{s_{n-1}}; \quad 3 \leq n \quad (17)$$

Si las observaciones tienen distribución Normal la distribución de $\sqrt{\frac{n-1}{n}}T_n$ tiene distribución t con $n - 1$ grados de libertad. La distribución acumulada de T_n es:

$$P(T_n \leq t) = F_{n-2} \left(t \sqrt{\frac{n-1}{n}} \right) \quad (18)$$

Donde F_n es la distribución acumulada t con $n - 1$ grados de libertad.

Resulta que si el área de la cola para cualquier variable aleatoria continua se convierte en un eje de ordenadas normal, se obtiene una nueva variable aleatoria que se distribuye exactamente como una variable aleatoria normal estándar. Es decir, si Φ^{-1} es la distribución Normal acumulada inversa, entonces la transformación hacia la normal se realiza de la siguiente manera:

$$U_n = \Phi^{-1}[F_{n-2}(a_n T_n)]; a_n = \sqrt{\frac{n-1}{n}} \quad (19)$$

Si los datos comienzan a mostrar alguna tendencia ya sea positiva o negativa la suma crecerá muy rápido. Y los límites están dados en magnitud del cambio.

$$\begin{aligned} S_i^+ &= \text{máx}[0, v_i - k + S_{i-1}^+] \\ S_i^- &= \text{máx}[0, -v_i - k + S_{i-1}^-] \end{aligned} \quad (20)$$

Donde k suele ser 0.5, y los límites de control $H = h\sigma$, donde los valores más usados de h son 4 o 5, dependiendo del ARL deseado.

2.2.6. EWMA para varianzas

Suponga que x_i muestras aleatorias independientes e idénticamente distribuida con distribución Normal con media μ y desviación estándar σ . El error cuadrático medio

ponderado exponencialmente (EWMS) discutido por (MacGregor & Harris, 1993) se define como:

$$S_i^2 = \lambda(x_i - \mu)^2 + (1 - \lambda)S_{i-1}^2 \quad (21)$$

Para valores grandes de i se puede demostrar que $E(S_i^2) = \sigma^2$ y si las observaciones son independientes la estimación tiene una distribución ji-cuadrada aproximada con $\nu = \frac{2-\lambda}{\lambda}$ grados de libertad. Y sus límites son:

$$UCL = \sigma_0 \sqrt{\frac{\chi_{\nu, \alpha/2}^2}{\nu}} \quad (22)$$

$$LCL = \sigma_0 \sqrt{\frac{\chi_{\nu, 1-(\alpha/2)}^2}{\nu}}$$

Para más información de estas cartas de control favor de revisar Montgomery (2008).

2.3. Cartas No paramétricas para la varianza

El problema de asumir una distribución específica es que los datos tiene que cumplir con el supuesto, de lo contrario no se asegura el desempeño, y en casos reales es muy difícil saber la distribución de los datos. Además si se conoce la distribución, encontrar los estimadores tal vez no sea una tarea fácil. Una forma de solucionar este problema es utilizar métodos no paramétricos

Lehmann (1975) usó pruebas no paramétricas como estadísticos de cartas de control en pruebas de igualdad de varianzas. Cuando el proceso se considera bajo control es necesario tener una mínima cantidad de datos m iniciales para comparar varianzas de muestras.

Bradley (1968) dio otro enfoque basado en la prueba de rango intercuartil de dos muestras.

Das & Bhattacharya (2008) propusieron una carta de control no paramétrica para monitorear la variabilidad basada la suma de rangos al cuadrado de Conover & Iman.

Das (2008a) propuso una carta de control no paramétrica para monitorear la variabilidad basada en dos sumas de rangos en dispersión por Ansari & Bradley (1960). Esta carta tiene un desempeño en ARL_0 muy bueno, pero el desempeño fuera de control no es mejor al propuesto por Das (2008b).

Das (2008b) propuso una carta de control no paramétrica para monitorear la variabilidad basada la suma de rangos al cuadrado de Conover & Iman, (esta prueba resulta ser mejor que su anterior carta con el estadístico de Ansari & Bradley) misma que utilizaremos en nuestra carta de control de punto de cambio y se explica en más adelante en la sección 2.6, página 36.

2.3.1. Carta de Control No-Paramétrica para controlar la variabilidad basada en la prueba de rangos al cuadrado

Das (2008b) desarrolló una carta de control para la varianza basada en el estadístico de Conover & Iman (1978) que se detalla más adelante. Los pasos para la construcción de la carta son:

1. Recolectar k (al menos 20) muestras de tamaño n .
2. Para cada 2 muestras seguidas calcular T . De esta manera $k - 1$ valores de T serán calculados.
3. Encontrar $E(T)$ y $Var(T)$
4. Calcular el valor estandarizado de T or Z para cada par de muestras
5. Establezca los parámetros de la gráfica de control como:

$$UCL = 3$$

$$LC = 0 \quad (23)$$

$$LCL = -3$$

Podemos notar que son los mismos que se usarían en una carta tipo Shewhart con datos en control que siguen la distribución Normal estándar.

6. Grafica los valores Z en la carta de control. Si algún punto va más allá del límite de control el proceso está fuera de control con respecto a la variabilidad.

Das (2008b) concluye que:

“La carta de control no paramétrico propuesto para la varianza muestra que su rendimiento es mejor que la carta de control S paramétrica cuando el ARL en control es considerado para todas las distribuciones” (Traducción propia).

Como se puede apreciar en la literatura revisada en la siguiente sección. En años recientes ha ido destacando el enfoque de punto de cambio para la creación de cartas de control. Este enfoque se detalla en la siguiente subsección.

2.4. Cartas de control basadas en el enfoque de punto de cambio

Este enfoque amplía el método de razón de verosimilitud generalizada para incluir la utilización de pruebas de hipótesis en el desarrollo de estadísticos para ser utilizados en cartas de control. Este relativamente nuevo procedimiento paramétrico fue desarrollado por Hawkins et al. (2003) quien propuso una carta de control para la detección de los cambios en la media cuando los parámetros del proceso son desconocidos, dos años más tarde Hawkins & Zamba (2005) desarrollaron un proceso similar pero ahora para cambios en la varianza (que describiremos más adelante) y otro para cambios en media y varianza, estos tres procedimientos bajo los supuestos de normalidad, Zhou, Zou, Zhang, & Wang,

(2009) desarrollaron por primera vez una carta de control no paramétrica basada en el enfoque de punto de cambio de Hawkins et al. (2003), usando el estadístico de Mann-Whitney y agregando un EWMA para detectar cambios en la media. Hawkins & Deng (2010) usaron un estadístico no paramétrico basado en el estadístico de Mann-Whitney para comparar los resultados de Zhou et al. (2009), en este último trabajo se concluye que es mejor el uso de Hawkins & Deng (2010) en la mayoría de los casos. Ross (2013) hace una mayor revisión de literatura actual y usa ayuda del lenguaje de programación R para una rápida implementación.

Enseguida se mostrará simbólicamente la metodología usada en las anteriores investigaciones. La diferencia en cada una de estas investigaciones es el estadístico de prueba.

El límite de control para la i – esima observación con determinado valor α de error es $h_{i,\alpha}$

Donde $h_{i,\alpha}$ se define resolviendo:

$$P(T_{\max,i} \geq h_{i,\alpha} | T_{\max,j} < h_{j,\alpha}, \forall j < i) = \alpha \quad (24)$$

2.4.1. Secuencia para validar la carta de control:

Donde T_i es el estadístico de la prueba de igualdad de varianzas de Conover & Iman (1978) para cada una de las particiones que van de $2 \leq i < n - 2$, y $T_{\max,n} = \max_{2 \leq i < n-2} |T_i|$ máximo de entre todas las particiones hasta la observación N .

$$T_i = \sqrt{\frac{5}{(N+1)(2N+1)(8N+11)}} \left[\frac{6T - \tau(N+1)(2N+1)}{\sqrt{\tau(N-\tau)}} \right] \quad (25)$$

- Calcular $T_{\max,n}$ el cuál es el estadístico T máximo de entre todas las particiones hasta la observación N .

- Si $T_{\max,n} \leq h_{i,\alpha}$ concluimos que el proceso está en control y regresamos al paso anterior con $n = n + 1$, en caso contrario ir al siguiente paso.
- $T_{\max,n} > h_{i,\alpha}$ entonces concluimos que ha ocurrido un cambio en la distribución del proceso, específicamente en la varianza. Se detiene la línea de producción para diagnóstico. Se estima el tiempo que ocurrió el cambio.

Lo que quiere decir, se calculará un estadístico por cada observación extra a las estables, si este estadístico resulta estar bajo los límites de control, agregamos una nueva observación y repetimos el proceso, de lo contrario si resulta estar fuera de control se manda una señal de alerta y se detiene el proceso para buscar causas asignables y tomar acciones correctivas.

2.5. Modelo de punto de cambio para cambios en varianza.

Hawkins y Zamba (2005) extendieron el trabajo de Hawkins et al. (2003) para desarrollar una carta de control para varianza con la nueva metodología de punto de cambio, y basado en la prueba de razón de verosimilitudes generalizada (GLR) para cambios en varianza con la corrección de Barlett adaptado para su uso secuencial.

La formulación se da en términos de punto de cambio

$$X_i \sim \begin{cases} N(\mu_1, \sigma_1); & 1 \leq i \leq \tau \\ N(\mu_2, \sigma_2); & \tau + 1 \leq i \end{cases} \quad (26)$$

Donde las X_i siguen la distribución Normal con media μ_1 y desviación estándar σ_1 de las lecturas de proceso antes del cambio, media μ_2 y desviación estándar σ_2 en su etapa fuera de control, τ es el momento en que ocurrió el cambio. Se dice que ha ocurrido un cambio en la media si $\mu_1 \neq \mu_2$, cambio en la varianza si $\sigma_1 \neq \sigma_2$, o en ambas si ocurren simultáneamente.

Suponiendo que han ocurrido n lecturas de proceso $0 \leq i < k \leq n$, se definen los siguientes estadísticos:

$$\bar{x}_{i,k} = \sum_{j=1+i}^k \frac{x_j}{k-1} \quad (27)$$

$$V_{i,k} = \sum_{j=1+i}^k (x_j - \bar{x}_{i,k})^2 \quad (28)$$

Suponiendo que se conoce el momento del cambio $\tau = k$. Se estiman los siguientes parámetros.

$$\begin{aligned} \hat{\mu}_1 &= \bar{x}_{0,k} \\ \hat{\mu}_2 &= \bar{x}_{k,n} \\ \hat{\sigma}_1^2 &= V_{0,k}/(k-1) \\ \hat{\sigma}_2^2 &= V_{k,n}/(n-k-1) \end{aligned} \quad (29)$$

Para el uso de la varianza constante, la varianza común se estima de esta manera:

$$\hat{\sigma}^2 = (V_{0,k} + V_{k,n})/(n-2) \quad (30)$$

La varianza estimada tiene los grados de libertad usados para hacerlo insesgado. La prueba de razón de verosimilitud generalizada para cambios en varianza, dónde $G_{k,n}$ se escribe en función de familias de razón F para probar igualdad de varianzas: si $F = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2}$ entonces

$$\begin{aligned} G_{k,n} = C^{-1} \{ & (k-1)\ln[k-1 + (n-k-1)F] \\ & + (n-k-1)\ln[(k-1)F^{-1} + n-k-1] - (n-2)\ln(n-2) \} \quad (31) \end{aligned}$$

Donde C es el factor de corrección de Barlett:

$$C = 1 + [(k - 1)^{-1} + (n - k - 1)^{-1} - (n - 2)^{-1}] / 3 \quad (32)$$

Por otra parte para la situación de un posible punto de cambio en la varianza desconocido, el GLR se da evaluando para todos los posibles valores de k . Tomando en cuenta que k no puede tomar algunos valores para asegurar la existencia del GLR, el rango de k es: $2 \leq k \leq n - 2$.

El estadístico del GLR para cambios en varianza es:

$$G_{max,n} = \max_{2 \leq k \leq n-2} G_{k,n} \quad (33)$$

El cálculo de los límites de control se realiza como lo explicamos en la sección anterior 2.4 haciendo cumplir la ecuación 32. Hawkins y Zamba (2005) usaron tablas para los $10 \leq n \leq 15$ y para $n > 15$ la siguiente ecuación de regresión.

$$h_{n,\alpha} = \begin{cases} -1.38 - 2.24\ln(\alpha) + \frac{1.61 + .691\ln(\alpha)}{\sqrt{n-9}} & .001 \leq \alpha < .05 \\ 5 + .066\ln(n-9) & \alpha = .05 \end{cases} \quad (34)$$

El desempeño del modelo de punto de cambio para cambios en varianza de Hawkins & Zamba (2005) se muestra en la Tabla 2. Usando en cada experimento se usó un $\alpha = .002$ el cuál corresponde a un $ARL_0 = 500$, teórico en la siguiente tabla se muestran los resultados del ARL de la carta de Hawkins & Zamba (2005), es fácil ver que el ARL_0 simulado no es parecido al teórico. Aunque, para efectos prácticos, sigue siendo aceptable.

Tabla 2. Desempeño en términos de **ARL** de la carta de ...

Teórico en Control	$\tau+1$	σ_2/σ_1 (Normal)						
		0.512	0.64	0.8	1	1.25	1.5625	1.953125
ARL = 500	20	61.20	180.20	302.40	329.20	274.00	178.60	78.20
	50	22.40	75.90	278.20	325.50	202.40	55.40	13.20
$\alpha = .002$	250	17.60	32.50	208.10	405.70	87.90	19.60	8.60

La investigación en esta tesis es una versión no paramétrica de la carta de control de Hawkins & Zamba (2005), se utiliza la metodología creada por ... para desarrollar una carta de control no paramétrica basada en el estadístico de la prueba de hipótesis de Conover & Iman desarrollada en la siguiente sección.

2.6. Prueba de igualdad de varianzas de Conover & Iman (1978)

Un método no paramétrico no está limitado a otra línea donde la distribución no se conoce. También se puede aplicar a datos con parámetros y distribuciones conocidas para los cuales no se tiene algún estadístico de prueba desarrollado. A continuación presentamos la prueba no paramétrica de igualdad de varianzas de Conover para dos muestras.

En la mayoría de los casos se busca determinar si existe diferencia entre las medias de dos poblaciones. En este caso hacemos la prueba para la varianza de las poblaciones de manera no paramétrica, que es muy similar a la prueba de medias de Mann-Whitney, pero en lugar de tomar como prueba de hipótesis la igualdad de los valores esperados de las poblaciones usaremos:

$$\begin{aligned}
 H_0: E[(X_i - \mu_1)^2] &= E[(Y_i - \mu_2)^2]; Var(X_i) = Var(Y_i) \\
 H_1: E[(X_i - \mu_1)^2] &\neq E[(Y_i - \mu_2)^2]; Var(X_i) \neq Var(Y_i)
 \end{aligned}
 \tag{35}$$

Donde X_1, \dots, X_n denota una muestra aleatoria de tamaño n de la primer población, y Y_1, \dots, Y_m denota una muestra aleatoria de tamaño m de la segunda población.

Para esto, dos muestras independientes son combinadas, rankeadas y la suma de los rangos de la población X es usado como estadístico de prueba. Cómo el estadístico de prueba es la suma de rangos no es necesario obtener la desviación estándar, ya que al momento de obtener los rangos el resultado de simplemente tomar el valor absoluto de las diferencias entre las observaciones individuales y la media de la población:

$$\begin{aligned} U_i &= |X_i - \mu_1|, i = 1, \dots, n \\ V_j &= |Y_j - \mu_2|, j = 1, \dots, m \end{aligned} \quad (36)$$

Donde μ_1 y μ_2 representan las medias de las poblaciones, si estas son desconocidas de usan las estimaciones \bar{X} y \bar{Y} , N es el total de las observaciones $N = n + m$.

Supuestos:

- Ambas muestras son muestras aleatorias de su respectiva población.
- Las observaciones son mutuamente independientes.
- La escala de medición es al menos de intervalo.

Si no hay valores de U empatados con valores de V , la suma de los cuadrados de las filas asignadas a la población 1 se puede utilizar como estadístico de prueba.

$$T_1 = \sum_{i=1}^n [R(U_i)]^2 \quad (37)$$

Si hay empates, se resta la media de T_1 y se divide entre la desviación estándar para obtener:

$$T = \frac{T - n\bar{R}^2}{\left[\frac{nm}{N(N-1)} \sum_{i=1}^N R^4 - \frac{nm}{N-1} (\bar{R}^2)^2 \right]^{1/2}} \quad (38)$$

Donde $\overline{R^2}$ representa el promedio de los rangos al cuadrado de las muestras combinadas:

$$\overline{R^2} = \frac{1}{N} \left\{ \sum_{i=1}^n [R(U_i)]^2 + \sum_{j=1}^m [R(V_j)]^2 \right\} \quad (39)$$

$\sum_{i=1}^N R^4$ representa la suma de los rangos a la cuarta potencia de ambas muestras:

$$\sum_{i=1}^N R^4 = \sum_{i=1}^n [R(U_i)]^4 + \sum_{j=1}^m [R(V_j)]^4 \quad (40)$$

La distribución nula: Cuantiles de la distribución de T_1 se obtiene por medio de permutaciones (Conover, 1999) para el caso sin empates y tanto n como m menores a 10. Para tamaños más grandes se toma en cuenta la aproximación de muestras grandes, basada en los cuantiles de la distribución normal estándar z_p , y se usa la siguiente aproximación para calcular los cuantiles w_p de T_1 .

$$w_p = \frac{n(N+1)(2N+1)}{6} + z_p \sqrt{\frac{nm(N+1)(2N+1)(8N+11)}{180}} \quad (41)$$

La distribución nula aproximada de T es la distribución Normal estándar.

Las Hipótesis se usa T_1 o T en caso de empates, la H_0 es la misma para cada caso.

A. Prueba de dos colas:

$$H_0: X \text{ y } Y \text{ son idénticamente distribuidas, salvo algún cambio en media} \quad (42)$$

$$H_1: \text{Var}(X) \neq \text{Var}(Y)$$

Se rechaza H_0 con un nivel de error α si:

$$w_{1-\frac{\alpha}{2}} < T \text{ ó } T < w_{\frac{\alpha}{2}} \quad (43)$$

B. Prueba de cola inferior:

$$H_1: \text{Var}(X) < \text{Var}(Y) \quad (44)$$

Se rechaza H_0 con un nivel de error α si:

$$T < w_{\frac{\alpha}{2}} \quad (45)$$

C. Prueba de cola superior:

$$H_1: \text{Var}(X) > \text{Var}(Y) \quad (46)$$

Se rechaza H_0 con un nivel de error α si:

$$w_{1-\frac{\alpha}{2}} < T \quad (47)$$

Esta prueba fue desarrollada originalmente por Taha (1964) quién propuso el estimador por primera vez, en lugar de usar los rangos por si solos. James (1967) examinó la eficiencia asintótica del estadístico. Whiteside, Duran, y Boullion, (1975) aportaron que R_i denota los rangos de $|x_i|$ entre $|x_1|, \dots, |x_N|$. Fligner y Killeen (1976) sugieren substraer la mediana común M de $H(x)$ y $G(x)$, R_i denota el rango de $|x_i - M|$ entre $|x_1 - M|, \dots, |x_N - M|$. Duran y Mielke Jr (1968) muestran que esta prueba es más ponderosa que la prueba t truncada con 2 grados de libertad. Mielke (1972) presentó las aproximaciones asintóticas para los primeros 4 momentos para investigar el supuesto de que la distribución del estadístico es aparentemente Normal. Estas variaciones tiene la misma distribución nula que la original.

CAPÍTULO 3. MODELO

3.1. Propuesta

Se desarrolla una carta de control no-paramétrica para el monitoreo de la varianza, eliminando así el supuesto de que se conoce a priori la distribución de las observaciones. Se sigue la formulación de las cartas de control basadas en el enfoque de punto de cambio propuesta por Hawkins et al. (2003), y así crear una versión no-paramétrica de la carta de control de Hawkins & Zamba (2005).

Usaremos el estadístico de prueba para varianzas iguales de Conover & Iman (1978), cuya metodología ya fue explicada en la sección 2.6 y ahora la retomaremos para el modelo.

3.2. Modelo

Sean X_i independientes y distribuidas de manera:

$$X_i \sim \begin{cases} F_1; & 1 \leq i \leq \tau \\ F_2; & \tau + 1 \leq i \leq N \end{cases} \quad (48)$$

Donde las X_i representan las lecturas de proceso, $E(X_i) = \mu$ es la media de las lecturas de proceso en el cuál permanece constante, $Var(X_i | 1 \leq i \leq \tau) = \sigma_1^2$ que es la varianza del proceso bajo control, $Var(X_i | \tau + 1 \leq i \leq N) = \sigma_2^2$ que es el valor fuera de control de la varianza, τ es el momento en que ocurrió el cambio. N es la última observación del proceso. De esta manera, se desea probar por medio del desarrollo de una carta de control si para algún momento N .

$$\begin{aligned} H_0: \tau &\notin [2, N - 2] \\ H_1: \tau &\in [2, N - 2] \end{aligned} \quad (49)$$

Esto es, si el proceso está en control, significa que no hay indicios de que exista un cambio en la varianza y uno se inclinará por la hipótesis nula, de lo contrario por la alterna.

Siguiendo los lineamientos de Conover (1999) se tiene que para una muestra $X_1, \dots, X_\tau, X_{\tau+1}, \dots, X_N$

Los datos deben cumplir con los supuestos mencionados en la sección previa 2.6, que recapitularemos para este caso:

- Ambas muestras son muestras aleatorias de su respectiva población.
- Las observaciones son mutuamente independientes.
- La escala de medición es al menos de intervalo.

Primer bloque o partición:

$$U_i = |X_i - \bar{x}|, i = 1, \dots, \tau \quad (50)$$

Segundo bloque o partición:

$$V_j = |X_j - \bar{x}|, j = \tau + 1, \dots, N \quad (51)$$

Donde \bar{x} es la media muestral de toda la serie, se jerarquiza toda la serie. Si hay valores de U empatados con valores de V se usa el promedio de los elementos empatados y la suma de los cuadrados de las jerarquías del primer bloque se puede utilizar como la estadístico de prueba.

$$T_{(\tau)} = \sum_{i=1}^{\tau} [R(U_i)]^2 \quad (52)$$

Si hay empates, se resta la media de $T_{(\tau)}$ y se divide por la desviación estándar para obtener:

$$T_{\tau} = \frac{T_{(\tau)} - \tau \overline{R^2}}{\left[\frac{\tau(N-\tau)}{N(N-1)} \sum_{i=1}^N R^4 - \frac{\tau(N-\tau)}{N-1} (\overline{R^2})^2 \right]^{1/2}}; 2 \leq \tau < N-2 \quad (53)$$

Donde $\overline{R^2}$ representa el promedio de los rangos al cuadrado de las muestras combinadas:

$$\overline{R^2} = \frac{1}{N} \left\{ \sum_{i=1}^{\tau} [R(U_i)]^2 + \sum_{j=\tau+1}^N [R(V_j)]^2 \right\} \quad (54)$$

$\sum_{i=1}^N R^4$ representa la suma de los rangos a la cuarta potencia de ambas muestras:

$$\sum_{i=1}^N R^4 = \sum_{i=1}^{\tau} [R(U_i)]^4 + \sum_{j=\tau+1}^N [R(V_j)]^4 \quad (55)$$

Substituyendo las ecuaciones en el estadístico de prueba y asumiendo que no existen empates:

$$T_{\tau} = \sqrt{\frac{5}{(N+1)(2N+1)(8N+11)}} \left[\frac{6T - \tau(N+1)(2N+1)}{\sqrt{\tau(N-\tau)}} \right] \quad (56)$$

Nota:

Se modificó en el diseño del estadístico fue propuesta por el Dr. Conover en una comunicación personal. La media global será tomada en lugar de la media de subgrupos. Tomando en cuenta el supuesto de media constante en el proceso

Usando la forma estandarizada del estadístico de prueba basada en rangos al cuadrado se desarrolla un nuevo estadístico.

$$T'_{\tau} = \max_{2 \leq \tau < N-2} |T_{\tau}| \quad (57)$$

Adicionalmente a la carta de control se puede estimar el momento en que ocurrió el cambio, cuándo la señal de alerta se da es fácil de implementar al encontrar el lugar \hat{t}_T de partición que maximice T_τ :

$$\hat{t}_T = \underset{2 \leq i < N-2}{\operatorname{argmax}} |T_{1,i}| \quad (58)$$

Un proceso es considerado bajo control mientras:

$$T'_i < h_{i,\alpha} \quad (59)$$

Se aplica la secuencia de pasos descrita en la sección 2.4.1 donde se describe el modelo general de Hawkins et al. (2003).

El límite de control $h_{i,\alpha}$ se define resolviendo:

$$P(T'_i \geq h_{i,\alpha} | T'_j < h_{j,\alpha}, \forall j < i) = \alpha \quad (60)$$

La ecuación anterior es demasiado compleja para ser resuelta analíticamente. Incluso en otras cartas de control que usan esta misma metodología mencionan esta problemática. Por tal razón, se simuló una aproximación para estimar los valores de $h_{i,\alpha}$ para diferentes valores de α . Las simulaciones para el cálculo de estos fueron hechas con promedios de series de tamaño 1,000,000 para tener una noción de la variabilidad de los límites y en base a esto se disminuyó (a 100,000 y después a 10,000) el número de iteraciones para agilizar los resultados.

El tiempo de las simulaciones crece conforme aumentan las observaciones que queremos mantener bajo control para obtener el límite $h_{n,\alpha}$, por tal motivo con los datos recabados al momento se buscó una alternativa a la simulación, estos límites a partir de la observación 50 comienzan a estabilizar, por lo cual sólo se estimaron a partir del dato 50 los coeficientes la ecuación de regresión propuesta por Hawkins & Zamba (2005). Los resultados estimados con una suma de cuadrados del error = 0.82.

$$h_{n,\alpha} = \begin{cases} a + b * \ln(\alpha) + \frac{c + d * \ln(\alpha)}{\sqrt{n - e}}, & .001 \leq \alpha < .05 \\ a' + b' * \ln(\alpha) + \frac{c' + d' * \ln(\alpha)}{\sqrt{n - e'}}, & \alpha = .05 \end{cases} ; n \in \mathbb{N} > 50 \quad (61)$$

Dónde:

a = 1.162286035;	a' = 2.134341751;
b = -0.356274258;	b' = 0.016245723;
c = 1.136626645;	c' = 0.002997212;
d = 0.235276633;	d' = -0.159123517;
e = 0.00046156;	e' = -2.60215E-05;

Los coeficientes fueron determinados mediante una optimización bajo la herramienta Solver complemento de Microsoft Excel®. La optimización se hace minimizando la suma de errores cuadrados entre la regresión y los valores reales, las celdas modificadas fueron los coeficientes o parámetros de la función propuesta Hawkins & Zamba (2005).

Los límites de control son presentados en la Tabla 3 las columnas corresponden a los a los $ARL_0 = 20, 50, 100, 200, 500, 1000$ respectivamente, los primeros $10 \leq n \leq 50$, límites de control fueron encontrados por simulación y los restantes fueron estimados con la ecuación (61). Con forme el numero n aumenta se estabilizan los límites.

Tabla 3. Límites de control $h_{n,\alpha}$ a partir de la muestra número 10. Los primeros 40 límites se calcularon mediante simulación y del 50 en adelante se usó la ecuación de regresión.

Índice de observación	n	ARL ₀					
		20	50	100	200	500	1000
	10	2.4059	2.6150	2.6150	2.6444	2.6444	2.6444
	11	2.3008	2.4678	2.5932	2.7414	2.7916	2.7916
	12	2.2396	2.5204	2.6247	2.8038	2.9085	2.9167
	13	2.2500	2.5655	2.6317	2.8784	2.9887	3.0244
	14	2.2248	2.5636	2.6706	2.8799	3.0055	3.1181
	15	2.2132	2.5730	2.6644	2.8606	3.0179	3.2004
	16	2.1888	2.6011	2.6733	2.9126	3.0740	3.2359
	17	2.2087	2.5451	2.7035	2.9148	3.0952	3.2161
	18	2.2273	2.5507	2.7382	2.9448	3.1354	3.2619
	19	2.2053	2.5477	2.7425	2.9763	3.1717	3.2983
	20	2.1893	2.5601	2.7623	2.9290	3.1997	3.3117
	22	2.1949	2.5445	2.7967	2.9494	3.2042	3.3366
	24	2.1802	2.5545	2.8193	3.0033	3.2284	3.3846
	26	2.1833	2.5548	2.8202	3.0288	3.2497	3.4009
	28	2.1735	2.5570	2.8158	3.0038	3.2599	3.4308
	30	2.1699	2.5712	2.8140	3.0240	3.2743	3.4372
	35	2.1695	2.5705	2.8007	3.0104	3.2985	3.4803
	40	2.1585	2.5816	2.8176	3.0134	3.3252	3.5277
	45	2.1508	2.5792	2.8177	3.0510	3.3398	3.4896
	50	2.1495	2.5702	2.8158	3.0663	3.3186	3.5402
	60	2.1476	2.5827	2.8089	3.0351	3.3341	3.5603
	70	2.1430	2.5808	2.8085	3.0362	3.3371	3.5648
	80	2.1393	2.5793	2.8082	3.0370	3.3396	3.5685
	90	2.1362	2.5780	2.8079	3.0378	3.3416	3.5715
	100	2.1336	2.5770	2.8077	3.0384	3.3434	3.5741
	125	2.1286	2.5749	2.8072	3.0396	3.3468	3.5791
	150	2.1248	2.5733	2.8069	3.0405	3.3492	3.5828
	175	2.1219	2.5721	2.8067	3.0412	3.3512	3.5857
	200	2.1196	2.5712	2.8065	3.0417	3.3527	3.5880
	250	2.1160	2.5697	2.8061	3.0426	3.3551	3.5916
	300	2.1134	2.5686	2.8059	3.0432	3.3569	3.5942
	350	2.1113	2.5678	2.8057	3.0437	3.3583	3.5962
	400	2.1097	2.5671	2.8056	3.0441	3.3594	3.5979
	500	2.1071	2.5661	2.8054	3.0447	3.3611	3.6004

CAPÍTULO 4. DESEMPEÑO

4.1. Diseño de la experimentación Monte Carlo

Con los límites de control (ver Tabla 3) y la ecuación de regresión (ecuación (61)) se dispuso a calcular el desempeño de la carta propuesta en términos de su **ARL** en control y fuera de control. Para el cálculo del **ARL₀** se promedió el momento en el que la carta dio la primer señal de alarma en 10,000 series de datos replicados, todos bajo control, y para el **ARL₁** se hicieron 30 diferentes experimentaciones para cada nivel de error tipo I. En cada experimentación se combinaron 6 diferentes proporciones de cambio y 5 momentos diferentes de cambio. Al momento de calcular el **ARL₁** se tomaron en cuenta sólo las series de datos en que resultaran en control hasta el momento del verdadero cambio, y si resultaba una falsa alarma se desechaba dicha serie (esto para evitar mezclar resultados del desempeño en control con el análisis fuera de control).

Se buscó tener diferentes distribuciones para hacer la comparativa, donde sea adecuado usar otra carta de control paramétrica y donde no es adecuado usar alguna existente en la literatura. En este caso la Normal es bien estudiada y ya se cuentan con cartas de control específicamente para ella, en cambio para la distribución Gamma no se tiene ninguna carta conocida por el autor cuando los parámetros son desconocidos. Estas fueron las dos distribuciones usadas. En todas se cambiaron los parámetros para la hora de calcular el **ARL₁** para obtener la misma media y diferentes desviaciones. Con la normal no hay problema ya que eso son sus parámetros en cambio para la Gamma se hizo uso de un ajuste para lograr mantener la media cambiando sólo la desviación.

Para hacer la experimentación se usó el software MatLab®.

4.2. Desempeño de la carta de control propuesta

En esta investigación se hace experimentación con series de observaciones normales para verificar si existe un similar desempeño en comparación con las cartas de control que necesitan los supuestos de Normalidad en los datos.

En las siguientes tablas se muestran el desempeño en términos de **ARL** de la carta de control propuesta para la distribución Normal. En la tabla se indican por bloques de filas cada α (.05, .02, .01, .005, .002, .001), y cada fila de cada bloque muestra el desempeño según el momento donde ha ocurrido el cambio, $\tau+1$ (20, 50, 80, 100, 250). Las columnas representan la proporción del cambio dónde se realiza el cociente de la segunda desviación entre la primera, σ_2/σ_1 (0.512, 0.64, 0.8, 1, 1.25, 1.563, 1.95313). Si el cociente es 1 significa que no ocurrió cambio ambas desviaciones son iguales. Si el resultado es menor a uno quiere decir que el cambio fue hacia abajo, (disminuyó la variabilidad), De lo contrario si la división resulta mayor a uno el cambio fue hacia a arriba (aumentó la variabilidad). La columna central muestra el **ARL₀** de la carta, donde se observa que sobrepasa en casi todos los casos el **ARL** teórico. Las columnas alrededor del **ARL₀** muestran el **ARL₁** que disminuye conforme el cambio es más evidente.

En la Tabla 4 se muestran los resultados del desempeño cuándo los datos tienen distribución Normal. No se cuentan con algunos resultados, ya que el tiempo de estimación de ellos era demasiado grande y se decidió no realizarlos. Estos están representados con un (*) en las tablas.

Como se puede observar el **ARL₀** de la carta en casi todos los casos se aproxima por la derecha al **ARL₀** teórico, y **ARL₁** disminuye con respecto el cambio aumenta. Además se puede observar que el desempeño en detectar cambios inmediatos no es tan bueno como detectar un cambio a partir de 80 observaciones en control. Este es un problema “típico” de las cartas de control para varianzas por la falta de poder cuando no se conocen los parámetros iniciales y hay que estimarlos.

Tabla 4. Desempeño en términos de ARL de la carta propuesta. Se marca con el símbolo

Teórico en Control	$\tau+1$	σ_2/σ_1 (Normal)						
		0.512	0.64	0.8	1	1.25	1.563	1.95313
ARL = 20 $\alpha = .05$	20	14.15	18.90	21.30	28.56	14.89	10.75	7.52
	50	8.45	14.51	20.91	28.52	11.53	6.29	3.96
	80	7.63	12.87	20.60	29.06	10.44	5.90	3.68
	100	7.40	12.46	20.56	28.90	10.32	5.66	3.63
	250	*	*	*	28.84	*	*	*
ARL = 50 $\alpha = .02$	20	42.51	55.07	57.86	61.27	40.80	30.52	22.53
	50	16.79	35.31	58.80	60.80	28.13	13.04	6.77
	80	13.33	27.21	55.39	60.52	23.94	10.34	5.72
	100	12.57	24.92	55.08	60.84	22.45	9.69	5.48
	250	11.23	20.37	49.19	60.72	18.84	8.85	4.97
ARL = 100 $\alpha = .01$	20	83.28	103.08	108.63	104.26	80.77	64.58	49.17
	50	25.36	59.56	106.59	104.79	51.57	21.63	10.21
	80	17.80	41.45	99.89	104.85	40.62	15.38	7.72
	100	16.37	35.71	96.97	105.30	36.23	13.80	7.23
	250	14.38	26.71	77.53	105.31	29.31	11.66	6.29
ARL = 200 $\alpha = .005$	20	174.85	204.14	212.02	203.39	168.50	143.19	117.54
	50	41.63	115.08	207.47	203.48	109.10	44.31	16.82
	80	24.26	67.14	194.38	205.34	81.95	25.13	10.96
	100	21.62	52.63	183.02	206.73	71.42	21.98	9.92
	250	17.82	34.51	125.03	205.08	47.75	16.19	8.12
ARL = 500 $\alpha = .002$	20	455.53	518.52	519.98	491.40	438.60	392.04	351.03
	50	86.93	287.64	506.86	490.18	305.07	126.95	37.17
	80	35.62	139.36	482.75	492.63	225.92	53.43	17.50
	100	30.32	96.35	437.57	503.79	190.68	39.66	15.02
	250	23.45	46.70	248.11	489.86	92.09	23.40	11.36
ARL = 1000 $\alpha = .001$	20	1005.74	1092.20	1091.47	1055.82	994.23	910.97	832.36
	50	180.21	650.29	1137.90	1058.30	733.79	339.93	91.83
	80	50.26	286.11	1012.28	1048.59	556.52	112.27	24.78
	100	38.31	159.33	933.85	1065.97	458.28	72.79	19.72
	250	28.17	58.51	445.89	1067.59	174.76	31.58	14.30

“*” los datos con los que no se cuenta por exceso de tiempo de cómputo.

En la tabla 2

Tabla 2, que repetimos de la sección 2.5 se muestran los resultados publicados por Hawkins y Zamba (2005) del desempeño en términos del **ARL** de su carta de control paramétrica de punto de cambio para varianzas. Con un error tipo I predeterminado de .002 consigue un **ARL** alrededor de 350, con cambio sólo en la varianza.

). Desempeño en términos de **ARL** de la carta de Hawkins & Zamba (2005).

Teórico en Control	$\tau+1$	σ_2/σ_1 (Normal)						
		0.512	0.64	0.8	1	1.25	1.5625	1.953125
ARL = 500	20	61.20	180.20	302.40	329.20	274.00	178.60	78.20
	50	22.40	75.90	278.20	325.50	202.40	55.40	13.20
$\alpha = .002$	250	17.60	32.50	208.10	405.70	87.90	19.60	8.60

No se puede hacer una comparación directa con la versión paramétrica Hawkins y Zamba (2005) de esta carta debido al bajo nivel de falsas alarmas con las que cuenta la carta, que con la misma probabilidad de error tipo I en ambas se tienen diferencias muy marcadas de ARL_0 . La comparación no se puede dar ya que como el nivel de falsas alarmas es mayor, al momento de calcular el ARL_1 esta carta tendrá un nivel más pequeño, el cual se busca comparar en este caso. Por esto se procede a comparar con una que tenga nivel comparable a la propuesta en este trabajo. Se decide evaluar el CUSUM auto-inicial de Hawkins (1981) que elimina el supuesto de los parámetros conocidos pero sigue bajo la necesidad de normalidad en los datos.

Por otro lado para comparar el desempeño en distribuciones con muchos puntos atípicos, donde las actuales cartas de control no son apropiadas y nuestra carta puede tener un mejor desempeño se procede a utilizar la distribución Gamma.

Tabla 5. Desempeño en términos de ARL de la carta propuesta para la distribución Gamma.

Teórico en Control	$\tau+1$	σ_2/σ_1 (Gamma(0.5))						
		0.512	0.64	0.8	1	1.25	1.5625	1.95313
ARL = 20 $\alpha = .05$	20	6.73	9.75	15.01	28.62	11.78	7.34	4.73
	50	5.26	7.17	12.38	28.90	7.92	4.37	2.83
	80	4.95	6.63	11.08	28.70	7.02	3.90	2.56
	100	4.76	6.39	10.91	28.57	6.51	3.71	2.45
	250	*	*	*	28.73	*	*	*
ARL = 50 $\alpha = .02$	20	12.63	21.63	37.34	59.60	30.30	17.33	10.12
	50	9.37	14.01	27.89	59.31	19.56	8.79	5.08
	80	8.84	12.58	24.46	60.05	15.62	7.19	4.38
	100	8.63	12.16	23.10	59.96	14.65	6.62	4.11
	250	8.12	11.26	20.42	60.09	12.00	5.90	3.75
ARL = 100 $\alpha = .01$	20	20.07	37.36	72.24	104.14	59.09	32.19	18.22
	50	12.19	19.79	48.44	104.87	32.16	12.49	6.87
	80	11.38	17.09	37.80	105.70	23.14	9.55	5.62
	100	11.10	16.41	35.76	102.26	21.42	8.99	5.44
	250	10.65	14.80	29.04	105.26	16.68	7.63	4.84
ARL = 200 $\alpha = .005$	20	31.78	69.23	144.02	204.53	124.10	68.74	35.54
	50	15.68	28.12	84.77	203.96	60.84	19.31	8.94
	80	14.35	22.13	62.83	201.79	39.60	13.20	7.18
	100	14.03	21.43	55.96	205.66	33.41	12.00	6.84
	250	13.33	19.05	39.55	201.13	22.54	9.96	6.11
ARL = 500 $\alpha = .002$	20	68.58	175.21	372.76	490.53	345.32	205.76	106.75
	50	21.06	44.29	209.95	497.94	161.80	40.35	13.73
	80	18.85	30.61	123.42	490.11	90.01	20.18	9.78
	100	18.25	28.39	98.01	483.66	64.95	17.03	9.17
	250	17.05	25.10	56.21	489.55	33.17	13.26	7.92
ARL = 1000 $\alpha = .001$	20	155.97	424.04	863.83	1056.56	791.14	535.89	291.99
	50	26.65	69.06	476.85	1040.95	405.38	85.94	19.54
	80	22.64	38.65	242.40	1045.23	203.40	29.03	12.37
	100	21.65	34.58	174.71	1042.18	133.36	21.83	11.24
	250	20.11	30.09	72.07	1048.77	43.82	16.51	9.51

“*” los datos con los que no se cuenta por exceso de tiempo de cómputo.

En la Tabla 5 se muestran los datos cuándo tienen distribución Gamma. Como se puede apreciar los resultados del ARL_0 son muy similares al de la Normal, pero existe alguna diferencia muy ligera al comparar el ARL_1 de la Gamma con el de la Normal.

4.3. Carta de control Auto-inicial CUSUM para varianzas

El desempeño en términos de ARL de la SSCUSUM bajo la distribución Normal se muestra en la Tabla 6 donde es debidamente aplicada al cumplirse el supuesto de normalidad. Se marca con “*” los datos con los que no se cuenta por exceso de tiempo de cómputo.

El desempeño en términos de ARL de la SSCUSUM bajo la distribución Gamma se muestra en la Tabla 7 donde no es apropiada ya que no cumple con los supuestos de normalidad. Se marca con “*” los datos con los que no se cuenta por exceso de tiempo de cómputo.

En ambas cartas se usa una SSCUSUM fijando un $\alpha = .002$. Para observaciones H_0 mala de donde se espera tener un $ARL_0 = 500$. En las tablas siguientes se muestran los resultados del desempeño por simulación. Se puede notar como el desempeño con datos Normales es bueno, pero al verificar el desempeño de la carta SSCUSUM con datos no normales, se observa que estos no coinciden con los teóricos.

En conclusión de la carta no paramétrica propuesta es robusta a la distribución de los datos y arroja desempeños aceptables en términos de ARL . En cambio la carta paramétrica SSCUSUM sólo arroja un buen desempeño cuándo las condiciones de normalidad se cumplen. Si tratamos de ignorar el supuesto de normalidad el desempeño cae y las falsas alarmas aumentan de manera considerable (casi el doble).

Tabla 6. Desempeño en términos de **ARL** del SSCUSUM para la distribución Normal.

Teórico en Control	$\tau+1$	σ_2/σ_1 (SSCUSUM-Normal)						
		0.512	0.64	0.8	1	1.25	1.5625	1.95313
ARL = 20 $\alpha = .05$	20	11.57	16.61	19.46	22.25	11.74	6.82	3.82
	50	9.05	14.77	19.75	22.35	10.32	5.44	3.02
	80	8.20	13.77	19.80	22.10	9.89	5.04	2.98
	100	8.11	13.63	19.95	22.27	9.72	5.02	2.83
	250	*	*	*	22.41	*	*	*
ARL = 50 $\alpha = .02$	20	25.19	41.70	53.06	53.68	32.52	17.81	8.23
	50	16.04	33.35	51.76	53.55	26.59	10.83	5.25
	80	13.84	29.24	51.68	54.39	23.48	9.58	4.81
	100	13.30	28.00	51.13	53.50	22.43	9.32	4.67
	250	11.86	23.80	48.21	53.82	19.73	8.29	4.39
ARL = 100 $\alpha = .01$	20	51.04	88.36	110.06	108.89	73.66	40.52	17.57
	50	25.83	66.32	108.52	110.25	54.43	19.10	7.76
	80	20.30	54.31	102.37	111.19	46.33	15.16	6.54
	100	19.09	48.52	104.30	109.09	43.20	13.98	6.25
	250	15.90	37.00	93.84	110.23	35.02	11.98	5.84
ARL = 200 $\alpha = .005$	20	107.75	183.93	226.68	222.59	160.69	96.58	46.57
	50	45.86	135.21	223.57	220.85	121.96	39.81	12.12
	80	29.96	108.15	218.09	222.10	99.23	25.15	9.07
	100	25.82	93.14	215.21	219.36	89.37	21.95	8.38
	250	20.20	60.60	186.28	220.06	60.82	16.41	7.43
ARL = 500 $\alpha = .002$	20	314.73	509.34	587.90	567.98	465.78	329.63	182.04
	50	125.43	393.47	584.41	570.45	367.73	138.94	29.19
	80	61.76	297.69	584.39	579.24	294.88	68.17	14.16
	100	44.80	257.89	571.80	573.10	257.16	49.89	12.47
	250	27.51	126.97	501.03	563.79	144.32	25.30	9.59
ARL = 1000 $\alpha = .001$	20	747.06	1088.75	1196.25	1153.91	1023.23	783.39	510.60
	50	318.40	882.98	1228.35	1178.86	851.48	366.68	85.73
	80	138.54	709.14	1194.91	1197.42	705.94	180.05	25.89
	100	89.02	617.02	1216.24	1175.02	635.48	129.43	17.51
	250	33.52	259.37	1104.23	1144.59	338.94	33.88	11.59

“*” los datos con los que no se cuenta por exceso de tiempo de cómputo.

Tabla 7. Desempeño en términos de **ARL** del SSCUSUM para la distribución Gamma.

Teórico en Control	$\tau+1$	σ_2/σ_1 (SSCUSUM-Gamma)						
		0.512	0.64	0.8	1	1.25	1.563	1.95313
ARL = 20 $\alpha = .05$	20	9.11	12.43	15.52	17.96	12.93	9.99	6.82
	50	7.12	11.56	15.44	17.87	13.62	10.04	6.12
	80	6.55	11.06	15.63	17.87	14.28	9.94	6.02
	100	6.29	10.58	15.39	17.82	13.97	9.83	5.75
	250	*	*	*	17.85	*	*	*
ARL = 50 $\alpha = .02$	20	17.13	26.78	33.19	36.05	27.12	19.87	13.24
	50	12.01	22.74	33.78	35.66	26.91	17.74	10.27
	80	10.55	21.02	33.19	35.93	26.88	16.95	9.32
	100	9.91	20.08	33.53	36.14	27.31	16.91	8.86
	250	9.97	19.90	33.70	35.79	26.60	16.59	9.03
ARL = 100 $\alpha = .01$	20	27.79	46.57	59.53	65.38	51.06	37.50	23.40
	50	16.74	36.71	60.66	62.83	48.14	29.41	15.15
	80	13.39	31.54	59.34	64.12	46.60	26.82	13.31
	100	12.65	29.87	58.43	64.72	45.51	26.17	12.45
	250	11.02	23.88	54.97	64.06	44.44	22.53	10.84
ARL = 200 $\alpha = .005$	20	47.65	84.43	107.89	114.72	95.19	72.40	48.77
	50	22.33	60.29	105.67	117.52	89.06	54.54	28.10
	80	17.28	50.22	101.59	114.49	84.88	45.49	20.46
	100	15.12	43.22	102.09	115.57	81.80	43.18	18.50
	250	13.18	30.64	90.40	115.93	73.25	34.47	14.58
ARL = 500 $\alpha = .002$	20	98.72	181.54	238.87	257.99	224.50	186.12	128.22
	50	38.11	123.91	229.36	253.88	206.81	134.56	70.02
	80	23.63	89.19	224.37	255.58	194.32	107.20	43.90
	100	20.30	77.43	212.65	256.89	186.41	96.27	35.93
	250	16.21	42.22	182.16	253.79	157.92	64.07	21.55
ARL = 1000 $\alpha = .001$	20	189.22	316.95	409.89	434.72	395.48	340.37	262.20
	50	61.01	216.56	383.90	436.90	368.77	259.89	146.01
	80	31.38	158.37	379.44	441.23	347.33	211.74	87.40
	100	24.72	122.45	368.74	445.99	333.77	186.66	65.27
	250	18.73	54.96	309.02	437.56	278.40	105.13	29.73

“*” los datos con los que no se cuenta por exceso de tiempo de cómputo.

4.4. Ejemplo Numérico

En este ejemplo se comparará la carta de control SSCUSUM contra la carta propuesta, tomando datos comprendidos en el periodo de Julio 2004 a Julio 2009 de uno de los índices bursátiles más importantes de EEUU el conocido **S&P 500**. Se analizará mediante el software Minitab® el posible cambio en la varianza de los datos. Los datos se muestran en la Tabla 8 y gráficamente en la **Figura 2** es evidente una disminución en la varianza de los datos.

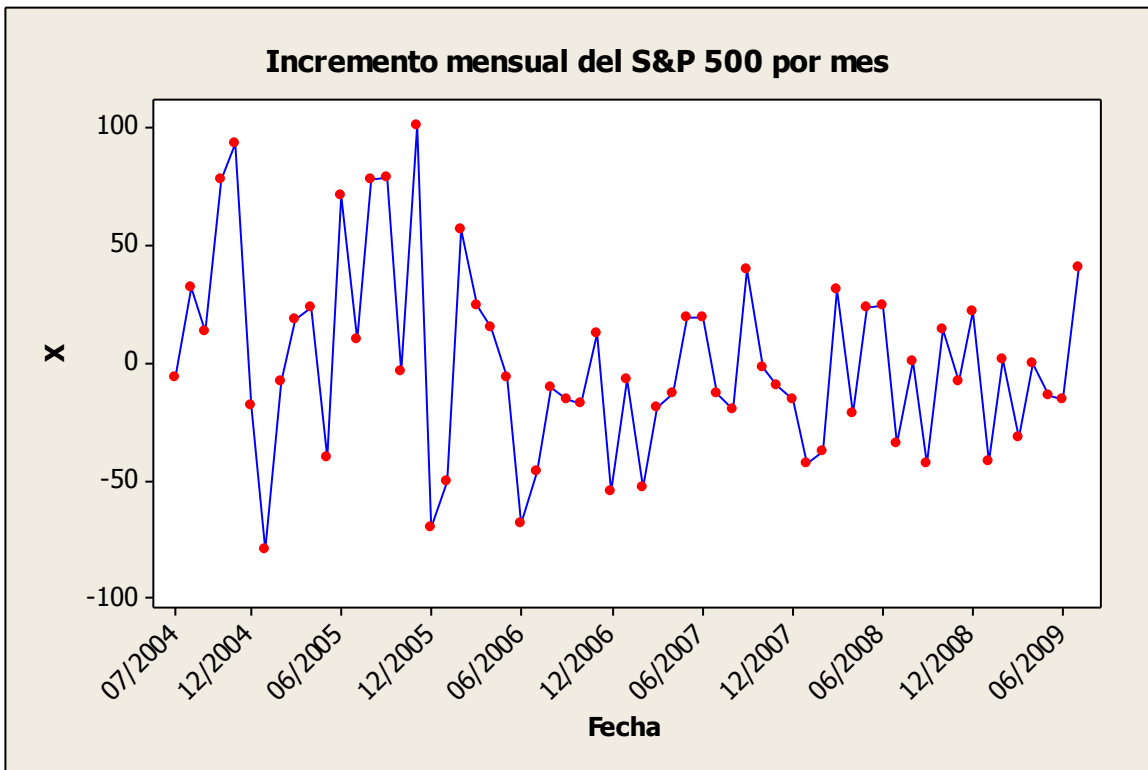


Figura 2. Datos comprendidos en el periodo de Julio 2004 a Julio 2009 del índice S&P 500.

Tabla 8. Datos del S&P 500 fechados para la SSCUSUM

SSCUSUM			Fecha	X	V
Fecha	X	V	01/2007	-7.49	-0.87
07/2004	-6.36		02/2007	-53.71	0.72
08/2004	31.40		03/2007	-19.21	-0.43
09/2004	13.19	-2.02	04/2007	-13.81	-0.68
10/2004	77.65	1.19	05/2007	18.73	-0.76
11/2004	92.64	0.72	06/2007	18.91	-0.76
12/2004	-18.84	0.57	07/2007	-13.38	-0.65
01/2005	-79.67	1.39	08/2007	-20.15	-0.35
02/2005	-8.63	-0.67	09/2007	39.12	0.14
03/2005	17.88	-1.54	10/2007	-2.52	-1.31
04/2005	23.47	-1.16	11/2007	-10.35	-0.79
05/2005	-40.65	0.47	12/2007	-15.62	-0.53
06/2005	70.46	0.62	01/2008	-43.63	0.52
07/2005	9.79	-1.53	02/2008	-38.09	0.31
08/2005	77.32	0.74	03/2008	30.65	-0.04
09/2005	78.20	0.63	04/2008	-22.33	-0.28
10/2005	-4.45	-0.33	05/2008	23.01	-0.35
11/2005	100.79	1.12	06/2008	23.74	-0.33
12/2005	-70.48	1.36	07/2008	-34.65	0.25
01/2006	-50.55	0.77	08/2008	0.17	-1.92
02/2006	56.49	0.00	09/2008	-42.85	0.52
03/2006	24.12	-1.46	10/2008	13.85	-0.76
04/2006	14.55	-1.57	11/2008	-8.48	-1.04
05/2006	-7.03	-0.40	12/2008	21.81	-0.33
06/2006	-68.74	1.22	01/2009	-42.47	0.53
07/2006	-46.67	0.62	02/2009	1.19	-1.88
08/2006	-10.91	-0.54	03/2009	-31.79	0.13
09/2006	-15.81	-0.37	04/2009	-0.58	-2.26
10/2006	-17.70	-0.32	05/2009	-14.16	-0.71
11/2006	12.05	-1.63	06/2009	-15.78	-0.62
12/2006	-54.75	0.79	07/2009	40.52	0.53

Los datos mostrados en la Tabla 8 fechados y los datos de la columna V son las transformaciones para monitorear la variabilidad de la ecuación (13), mostrada en secciones anteriores.

La carta de control SSCUSUM se presenta en la siguiente Figura 3. Con los parámetros $k = .25$ y $h = 6$, se calcula un $ARL \approx 126$.

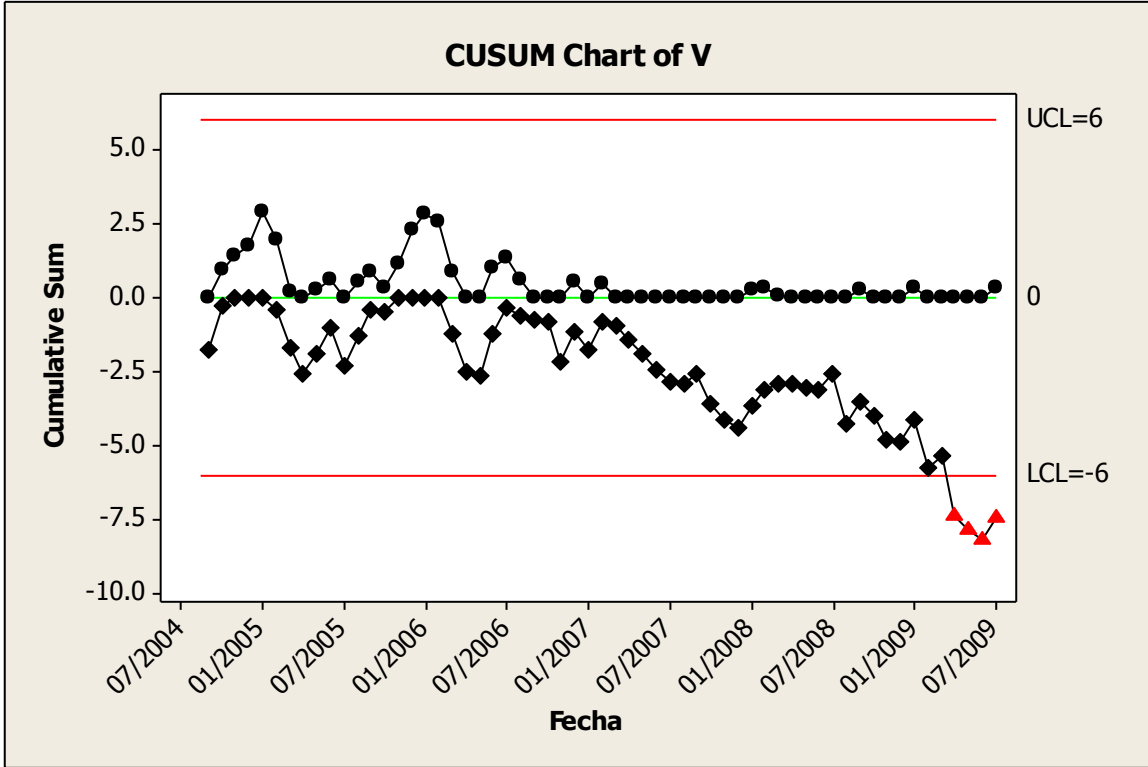


Figura 3. Carta de control SSCUSUM para el índice bursátil S&P 500.

La SSCUSUM nos mostró alarmas en los datos 58, 59, 60, 61, de aquí podemos concluir que en el dato 58 la serie de datos está fuera de control, se da una señal de que la varianza disminuyó.

Ahora se presenta el desempeño de la carta de control propuesta para comparar resultados. Se calculan los seis niveles de α con los que contamos, esto para tener una idea del desempeño de la carta. Se puede observar de la Figura 4 cómo se comportan los límites de control (en línea roja), el estadístico para observación (en línea azul) y las señares de alarma (* rojo) para cada nivel de α . Aquí se muestran todos los límites disponibles, pero para comparar contra un $ARL \approx 126$ de la SSCUSUM podemos escoger $\alpha = .01$ que tiene un $ARL = 100$ teórico, que se acerca por la izquierda eso nos da a entender que la señal de alarma de la carta propuesta puede estar en promedio un poco más abajo de la señal de

alarma de la SSCUSUM. En conclusión ambas cartas tienen un desempeño similar, pero la carta propuesta tienen un mejor desempeño y un poco mejor, y con menos supuestos.

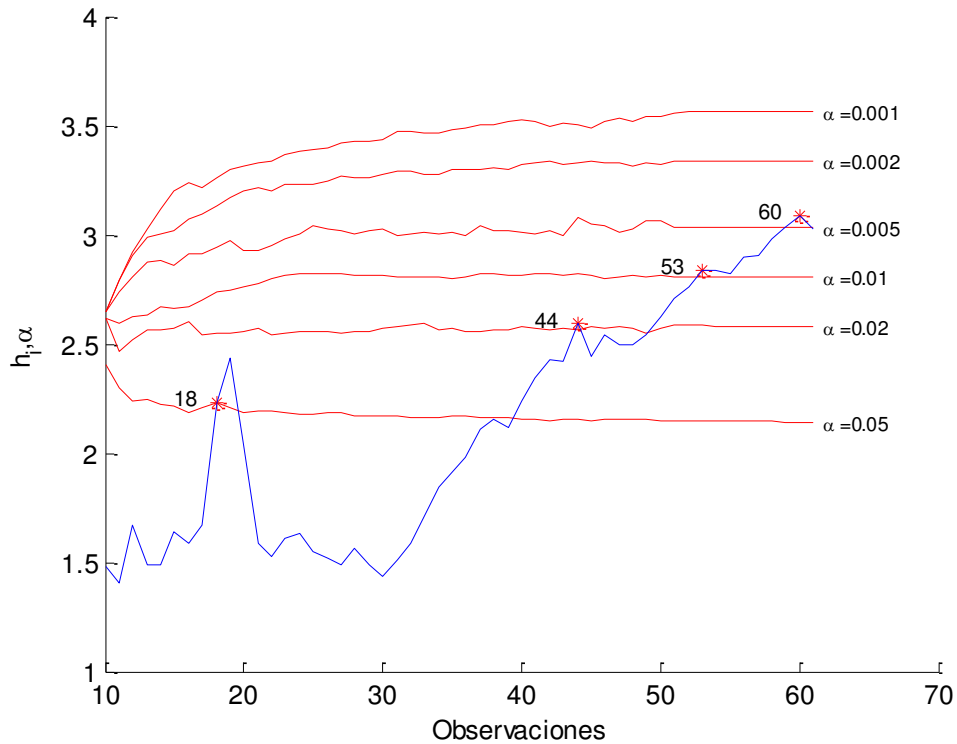


Figura 4. Desempeño de la carta propuesta con los diferentes niveles α

En la siguiente tabla mostramos los datos de la experimentación, se presentan los 10 primeros datos estables que se usan de calentamiento, se comprueba que estos datos esten bajo control y se evalúan con el primer límite de control, se muestran también los límites de control $h_{i,\alpha}$ que deben ser superados para considerar que existe un cambio bajo cierto nivel de significancia en este caso el error tipo I lo mostramos con los valores α (.05, .02, .01, .005), que fue donde se detectaron alarmas (18, 44, 53 y 60) respectivamente.

Tabla 9. Experimentación de la tabla de control propuesta.

Propuesta											
X	T_i	$h_{i,.05}$	$h_{i,.02}$	$h_{i,.01}$	$h_{i,.005}$	X	T_i	$h_{i,.05}$	$h_{i,.02}$	$h_{i,.01}$	$h_{i,.005}$
-6.36	**	**	**	**	**	-7.49	1.51	*	2.58	2.81	2.99
31.40	**	**	**	**	**	-53.71	1.59	*	2.59	2.81	3.00
13.19	**	**	**	**	**	-19.21	1.70	*	2.59	2.81	3.01
77.65	**	**	**	**	**	-13.81	1.84	*	2.57	2.80	3.00
92.64	**	**	**	**	**	18.73	1.91	*	2.57	2.80	3.01
-18.84	**	**	**	**	**	18.91	1.98	*	2.56	2.81	3.00
-79.67	**	**	**	**	**	-13.38	2.11	*	2.56	2.82	3.04
-8.63	**	**	**	**	**	-20.15	2.16	*	2.56	2.82	3.02
17.88	**	**	**	**	**	39.12	2.12	*	2.56	2.82	3.02
23.47	1.48	2.41	2.62	2.62	2.64	-2.52	2.24	*	2.58	2.82	3.01
-40.65	1.41	2.30	2.47	2.59	2.74	-10.35	2.34	*	2.57	2.82	3.00
70.46	1.67	2.24	2.52	2.62	2.80	-15.62	2.43	*	2.56	2.82	3.02
9.79	1.49	2.25	2.57	2.63	2.88	-43.63	2.42	*	2.57	2.81	3.00
77.32	1.49	2.22	2.56	2.67	2.88	-38.09	2.59	*	2.56	2.82	3.08
78.20	1.64	2.21	2.57	2.66	2.86	30.65	2.44	*	*	2.82	3.05
-4.45	1.59	2.19	2.60	2.67	2.91	-22.33	2.54	*	*	2.80	3.04
100.79	1.67	2.21	2.55	2.70	2.91	23.01	2.49	*	*	2.81	3.01
-70.48	2.23	2.23	2.55	2.74	2.94	23.74	2.49	*	*	2.81	3.03
-50.55	2.43	*	2.55	2.74	2.98	-34.65	2.54	*	*	2.81	3.06
56.49	2.04	*	2.56	2.76	2.93	0.17	2.63	*	*	2.82	3.07
24.12	1.58	*	2.57	2.78	2.93	-42.85	2.71	*	*	2.81	3.04
14.55	1.52	*	2.54	2.80	2.95	13.85	2.76	*	*	2.81	3.04
-7.03	1.61	*	2.55	2.81	2.98	-8.48	2.84	*	*	2.81	3.04
-68.74	1.63	*	2.55	2.82	3.00	21.81	2.84	*	*	*	3.04
-46.67	1.55	*	2.56	2.82	3.04	-42.47	2.82	*	*	*	3.04
-10.91	1.52	*	2.55	2.82	3.03	1.19	2.89	*	*	*	3.04
-15.81	1.49	*	2.55	2.82	3.02	-31.79	2.91	*	*	*	3.04
-17.70	1.56	*	2.56	2.82	3.00	-0.58	2.98	*	*	*	3.04
12.05	1.49	*	2.56	2.81	3.02	-14.16	3.03	*	*	*	3.04
-54.75	1.43	*	2.57	2.81	3.02	-15.78	3.09	*	*	*	3.04
						40.52	3.03	*	*	*	*

Los datos “**” no aplica limite de control, y los reemplazados con “*” no son necesarios para el ejemplo.

CAPÍTULO 5. CONCLUSIONES

5.1. Conclusiones

En desarrollo de procesos y servicios (entre otras cosas) cuándo existe un déficit de calidad que se desea disminuir, una opción para incrementar la calidad es reducir la variabilidad de las características que se estén midiendo de los procesos y/o servicios. En este caso las cartas de control tradicionales paramétricas nos son útiles para la tarea de monitorear y controlar la variabilidad cuando los procesos no siguen una distribución normal. En la actualidad no existen muchas maneras de controlar esta variabilidad de forma eficiente y sin asumir ciertas condiciones que deben cumplir los datos, ya que en el mundo real esto es necesario. El problema tratado en esta investigación el de crear una carta de control no paramétrica siguiendo los lineamientos del punto de cambio, para detectar cambios sostenidos en la varianza. Se definieron los objetivos de: crear la carta de control, evaluar el desempeño en casos donde exista normalidad y evaluarla cuando no se cumplen los supuestos de normalidad. Cumpliendo de forma favorable los tres objetivos. Se es consciente de algunas limitantes, pero las ventajas son muy favorecedoras.

Se necesitó una revisión de literatura para mostrar las cartas de control existentes, se describe sus características, se describe como se evalúa el desempeño en términos del *ARL* y se muestra evidencia de que existe una necesidad por atender el problema planteado. También se mostraron las diferentes metodologías para poder seleccionar la carta de control contra la que se comparó la propuesta. Se mostró la metodología usada para la carta propuesta, y la versión paramétrica de nuestra propuesta, contra la que no pudimos hacer una mejor comparación ya que el desempeño no es el mismo. Se explica el estadístico no paramétrico usado.

Durante el desarrollo de la carta de control se hicieron ajustes para realizar de mejor manera los cálculos tomando en cuenta una observación del autor del estadístico no

paramétrico utilizado. En las hipótesis de la prueba solo existe un cambio o no existe un cambio en la dispersión de los datos, siguiendo la metodología de Conover con sus respectivos supuestos, se calculan los límites de acuerdo con la metodología de punto de cambio que fue introducida por Hawkins et al. (2003). La carta de control evalúa un estadístico con los primeros 10 datos estables para asegurarnos que estén bajo control, y por cada observación nueva que agregamos se calcula dicho estadístico, se grafican los estadísticos por cada observación.

Para medir el desempeño en términos del ARL se calculó el estadístico para cada observación nueva agregada y se evalúa si es menor a los límites de control, si esto es cierto seguimos haciendo repetidamente pruebas de hipótesis hasta que nos arroje la primer señal de alarma, en teoría si no hay cambio existe una probabilidad de que nos señale una alarma, esa probabilidad es muy baja y se representa con la letra griega alpha (α) que nosotros establecemos, el ARL_0 entre más grande mejor. En caso contrario si existe un cambio en la distribución se espera que el ARL_1 sea lo más pequeño posible. Se deben comparar ambos parámetros, por lo general se puede definir un ARL_0 específico para ambas cartas a comparar, con el mismo nivel de falsas alarmas en las cartas se procede a comparar el ARL_1 y el que resulte más pequeño tiene mejor desempeño, pero si dos cartas de control con el mismo nivel de significancia no tienen el mismo ARL_0 no es comparable el ARL_1 , aquí la ventaja es de la carta con el ARL_0 más grande.

Del primer objetivo se usó la metodología de Hawkins & Zamba, 2005, para encontrar los límites de la carta control para varianzas con seis niveles de error tipo I (α), usando el estadístico modificado de Conover & Iman, esto nos dio como resultado una carta de control para varianzas, donde una falsa alarma indica que el sistema salió de control en un momento previo y se cuenta con la posibilidad de poder estimar el momento en que ocurrió el cambio, tomando las ventajas de no necesitar el conocimiento a priori de la distribución de los datos ni de los parámetros, a menos que tengamos indicios de cambios

en la media. Las simulaciones para la creación de los límites de control fue la parte más exhaustiva del proyecto.

Para cumplir con el segundo objetivo se realizaron una serie de simulaciones de las cuales sólo tenemos material para comparar una de ellas con el homólogo paramétrico que queremos comparar, los resultados nos favorecen en el sentido de las falsas alarmas, ya que nuestra carta parece ser robusta en este sentido, pero no se puede dar una completa comparación de los desempeños cuándo existe un cambio, de este modo en la carta que queremos comparar crece la probabilidad de encontrar falsos positivos. Lo cual sesga los resultados y no representa una verdadera comparación. Se buscó otra metodología que presente similitudes en los desempeños, usamos una variante de la ya muy usada CUSUM, usamos esta metodología para aplicarla a la varianza y que fuera de auto-inicio, para omitir el conocimiento de los parámetros iniciales. Los resultados son muy similares y nuestra carta nos da la satisfacción de estar al nivel de las mejores cartas de control paramétricas.

Lo más relevante de la investigación viene en el tercer objetivo, ya que tener resultados favorables en la comparación del método con distribuciones para las cuales no se han desarrollado métodos, pone esta propuesta como la primera en el tipo con buenos resultados y podríamos decir que sería la próxima a vencer. En efecto se dieron resultados son positivos y se podría decir que el desempeño de la carta mejora con distribuciones con muchos puntos atípicos como lo es la distribución Gamma con un parámetro $\alpha = .5$ (parecida a una exponencial). En esta investigación se concluye que el desempeño en control de nuestra carta en control es casi perfecto al ser nuestro $ARL_0 \cong 1/\alpha$, cosa que no se puede decir de la versión paramétrica de Hawkins & Zamba (2005), y al compararla con la SSCUSUM para varianzas tienen ambos desempeños en el caso de usar datos con distribución Normal, y al usar distribuciones sesgadas nuestra carta tiene completamente la ventaja.

5.2. Trabajo futuro

5.2.1. Carta de control de una cola

La carta de control propuesta es de dos colas esto quiere decir que detecta cambios hacia ambos lados, la propuesta es desarrollar los límites de control de una cola, esto es para casos específicos donde tal vez no nos interese si la variabilidad baja, pero es de gran importancia que esta no aumente.

5.2.2. Usar diferentes estadísticos para el cálculo de los límites de control

Así como se usó para la prueba el estadístico de Conover, se puede hacer uso de algún otro estadístico sensible a cambios en varianzas y se comparan los desempeños de ambas cartas.

5.2.3. Considerando cambios en media y varianza simultáneos.

La carta propuesta considera que sólo existen cambios en la varianza, que la media queda constante, o si sabemos cuándo cambia la media y la magnitud del cambio podemos transformar los datos para detectar cambios en varianza. Ahora se pretende que no sea necesario tener este conocimiento para así sin importar si cambia o no la media poder detectar los cambios en la varianza, y posiblemente también se encuentren cambios en media.

5.2.4. Analizar el desempeño de \hat{t}

En los alcances de la investigación no se encuentra el análisis de este parámetro que estima el momento de cambio, se puede extender para mostrar resultados.

5.2.5. Nota para los trabajos de Zhou et al. (2009) y Hawkins & Deng (2010), para implementarlos en monitoreo de varianzas.

Los trabajos de Zhou et al. (2009) y Hawkins & Deng (2010) son muy similares a este trabajo en cuanto a metodología y supuestos, usan un estadístico de prueba no paramétrico basado en la prueba de Mann-Whitney que es una prueba para diferencias entre medias, esa es la diferencia principal con la investigación actual, la propuesta de trabajo futuro es aplicar una transformación a los datos originales para en lugar de monitorear la media se logre monitorear la varianza con el mismo estadístico de prueba de Mann-Whitney o alguno similar.

REFERENCIAS

- Amiri, & Allahyari. (2011). Change Point Estimation Methods for Control Chart Post signal Diagnostics A Literature Review. *Quality and Reliability Engineering International*.
- Ansari, A. R., & Bradley, R. A. (1960). Rank-Sum Tests for Dispersions. *The Annals of Mathematical Statistics, Institute of Mathematical Statistics*, 31(4), 1174–1189.
- Bradley, J. V. (1968). *Distribution-Free Statistical Tests*. Prentice-Hall.
- Conover, W. J. (1999). *Practical nonparametric statistics*. Wiley.
- Conover, W. J., & Iman, R. L. (1978). Some exact tables for the squared ranks test. *Communications in Statistics-Simulation and Computation*, 7(5), 491–513.
- Das, N. (2008a). A Note on the Efficiency of Nonparametric Control Chart for Monitoring Process Variability. *Economic Quality Control*, 23(1), 85–93. doi:10.1515/EQC.2008.85
- Das, N. (2008b). Non-parametric Control Chart for Controlling Variability Based on Rank Test. *Economic Quality Control*, 23(2). doi:10.1515/EQC.2008.227
- Das, N., & Bhattacharya, A. (2008). A New Non-Parametric Control Chart. *Quality Technology & Quantitative Management*, 5(4), 351–361.
- Duran, B. S., & Mielke Jr, P. W. (1968). Robustness of sum of squared ranks test. *Journal of the American Statistical Association*, 338–344.
- Fligner, M. A., & Killeen, T. J. (1976). Distribution-Free Two-Sample Tests for Scale. *Journal of the American Statistical Association*, 71(353), 210. doi:10.2307/2285771
- Hawkins, D. M. (1981). A CUSUM for a scale parameter.

- Hawkins, D. M. (1987). Self-starting CUSUM charts for location and scale. *The Statistician*, 299–316.
- Hawkins, D. M., & Deng, Q. (2010). A nonparametric change-point control chart. *Journal of Quality Technology*, 42(2), 165–173.
- Hawkins, D. M., Qiu, P., & Kang, C. W. (2003). The changepoint model for statistical process control. *Journal of quality technology*, 35(4), 355–366.
- Hawkins, D. M., & Zamba, K. D. (2005). A change-point model for a shift in variance. *Journal of Quality Technology*, 37(1), 21–31.
- James, B. R. (1967). On Pitman efficiency of some tests of scale for the gamma distribution. In *Proceedings of the Fifth Berkeley Symposium* (pp. 389–93).
- Lehmann, E. (n.d.). *Nonparametrics - Statistical Methods Based on Ranks*. Retrieved from <http://www.springer.com/statistics/statistical+theory+and+methods/book/978-0-387-35212-1>
- MacGregor, J. F., & Harris, T. J. (1993). The exponentially weighted moving variance. *Journal of Quality Technology*, 25(2). Retrieved from https://secure.asq.org/perl/msg.pl?prvurl=/data/subscriptions/jqt_open/1993/april/jqtv25i2macgregor.pdf
- Mielke, P. W. (1972). Asymptotic Behavior of Two-Sample Tests Based on Powers of Ranks for Detecting Scale and Location Alternatives. *Journal of the American Statistical Association*, 67(340), 850–854. doi:10.1080/01621459.1972.10481304

Montgomery, D. C. (2008). *Statistical quality control: a modern introduction*. Hoboken, N.J.: John Wiley.

Ross, G. J. (2013). Parametric and Nonparametric Sequential Change Detection in R: The cpm package. *Journal of Statistical Software, Forthcoming*. Retrieved from <ftp://ie.freshrpms.net/disk1/CRAN/web/packages/cpm/vignettes/cpm.pdf>

Shewhart, W. A. (1931). *Economic control of quality of manufactured product*. New York: D. Van Nostrand Company, Inc.

Taha, M. A. H. (1964). Rank test for scale parameter for asymmetrical one-sided distributions. *Publ. Inst. Statist. Univ. Paris, 13*, 169–180.

Whiteside, M., Duran, B. S., & Boullion, T. L. (1975). A comparison of some nonparametric tests for scale. *Journal of Statistical Computation and Simulation, 4(2)*, 121–132. doi:10.1080/00949657508810115

Zhou, C., Zou, C., Zhang, Y., & Wang, Z. (2009). Nonparametric control chart based on change-point model. *Statistical Papers, 50(1)*, 13–28.

CÓDIGOS DE MATLAB

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% %   Límites de Control para la varianza   % %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Funciones necesarias  %%%%%%%%%
clear all;           clc;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Datos de la simulación
replicas = 10000; p = 0.998; ni = 10; nf = 500; a = 380; colas = 1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Donde se guardará la información
tic
if a == 1
    for i = 1:replicas
        U = rand(1,ni) ;
        T(i) = Estimador_EqualV(U,colas);
    end
    h(1) = quantile(T,p);
end
for m = a:nf-ni          % para cada dato extra se obtiene un cuantil h
    for j = 1:replicas  % para cada replica se obtiene un Estimador T
        T(j)= NuevoEstimadorIC(ni,m,colas,h);
    end
    h(m+1,1) = quantile(T,p);
end
Tiempo = toc

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% %           Estimador Equal Variances           % %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Funciones necesarias %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function T2 = Estimador_EqualV(U,colas)
% rand('seed',1); % colas = 1; % U = rand(1,100);
t = length(U);
c1 = sqrt( 5/((t+1)*(2*t+1)*(8*t+11)) );
c2 = (2*t+1)*(t+1);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mm = mean(U) ;
x = abs(U - mm);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% % x = Estandarizar(U);
Rangos = FractionalRankings(x);
s = Rangos.^2; S = cumsum(s);

switch colas
case 1
T2(1,1) = -1000;
case 2
T2(1,1) = -1000;
T2(1,2) = 1000;
end

for n=2:t-2
m=t-n;
T = S(n); %cumsum(Rangos.^2);
Ti =c1*(6*T - n*c2)/(sqrt(n*m));
switch colas
case 2 % disp('colas = 2')
if Ti > T2(1,1) % Máximo
T2(1,1) = Ti;
end
if Ti < T2(1,2) % Minimo
T2(1,2) = Ti;
end
case 1 % disp('colas = 1')
if abs(Ti) > T2(1,1) % |Máximo|
T2(1,1) = abs(Ti);
end
end
end
end % Function

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% %   Límites de Control para la varianza   % %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Funciones necesarias %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all; close all; clc; % hold all;
ni = 10; replicas = 5; j=1; Falarm = 1;
sigma = 1; mu = 0; n = 1; % n = subgrupos;
ALPHA = .002; % [.05 .02 .01 .005 .002 .001];
SIGMA2 = [1.5625 1.9531]; % [0.5120 0.6400 0.8000 1.2500
1.5625 1.9531];
TAO = [0 50 250];
nn = numel(TAO);
mm = numel(SIGMA2);
kk = numel(ALPHA);
Est = 'EQ';
[stat, struc] = fileattrib;
CarpetaRaiz = struc.Name;
Fecha = clock;
for ia = 1:kk
    alpha = ALPHA(ia);
    for it = 1: nn
        tao = TAO(it);
        for s2 = 1:mm
            j=1; %
            sigma2 = SIGMA2(s2);
            tic
            while j <= replicas % Para una réplica se calcula el T
                i = ni; % Índice igual al número inicial IC
                x = sigma*subgrupos(ni-1,n)+mu; % Vector (ni-1) IC IC
                if tao == 0 % Si no hay cambio i.e. IC
                    while i ~= 0 % Mientras este IC [i >= ni]
                        x(1,i) = sigma*subgrupos(1,n)+mu; % nueva obs IC
                        T(i) = Estimador_EqualV(x(1,1:i),1); % Valor del est.
                        H(i) = TabRegEQ10(i,alpha); % L.C.
                        E(i) = T(i)-H(i); % Estimador - L.C.
                        if E(i) > 0
                            taoest(j) = i;
                            j= j+1;
                            i = 0;
                        else i = i+1;% % Bajo control
                        end%
                    end
                end

            else % if tao
                while i ~= 0
                    aux1 = i-tao;
                    if aux1 >= 0
                        sigma = sigma2; % miu = miu2;
                    else sigma = 1;
                    end
                    x(1,i) = sigma*subgrupos(1,n)+mu;
                    T(i) = Estimador_EqualV(x(1,1:i),1); % Valor del est.
                    H(i) = TabRegEQ10(i,alpha); % LC
                end
            end
        end
    end
end

```

```

        E(i) = T(i)-H(i);
        if E(i) > 0
            if aux1 >= 0
                taoest(j) = i
                j= j+1;
            end
            i = 0;           % Fuera de control
        else i = i+1;       % Bajo control
        end% if E(i)
    end % while i~=0
end % if tao
clear T x
end % end while replicas
Tiempo = toc;
ARL = mean(taoest)- tao;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Guardar los resultados de la simulación
    end %for it
end %for ia
Tiempo = toc;
end %for s2

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% % Límites de Control para la varianza Gamma % %
clear all; close all; clc; % hold all;
ni = 10; replicas = 1; mu = .5 ; a = .5;
ALPHA = .001; % [.001 .01]; % [.05 .02 .01 .005 .002 .001];
SIGMA2 = .8; % [0.5120 0.6400 0.8000 1.2500 1.5625 1.9531];
% s2 = 1:7
TAO =0;% [0 20 50 80 100 250]; % it = 1:6
nn = numel(TAO);
mm = numel(SIGMA2);
kk = numel(ALPHA);
[stat,estruc] = fileattrib;
Carpeta = struc.Name
Fecha = clock;
for ia = 1:kk
    alpha = ALPHA(ia);
    Archivo =
strcat('EEGamma', '13', '_', num2str(Fecha(2)), '_', num2str(Fecha(3)), '_', num
2str(Fecha(4)), '-', num2str(alpha), '.xls');
    for it = 1: nn
        tao = TAO(it);
        for s2 = 1:mm
            j=1; %
            sigma2 = SIGMA2(s2); sigma = 1;
            tic
            while j <= replicas
                i = ni;
                x = gamrnd(.5,1,1,ni-1
            if tao == 0 % Si no hay cambio i.e. IC
                while i ~= 0 % Mientras este IC [i >= ni]

```



```

x(1,i) = gamrnd(.5,1);% nueva observación IC
T(i) = Estimador_EqualV(x(1,1:i),1
H(i) = TabRegEQ10(i,alpha);    % L.C.
E(i) = T(i)-H(i);              % Estimador - L.C.
if E(i) > 0
    taoest(j) = i;
    j= j+1;
    i = 0;
else    i = i+1;                % Bajo control
end%
end

else                            % if tao
while i ~= 0
    aux1 = i-tao;
    if aux1 >= 0
        k = sigma2;
        g = .5-a*k;
    else g=0; k = 1;
    end
    x(1,i) = gamrnd(.5,k)+g;
    T(i) = Estimador_EqualV(x(1,1:i),1
    H(i) = TabRegEQ10(i,alpha);
    E(i) = T(i)-H(i);
    if E(i) > 0
        if aux1 >= 0
            taoest(j) = i;
            j= j+1;
        end
        i = 0;                  % Fuera de control
    else    i = i+1;            % Bajo control
    end% if E(i)
end % while i~=0
end % if tao
clear T x
end % end while replicas
Tiempo = toc;
ARL = mean(taoest)- tao;
fprintf('Tiempo = %d ARL = %d\n',Tiempo, ARL);
%%%%%%%%%%%%%% Guardar los resultados de la simulación
end %for it
end %for ia
Tiempo = toc;
end %for s2

```