

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICAS**

PROYECTO DE TITULACIÓN

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

**“MAGÍSTER EN ESTADÍSTICA CON MENCIÓN EN GESTIÓN DE
LA CALIDAD Y PRODUCTIVIDAD”**

TEMA:

DISEÑO DE GRÁFICO DE CONTROL MULTIVARIANTE CON
EFECTO MEMORIA PARA ANÁLISIS DE VARIABLES NORMALES

AUTOR:

CHRISTIAN RAFAEL NARANJO MENDOZA

Guayaquil - Ecuador

2019

DEDICATORIA

A mis padres, Fabiola y Gonzalo, que nunca pierden la fe en mí y atesoran cada uno de mis logros. No hay nada más satisfactorio que saber que estarán orgullosos de mí.

AGRADECIMIENTO

A mis padres y hermanas por apoyarme siempre y motivarme a culminar este ciclo de la mejor manera.

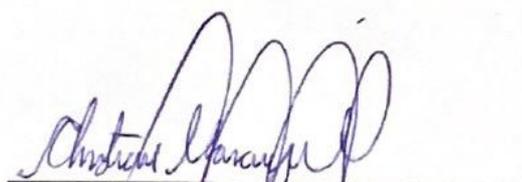
A mi directora de proyecto, Sandra García, por su paciencia y dedicación.

A mis profesores por ser inspiración y compartir sus conocimientos y experiencias desinteresadamente.

Y a mis compañeros de maestría por hacer de estos dos años de estudio un viaje lleno de sonrisas.

DECLARACIÓN EXPRESA

La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Graduación me corresponde exclusivamente; el patrimonio intelectual del mismo, corresponde exclusivamente a la **Facultad de Ciencias Naturales y Matemáticas, Departamento de Matemáticas** de la Escuela Superior Politécnica del Litoral.



Christian Rafael Maranjo Mendoza

TRIBUNAL DE GRADUACIÓN



M.Sc. Francisco Moreira Villegas

Presidente



Ph.D Sandra García Bustos

Director



Ph.D Omar Ruiz Barzola

Vocal

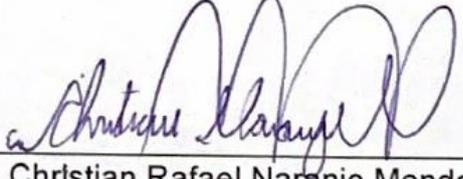


Ph.D María Nela Pastuizaca

Fernández

Vocal

AUTOR DEL PROYECTO



Christian Rafael Naranjo Mendoza

ÍNDICE GENERAL

| | |
|--|------|
| ÍNDICE GENERAL..... | vi |
| ÍNDICE DE TABLAS | viii |
| ÍNDICE DE GRÁFICOS | ix |
| ABREVIATURAS O SIGLAS..... | x |
| PRESENTACIÓN..... | xi |
| CAPÍTULO I | 1 |
| 1. ANTECEDENTES, PROBLEMA, OBJETIVO Y METODOLOGÍA | 1 |
| 1.1. ANTECEDENTES | 1 |
| 1.2. PLANTEAMIENTO DEL PROBLEMA | 3 |
| 1.3. JUSTIFICACIÓN | 4 |
| 1.4. ALCANCE | 5 |
| 1.5. OBJETIVOS DEL PROYECTO | 5 |
| 1.6. METODOLOGÍA..... | 6 |
| CAPÍTULO II | 8 |
| 2. MARCO TEÓRICO | 8 |
| 2.1. PRUEBAS DE HIPÓTESIS | 8 |
| 2.1.1. Errores en las pruebas de hipótesis..... | 8 |
| 2.2. ARL: MEDIDA DE RENDIMIENTO DE LOS GRÁFICOS DE CONTROL | 9 |
| 2.3. GRÁFICO DE CONTROL MULTIVARIADO T^2 DE HOTELLING | 10 |
| 2.4. GRÁFICO DE CONTROL EWMA | 12 |
| 2.5. CADENAS DE MARKOV Y SU APLICACIÓN PARA EVALUAR EL RENDIMIENTO DE LOS GRÁFICOS DE CONTROL..... | 14 |
| 2.5.1. Proceso Estocástico..... | 14 |
| 2.5.1. Cadenas de Markov | 15 |
| 2.5.2. Tipos de estados en una cadena de Markov | 16 |
| 2.5.3. Cadenas de Markov y Cartas de Control | 16 |
| 2.6. ALGORITMOS GENÉTICOS | 17 |
| 2.7. ESTADO DEL ARTE | 22 |
| CAPÍTULO III | 25 |
| 3. METODOLOGÍA DEL DISEÑO..... | 25 |
| 3.1. GRÁFICO DE CONTROL EWMA(T^2) | 25 |
| 3.2. DESEMPEÑO DEL GRÁFICO DE CONTROL EWMA(T^2)..... | 26 |

| | |
|---|----|
| 3.3. OPTIMIZACIÓN DEL GRÁFICO DE CONTROL EWMA(T^2)..... | 31 |
| 3.4. INTERFAZ GRÁFICA DEL OPTIMIZADOR | 32 |
| 3.5. EJEMPLO DE APLICACIÓN..... | 33 |
| CAPÍTULO IV | 36 |
| 4. RESULTADOS..... | 36 |
| 4.1. DESEMPEÑO DEL GRÁFICO EWMA(T^2) | 36 |
| 4.2. ANÁLISIS DE SENSIBILIDAD DEL GRÁFICO EWMA(T^2) | 40 |
| CAPÍTULO V | 44 |
| 5. CONCLUSIONES Y RECOMENDACIONES | 44 |
| 5.1. Conclusiones..... | 44 |
| 5.2. Recomendaciones..... | 45 |
| BIBLIOGRAFÍA | 46 |
| APÉNDICES | 49 |
| APÉNDICE A: Desempeño de la carta de control EWMA(T^2)..... | 49 |

ÍNDICE DE TABLAS

| | |
|--|----|
| Tabla 1: Tipos de errores en una prueba de hipótesis | 9 |
| Tabla 2: Escenarios de análisis del desempeño de la gráfica propuesta EWMA(T^2) | 37 |
| Tabla 3: Desempeño de la carta de controla EWMA(T^2) con $ARL_0=200$ y $p=2.38$ | |
| Tabla 4: Desempeño de la carta de controla EWMA(T^2) con $ARL_0=200$ y $p=4.39$ | |
| Tabla 5: Desempeño de la carta de controla EWMA(T^2) con $ARL_0=500$ y $p=10$ | 39 |
| Tabla 6: Casos definidos para el análisis de sensibilidad | 41 |
| Tabla 7: ARL fuera de control para el gráfico EWMA(T^2) optimizados para los cambios 0,50 y 3,00. | 41 |
| Tabla 8: ARL fuera de control para el gráfico EWMA(T^2) optimizados para los cambios 0,50 y 0,80. | 42 |
| Tabla 9: ARL fuera de control para el gráfico EWMA(T^2) optimizados para los cambios 3,00 y 3,50. | 42 |
| Tabla 10: Desempeño de la carta de controla EWMA(T^2) con $ARL_0=200$ y $p=6$ | 49 |
| Tabla 11: Desempeño de la carta de controla EWMA(T^2) con $ARL_0=200$ y $p=10$ | 49 |
| Tabla 12: Desempeño de la carta de controla EWMA(T^2) con $ARL_0=500$ y $p=4$ | 49 |

ÍNDICE DE GRÁFICOS

| | |
|---|----|
| Gráfico 1: Fases para la ejecución de un algoritmo genético..... | 20 |
| Gráfico 2: Diagrama de probabilidades de transición q_{ij} | 29 |
| Gráfico 3: Captura de pantalla del programa “Aplicación EWMA T^2 ” | 33 |
| Gráfico 4: Salida del software para el gráfico de control EWMA(T^2) optimizado para el ejemplo de aplicación propuesto..... | 35 |

ABREVIATURAS O SIGLAS

| | |
|------------------|---|
| ARL | Average Run Length. Longitud promedio de rachas |
| ARL ₀ | ARL bajo control deseado |
| ARL ₁ | ARL fuera de control |
| LCS | Límite de control superior de la carta de control |
| LCI | Límite de control inferior de la carta de control |
| GA | Algoritmos Genéticos |
| ARL(d=0) | ARL bajo control obtenido por el software |
| ARL(d=d*) | ARL fuera de control obtenida por el software para un cambio d* |

PRESENTACIÓN

En un ambiente competitivo como el que viven las industrias hoy en día en el Ecuador y en el mundo, una de las vías para mantener clientes y diferenciarse de la competencia es a través de la calidad de sus procesos y productos. Entre las herramientas y metodologías para garantizar la calidad de los procesos en las organizaciones se encuentra el control estadístico de calidad con técnicas como las cartas o gráficos de control que nos ayudan a detectar cuando un proceso está arrojando resultados diferentes de los esperados por la organización.

Desde su introducción en 1924 por Walter A. Shewhart los gráficos de control han sido de interés de muchos investigadores y han evolucionado para ajustarse a la realidad de los procesos, desde el control univariante hasta el control multivariante que es el análisis de varias variables de calidad a la vez y correlacionadas entre ellas.

Dentro de los gráficos de control multivariante se encuentra el gráfico T^2 de Hotelling que es justamente el objeto de estudio de este proyecto en donde buscamos desarrollar una nueva gráfica de control que combina el aspecto multivariante del gráfico T^2 de Hotelling con la capacidad para detectar cambios pequeños del gráfico EWMA, dando como resultado un gráfico de control multivariante, al que se llamó EWMA(T^2), y que presenta mejor desempeño que el gráfico T^2 de Hotelling.

CAPÍTULO I

1. ANTECEDENTES, PROBLEMA, OBJETIVO Y METODOLOGÍA

1.1. ANTECEDENTES

Con la Revolución Industrial y el avance tecnológico al inicio del siglo XX surgió la línea de ensamblaje y con ello la producción a bajo costo. La mayor demanda de productos hizo que la inspección total de la calidad de la producción no fuese posible debido a los altos costos en tiempo y dinero.

Bajo esta situación, en 1924, el matemático estadounidense Walter A. Shewhart introdujo el control estadístico de calidad, creando la base de los ahora conocidos gráficos de control. Su principal motivación era determinar cuándo un proceso presentaba variaciones normales y propias (es decir que se encontraba bajo control) y cuándo presentaba variaciones especiales que no eran propias de él (es decir que se encontraba fuera de control); con el fin de poder detectar posibles fallas en el producto y corregirlas antes de que el mismo sea entregado al cliente final.

El principal objetivo de un gráfico de control es monitorear y analizar el comportamiento de un proceso a lo largo del tiempo haciendo posible distinguir entre variaciones por causas comunes y variaciones por causas especiales (atribuibles), esto los ha constituido como una de las siete herramientas de la calidad y, por su aplicación en diferentes industrias, se han convertido actualmente en un importante y amplio campo de investigación. Siendo los gráficos de control multivariantes una de las áreas más activas en los estudios sobre control estadístico de calidad.

Los primeros gráficos de control propuestos por Shewhart (1924) para monitorear variables cuantitativas utilizaban como estadísticos: el promedio \bar{X} y la variación (R (Rango) o S (Desviación estándar)). Posteriormente, se desarrollaron gráficos para controlar variables cualitativas, como los gráficos c, u, p y np para monitorizar el número de defectos por lote, la proporción de defectos por lote, la proporción de unidades no conformes por lotes y el número de unidades defectuosas por lote, respectivamente (Alt, 1985). Estos gráficos detectan eficientemente grandes cambios en el proceso.

A fin de detectar pequeños cambios en el proceso, dos nuevos gráficos univariados surgieron, los cuales consideran toda la información muestral y no sólo el último valor como lo hacen los gráficos Shewhart:

- En 1954, E. S. Page diseñó un gráfico de control C CUSUM (Cumulative Sum Control Chart), basado en las sumas acumuladas de las desviaciones de los valores muestrales respecto de un valor objetivo (Montgomery, 2006)
- Un segundo gráfico fue introducido por S. W. Roberts en 1959, basado en los promedios móviles exponencialmente ponderados (EWMA (Exponentially Weighted Moving Average Chart)) (Montgomery, 2006).

Pronto nació la necesidad de controlar simultáneamente más de una característica de calidad en ciertos procesos, de esta manera surge el control de calidad multivariante o monitoreo del proceso (Umit & Cigdem, 2001). El primer estudio en control de calidad multivariante fue introducido por Hotelling (Hotelling, 1947). Tres de las gráficas de control multivariantes más populares son la T^2 de Hotelling, la gráfica MEWMA (Multivariate Exponentially-Weighted Moving Average) y la gráfica MCUSUM (Multivariate Cumulative Sum) (Montgomery, 2006).

Hoy en día, las computadoras modernas han permitido la ejecución de cálculos complejo durante la última década y las gráficas de control multivariantes han

ganado mayor atención. Diversos estudios se han desarrollado entre los cuáles se puede mencionar: “*Control chart for multivariate attribute processes*” (Lu, Xie, & Goh, 1998), “*Monitoreo y Control de un Proceso Normal Multivariado*” (Contreras, Marroquin, & Zertuche, 2011), “*Optimum Multiple and Multivariate Poisson Statistical Control Charts*” (Aparisi, García Bustos, & Epprecht, Optimum Multiple and Multivariate Poisson Statistical Control Charts, 2013), “*Optimal EWMA of Linear Combination of Poisson Variables for Multivariate Statistical Process Control*” (García Bustos, Aparisi, & Epprecht, 2014), “*Sistema de control multivariado no paramétrico de procesos*” (Zertuche Luis, Cantú Sifuentes, & Piña Monarrez, 2007), “*Control charts with variable dimension for linear combination of Poisson variables*” (García-Bustos, Mite, & Vera, 2016), “*One-sided cumulative sum control chart for monitoring shift in the scale parameter delta (δ), of the new Weibull-Pareto distribution*” (Baba, Jabir, & Abdul-Salam, 2017), entre otros.

1.2. PLANTEAMIENTO DEL PROBLEMA

En nuestro país varias industrias utilizan gráficos de control univariantes para controlar las salidas de sus procesos siendo los más utilizados, sin duda, los gráficos Shewhart. Sin embargo, existen procesos en los que intervienen muchas variables y además correlacionadas lo que hace necesario plantearse el uso de gráficos de control multivariantes con el fin de disminuir la probabilidad de tener falsas alarmas; es decir, que el gráfico nos muestre un punto fuera de control cuando en realidad el proceso está bajo control.

Uno de los gráficos de control multivariantes más usados y más investigados es el gráfico T^2 de Hotelling que presenta una gran potencia a la hora de detectar cambios en las características de calidad que se desean controlar. Sin embargo este gráfico tiene un problema y es que no es del todo potente para detectar cambios pequeños en el proceso.

Por ese motivo, en este proyecto se desea proponer una mejora al gráfico de control T^2 de Hotelling y desarrollar un gráfico que combine la capacidad de detección de cambios pequeños del gráfico $EWMA$ con el esquema multivariante del gráfico T^2 de Hotelling, es decir crear un nuevo gráfico de control que sea equivalente al $EWMA$ de un T^2 .

1.3. JUSTIFICACIÓN

Hoy en día es conocido que existen múltiples variables que definen la calidad de un proceso. Esto, sumado al rápido crecimiento de la tecnología para analizar datos y a la continua modernización y automatización en el monitoreo de los procesos han ocasionado que el interés en estudiar el comportamiento simultáneo de varias características de calidad (o variables del proceso) vaya en aumento.

Además de estudiar el comportamiento de varias variables al mismo tiempo, es importante también para las industrias detectar de forma oportuna cuando sus procesos están arrojando salidas fuera de los criterios de control establecidos, esto con el fin de tomar acciones oportunas con el mínimo de impacto para la organización.

Por facilidad, o desconocimiento, muchas industrias utilizan cartas de control múltiples para controlar la mayoría de sus procesos; es decir, cartas de control independientes para cada una de las variables que desean controlar lo que puede conllevar algunos problemas en el caso de que las variables monitoreadas se encuentren correlacionadas.

Como alternativa al esquema múltiple nace el esquema multivariado siendo la carta de control T^2 de Hotelling una de las más utilizadas, pero que presenta un problema al momento de detectar cambios pequeños (como se expuso en la sección de planteamiento del problema).

Todo lo antes expuesto motivó el desarrollo del presente proyecto, el interés de construir un gráfico de control multivariado para variables continuas que tenga como finalidad detectar rápidamente fluctuaciones anormales en un proceso y que sea una modificación del gráfico T^2 de Hotelling.

1.4. ALCANCE

El presente proyecto de titulación busca desarrollar un nuevo gráfico de control con efecto memoria para un conjunto de variables normales multivariadas. A través del: Diseño de un estadístico de prueba con efecto memoria, Construcción de una gráfica de control basada en este estadístico y el desarrollo de un software para obtener los parámetros de la carta de control optimizada, siendo estos parámetros el límite de control superior (*UCL* por sus siglas en inglés) y r el peso dado a la observación actual.

Además, este proyecto forma parte de una investigación en gráficos de control univariantes y multivariantes dirigida por la Dra. Sandra García.

1.5. OBJETIVOS DEL PROYECTO

El objetivo principal del proyecto es diseñar una carta de control multivariante con efecto memoria para análisis de variables normales que presente un mejor rendimiento que otros gráficos de control existentes. A partir de este objetivo general se plantean los siguientes objetivos específicos:

- Diseñar un estadístico de prueba con efecto memoria y usarlo en un nuevo gráfico de control.
- Construir una gráfica de control optimizada basada en el estadístico de prueba diseñado que presente un buen rendimiento.

- Desarrollar un software para obtener los parámetros de la carta de control optimizada.
- Comparar el desempeño de la carta de control diseñada con metodologías existentes a fin de que el usuario pueda seleccionar una alternativa adecuada para el control de un proceso en particular.

1.6. METODOLOGÍA

Para diseñar y construir gráficos o cartas de control se debe conocer la naturaleza de la variable a monitorear, luego pensar en estadísticos de control o de prueba cuyas funciones de probabilidad puedan obtenerse o estimarse; y, finalmente, calcular los límites de control para el gráfico diseñado, haciendo uso de algoritmo genético, que permitan lograr una probabilidad de falsa alarma deseada.

La probabilidad de falsa alarma viene a ser la probabilidad del error tipo I (α) de un contraste de hipótesis que es la probabilidad de que en el gráfico de control se presente una señal de que el proceso está fuera de control dado que en realidad está bajo control.

Es deseable también que los gráficos de control detecten de forma rápida cualquier cambio en el proceso, esto está relacionado con el concepto de potencia de la prueba, es decir con la probabilidad de no cometer el error tipo II (β) que es la probabilidad de no rechazar la hipótesis nula dado que es falsa. Un diseño óptimo de un gráfico de control sería aquel que tenga una tasa baja de falsas alarmas y una detección rápida de cambios en el proceso.

Para medir el rendimiento de los gráficos de control se utiliza el *ARL* (Average Run Length), que es el número promedio de muestras requeridas hasta que el gráfico de control presente una señal. Si el proceso se encuentra bajo control se denota como ARL_0 y si el proceso está fuera de control se denota como

ARL_1 . Un diseño óptimo de gráficos de control implica un valor grande de ARL_0 y un valor pequeño de ARL_1 .

En el presente proyecto para diseñar este gráfico y encontrar los parámetros del mismo se utilizará la técnica de algoritmo genético la cual evalúa diferentes soluciones hasta encontrar la óptima o la más cercana a ella, en función a un criterio conocido como “medida de aptitud”.

Para el desarrollo de este optimizador se utilizar el software libre R junto con una interfaz gráfica amigable al usuario desarrollada en Visual Studio 2017, de tal manera que sólo se ingresen los parámetros de entrada y se calculen automáticamente los parámetros necesarios.

Una vez construido el gráfico, se podrán identificar, analizar e interpretar señales cuando el estadístico de prueba presente valores fuera de los límites de control y se podrá realizar comparaciones con metodologías ya existentes para comprobar que la planteada tiene mejor desempeño.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. PRUEBAS DE HIPÓTESIS

Se utiliza el término “hipótesis” para hacer referencia a los supuestos que hacemos sobre un parámetro estadístico (media, proporción, entre otros) de una determinada población. Y, con “Prueba de Hipótesis” se entiende al procedimiento estadístico mediante el cual se determina si la hipótesis que se tiene sobre el parámetro de la población es comprobable con lo que se observa al tomar una muestra de dicha población.

Los elementos de una prueba estadística son:

- Hipótesis nula (H_0)
- Hipótesis alternativa (H_1)
- Estadístico de prueba
- Región de rechazo

La hipótesis alternativa es normalmente la hipótesis que se desea apoyar con base a la información contenida en la muestra.

2.1.1. Errores en las pruebas de hipótesis

Luego de efectuada la prueba de hipótesis el investigador habrá escogido alguna de las dos hipótesis (nula o alternativa), y aquella hipótesis escogida puede coincidir o no con la que en realidad es cierta dando lugar a que se presente alguno de los siguientes errores:

- Error Tipo I: Que se presenta con probabilidad α y es conocido también como falso positivo. Es el error que se comete cuando el investigador rechaza H_0 dado que esta es verdadera en la población.
- Error Tipo 2: Que se presenta con probabilidad β y es conocido también como falso negativo. Se comete cuando el investigador no rechaza H_0 y esta resulta ser falsa en la población.

En la siguiente tabla se resume lo explicado:

| Decisión | Eventos | |
|-------------------|--|---|
| | H_0 es falsa | H_0 es verdadera |
| Rechazar H_0 | Correcto | Error Tipo I $P(\text{Error Tipo I}) = \alpha$ |
| No Rechazar H_0 | Error Tipo II $P(\text{Error Tipo II}) = \beta$ | Correcto |

Tabla 1: Tipos de errores en una prueba de hipótesis

Fuente: Elaboración propia

2.2. ARL: MEDIDA DE RENDIMIENTO DE LOS GRÁFICOS DE CONTROL

El ARL (Average Run Length) es el número promedio de puntos que se deben graficar en una carta de control antes de que un punto indique una salida de control (Montgomery, 2006). Si las observaciones del proceso no se encuentran correlacionadas, entonces para cualquier gráfico de control de tipo Shewhart el ARL puede ser calculado fácilmente como la media de una variable aleatoria geométrica usando la siguiente fórmula:

$$ARL = \frac{1}{p}$$

En donde p es la probabilidad de que cualquier punto esté por fuera de los límites de control.

Se identifican dos tipos de ARL:

- ARL bajo control (ARL_0): Que es el número esperado de puntos o muestras hasta que el gráfico de control dé una señal de fuera de control, bajo la condición de que el proceso actual se encuentra bajo control. Se calcula como $1/\alpha$ siendo α la probabilidad de cometer el error tipo I.
- ARL fuera de control (ARL_1): Que es el número esperado de puntos o muestras hasta que el gráfico de control dé una señal de fuera de control, bajo la condición de que el proceso actual se encuentra fuera de control. Se calcula como $1/(1 - \beta)$ siendo β la probabilidad de cometer el error tipo II.

Un diseño óptimo de una carta de control es aquel en el que el valor de ARL_0 es lo más alto posible y el valor de ARL_1 es lo más pequeño posible garantizando así que detecte de forma rápida los cambios del proceso.

2.3. GRÁFICO DE CONTROL MULTIVARIADO T^2 DE HOTELLING

En 1936 Mahalanobis introdujo una expresión matemática para medir la distancia entre dos variables aleatorias (X, Y) de dimensión p con la misma función de distribución de probabilidades y matriz de varianzas y covarianzas Σ (Mahalanobis, 1936). Esta expresión estaba definida de la siguiente manera:

$$d_m(X, Y) = \sqrt{(X - Y)^T \Sigma^{-1} (X - Y)}$$

Siendo $X^T = (X_1, X_2, \dots, X_p)$ un conjunto de variables con un vector de medias $\mu^T = (\mu_1, \mu_2, \dots, \mu_p)$ y matriz de varianzas y covarianzas Σ (de dimensión $p \times p$)

simétrica y definida positiva, el cuadrado de la distancia de Mahalanobis estandarizada de X a μ estaría dado por:

$$d_m^2(X, Y) = (X - \mu)^T \Sigma^{-1} (X - \mu)$$

Dado el caso de que X provenga de una distribución normal multivariante con media μ y matriz de varianzas y covarianzas Σ , se define al estadístico χ_p^2 con una distribución chi-cuadrada con p grados de libertad de la siguiente manera:

$$\chi_p^2 = n(\bar{X} - \mu)^T \Sigma^{-1} (\bar{X} - \mu)$$

En donde \bar{X} es el vector de medias de la muestra.

Ya que por lo general no se conocen los parámetros del proceso hay que estimarlos, y para ello Harold Hotelling estableció el estadístico T^2 que se utiliza para un conjunto de variables que siguen una distribución normal multivariada y que se define de la siguiente manera:

$$T^2 = n(\bar{X} - \hat{\mu})^T S^{-1} (\bar{X} - \hat{\mu})$$

Donde S y $\hat{\mu}$ son los estimadores de la matriz de varianzas y covarianzas y del vector de medias respectivamente. Esta estimación se la realiza con m muestras de tamaño n y, si m tiende al infinito, la distribución del estadístico T^2 converge a una distribución Chi-cuadrado con p grados de libertad (Seber, 1984).

El gráfico de control del estadístico T^2 tiene como límite de control superior $LCS = \chi_{\alpha, p}^2$ y como límite de control inferior $LCI = 0$, siendo α el nivel de significancia y p el número de variables de calidad que serán controladas.

Cuando el proceso ha sufrido una desviación en al menos uno de los promedios de sus variables, entonces el vector μ_1 se aleja del vector de

promedios del proceso bajo control μ_0 en una distancia d de Mahalanobis. Entonces, el estadístico T^2 pasa a tener una distribución chi-cuadrado no centralizada con p grados de libertad y con parámetro de descentralizado definido por la siguiente ecuación:

$$\lambda = nd^2 = n(\mu_1 - \mu_0)^T \Sigma^{-1}(\mu_1 - \mu_0)$$

En donde n es el tamaño de la muestra y Σ es a matriz de varianzas y covarianzas de X cuando el proceso se encuentra bajo control. El ARL de este tipo de gráfico es calculado de la siguiente manera:

$$ARL = \frac{1}{1 - P(T^2 < LCS)}$$

En donde LCS es el límite de control superior. En el caso de que el estadístico caiga exactamente en el límite, entonces existe una señal de que el proceso pudiera estar fuera de control.

2.4. GRÁFICO DE CONTROL EWMA

El gráfico de control de medias móviles ponderadas exponencialmente (EWMA, por sus siglas en inglés) es una buena alternativa a los gráficos de control de Shewhart cuando estamos interesados en detectar cambios pequeños (Montgomery, 2006). Se utiliza usualmente sobre observaciones individuales.

Fue introducido en 1956 por Roberts, y consiste en acumular en cada período los valores de las observaciones anteriores; es decir que en cada período se representa un promedio de la observación actual y las observaciones anteriores dando mayor peso a las observaciones más recientes. Los gráficos EWMA utilizan una forma de hacer medias móviles que consiste en dar un peso a las informaciones históricas el cual decae exponencialmente con el tiempo. Esta media móvil se define de la siguiente manera:

$$y_i = \lambda x_i + (1 - \lambda)y_{i-1}$$

Donde el parámetro λ es un constante entre 0 y 1 y y_0 es el valor objetivo del proceso por lo que $y_0 = \mu_0$. En otras ocasiones y_0 también puede ser el promedio de data preliminar por lo que $y_0 = \bar{x}$ y. En cada período entonces se pondera la observación actual x_i con el valor de la media móvil anterior y_{i-1} y si sustituimos recursivamente llegaremos a la siguiente expresión:

$$y_i = \lambda x_i + (1 - \lambda)[\lambda x_{i-1} + (1 - \lambda)y_{i-2}] = \lambda x_i + \lambda(1 - \lambda)x_{i-1} + (1 - \lambda)^2 y_{i-2}$$

Si continuamos sustituyendo recursivamente los y_{i-j} en donde $j = 2, 3, \dots, t$ obtendremos que:

$$y_i = \lambda \sum_{j=0}^{i-1} (1 - \lambda)^j x_{i-j} + (1 - \lambda)^i y_0$$

Los pesos $\lambda(1 - \lambda)^j$ decrecen exponencialmente con j . Por ejemplo, si $\lambda = 0,20$ los pesos dados a las medias siguientes serían 0,16, 0,128, 0,1024, etcétera. La suma de todos estos pesos dará como resultado la unidad.

Si las observaciones x_i son variables aleatorias independientes con varianza σ^2 , entonces:

$$E(y_i) = \mu_0$$
$$Var(y_i) = \sigma^2 \left(\frac{\lambda}{2 - \lambda} \right) (1 - (1 - \lambda)^{2i})$$

Resultando los límites de control y la línea central de la siguiente manera:

$$LCS = E(y_i) + k \sqrt{Var(y_i)}$$
$$LC = E(y_i)$$
$$LCI = E(y_i) - k \sqrt{Var(y_i)}$$

Sustituyendo por los valores de la media y la varianza de y_i y para $k=3$, resulta:

$$\begin{aligned}LCS &= \mu_0 + 3\sigma \sqrt{\left(\frac{\lambda}{2-\lambda}\right) (1 - (1-\lambda)^{2i})} \\LC &= \mu_0 \\LCI &= \mu_0 - 3\sigma \sqrt{\left(\frac{\lambda}{2-\lambda}\right) (1 - (1-\lambda)^{2i})}\end{aligned}$$

Se puede apreciar que los límites de control superior e inferior varían para cada valor de i , y si i es muy elevado se tiene aproximadamente que:

$$\begin{aligned}LCS &= \mu_0 + 3\sigma \sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \\LC &= \mu_0 \\LCI &= \mu_0 - 3\sigma \sqrt{\left(\frac{\lambda}{2-\lambda}\right)}\end{aligned}$$

Al emplear este tipo de gráfico se suelen elegir valores de λ comprendidos entre 0,05 y 0,25, estando muy generalizado el uso de valor de λ de 0,05, 0,10 y 0,20. Cuanto más grande es el valor de λ menos importancia se le da a los valores más anteriores pero puesto a que la finalidad del gráfico EWMA es detectar pequeños cambios en el proceso se recomienda trabajar con valores de λ bajos.

2.5. CADENAS DE MARKOV Y SU APLICACIÓN PARA EVALUAR EL RENDIMIENTO DE LOS GRÁFICOS DE CONTROL

2.5.1. Proceso Estocástico

Previo a definir lo que es una cadena de Markov indicaremos la definición de un proceso estocástico. Un proceso estocástico es un concepto matemático utilizado para referirse a una sucesión de variables aleatorias (estocásticas) que evolucionan en función de otra variable que usualmente es el tiempo (Pérez R., 2007). Cada una de estas variables tiene su propia función de probabilidad.

2.5.1. Cadenas de Markov

Las cadenas de Markov (o modelo de Markov) fueron introducidas por el matemático ruso Andréi Márkov en 1907 y son un proceso estocástico discreto en el cual la probabilidad de ocurrencia de cada una de las variables está determinada únicamente por el comportamiento de la inmediata anterior (Gagniuc, 2017). Es decir:

$$P\{X_{t+1} = j / X_0 = i_0, X_1 = i_1, \dots, X_t = i_t\} = P\{X_{t+1} = j / X_t = i_t\} = p_{i,j}(t)$$

El dominio de la sucesión de variables $X_0, X_1, \dots, X_i, \dots, X_t$ es llamado espacio de estados $\{E_i\}$ ($i = 1, \dots, t$) esto quiere decir que el valor de X_t podría leerse como el estado del proceso en el tiempo t . También se define $p_{i,j}(t)$ como la probabilidad de que el proceso, en el tiempo t , pase del estado i al estado j en el momento siguiente. Si esta probabilidad no depende del tiempo t entonces $p_{i,j}(t) = p_{i,j}$, se dice que las probabilidades de transición son estacionarias.

La matriz de transición de un paso \mathbf{P} , está formada por las probabilidades $p_{i,j}$ de pasar de un estado a otro en un solo paso. Por otro lado, la probabilidad de transición del estado i al estado j en n pasos se denota por:

$$p_{ij}^{(n)} = P(X_n = j / X_0 = i)$$

Y son estas probabilidades las que conforman a la matriz de transición de n pasos $\mathbf{P}^{(n)}$, la cual puede calcularse multiplicando n veces la matriz \mathbf{P} o, lo que es lo mismo, como \mathbf{P}^n .

2.5.2. Tipos de estados en una cadena de Markov

Se pueden clasificar los estados de una cadena de Markov como: accesibles, comunicantes, recurrentes o sin retorno.

Un estado E_j es accesible desde el estado E_i si $p_{ij}^{(n)} > 0$, para algún valor de n . Si E_j es accesible desde E_i y viceversa se dice que estos estados están comunicados, son estados comunicantes.

Un estado E_i es recurrente, si una vez que el proceso lo alcanza, regresará a él. Un caso especial de este tipo de estados es el estado absorbente, el cual una vez que el proceso lo alcanzó es imposible salir de él.

Por último, un estado es sin retorno cuando no se comunica con ningún otro estado ni consigo mismo.

2.5.3. Cadenas de Markov y Cartas de Control

Se tiene una cadena de Markov con al menos un estado absorbente, de tal forma que la matriz de transición de un paso se puede particionar de la siguiente manera:

$$\mathbf{P} = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

Donde \mathbf{Q} y \mathbf{R} son sub matrices de \mathbf{P} . Siendo \mathbf{Q} la sub matriz de las probabilidades de ir de un estado no absorbente o transitorio a otro estado no absorbente en un paso, \mathbf{R} la sub matriz de las probabilidades de ir desde los

estados no absorbentes a los estados absorbentes en un paso, $\mathbf{0}$ la matriz nula e \mathbf{I} la matriz identidad.

Bajo estas condiciones, se puede calcular el vector de número de pasos promedio hasta que el proceso cae en un estado absorbente mediante la siguiente expresión:

$$\mathbf{T} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}$$

Donde $\mathbf{1}$ es un vector de unos. El i -ésimo elemento del vector \mathbf{T} tiene como valor el número promedio de pasos hasta que el proceso cae en un estado absorbente si se empezó en el estado i .

Este procedimiento planteado es el que se utiliza para calcular el ARL (Brook & Evans, 1972), a través la identificación de los estados no absorbentes del proceso, y considerando como estado absorbente aquel en el que el proceso está fuera de control; es decir cuando el valor del estadístico se ubica fuera de los límites de control.

Finalmente, se debe considerar también un vector \mathbf{b} de probabilidades iniciales para los estados de la cadena de Markov, quedando el valor del ARL de la siguiente manera:

$$ARL = \mathbf{b}'(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}$$

En el presente proyecto se utilizarán cadenas de Markov para calcular el valor de los ARLs de la nueva carta de control propuesta.

2.6. ALGORITMOS GENÉTICOS

En ocasiones, la obtención de los parámetros óptimos de un gráfico de control exige optimizaciones muy complicadas haciendo necesario el uso de métodos heurísticos como los algoritmos genéticos (GA por sus siglas en inglés). Los GA son un tipo de algoritmo de optimización utilizado para encontrar soluciones

óptimas a un problema computacional dado que maximiza o minimiza una función en particular. Son una rama de un amplio campo de estudio conocido como “computación evolutiva” y que imitan el proceso biológico de reproducción y selección natural para resolver problemas de optimización (Carr, 2014).

Se encuentran inspirados en el proceso de selección natural expuesto por Charles Darwin en su libro “El origen de las especies” en donde señala que las especies tiene su origen a través de la selección natural y que son las razas más fuertes las que sobreviven por sobre las más débiles, evolucionando sobre estas últimas.

A mediados del siglo XX, investigadores del área de la programación adoptaron este tipo de comportamiento y lo estudiaron, para utilizarlo como reglas empíricas-rationales y construir nuevos algoritmos. La técnica de algoritmo genético fue creada por John Henry Holland (Holland, 1975), quien desarrolló y creó varios algoritmos genéticos para su implementación en inteligencias artificiales, llegando a ser conocido como el padre de esta técnica .

Los algoritmos genéticos son métodos adaptativos que pueden utilizarse para resolver problemas de búsqueda y optimización, siendo capaces de ir creando soluciones, las mismas que irán evolucionando hacia los valores óptimos del problema planteado. Se los considera como una técnica robusta y su eficacia está probada en el caso de querer calcular funciones no derivables (o de derivación compleja) aunque su uso es posible en cualquier función. Para su aplicación es importante tener en cuenta lo siguiente:

- Si la función a optimizar tiene muchos máximos o mínimos locales, se requerirán más iteraciones del algoritmo para “asegurar” el máximo o mínimo global.

- Si la función a optimizar tiene varios puntos muy cercanos al valor óptimo, solamente se puede “asegurar” que se encontrará uno de ellos que no necesariamente será el óptimo.

Entre los elementos básicos de los algoritmos genéticos tenemos los siguientes:

- Individuo: Posible solución al problema planteado.
- Población: Conjunto de individuos o de posibles soluciones.
- Gen: Representación de un parámetro específico del individuo
- Cromosoma: Agrupación de genes de un individuo. La forma más utilizada para representar a los genes del individuo es a través de codificación binaria, la cual es superior que otros métodos de codificación existentes.

Las fases para la ejecución de un algoritmo genético son las siguientes:

- 1) Iniciación (i): Se crean individuos aleatoriamente y conjuntos de cromosomas que podrían dar solución al problema.
- 2) Evaluación (f(X)): A cada uno de los cromosomas del grupo de individuos, se le aplicará una función de actitud para determinar cuan buena es la solución que aporta el individuo.
- 3) Selección (Se): Los individuos con mejores resultados en la fase anterior tendrán mayor probabilidad de ser seleccionados. Las formas de selección más comunes son: Por ruleta, por rango y por torneo.
- 4) Cruzamiento (Cr): Constituye el principal operador a nivel genético, representa la reproducción de los individuos. Utiliza dos cromosomas a la vez para que se generen dos descendientes que combinen las

características de los cromosomas padres. Los operadores más utilizados son: Un punto, dos puntos, uniforme y aritmético.

- 5) Mutación (Mu): Como su nombre indica, se modifica al azar el cromosoma de los individuos, esto permite considerar nuevas opciones en el grupo de individuos que no están cubiertas en la población actual. Se lleva a cabo a través de una probabilidad de mutación establecida que suele ser baja para evitar hacer del algoritmo genético una búsqueda aleatoria.
- 6) Reemplazo (Re): Se seleccionan a los mejores individuos del grupo analizado, para generar el conjunto de cromosomas de la siguiente generación.

El siguiente gráfico resume las fases de un algoritmo genético.

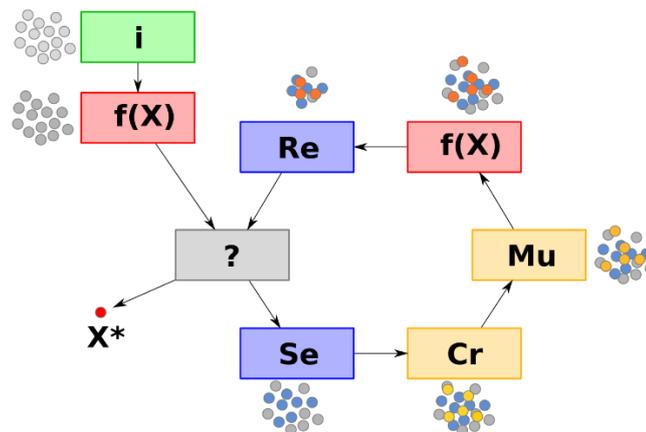


Gráfico 1: Fases para la ejecución de un algoritmo genético

Fuente: Wikipedia

Para diseñar el gráfico de control planteado en este proyecto se utilizaron algoritmos genéticos para encontrar los parámetros óptimos que minimicen el ARL fuera de control. La interfaz de usuario diseñada contempla los métodos de selección de torneo y ruleta; tipo de cruzamiento de un punto, dos puntos y uniforme con una probabilidad de ocurrencia de 0,99 y una probabilidad de mutación de 0,03.

Sobre los métodos de selección cabe mencionar lo siguiente: En la selección por torneo se escoge al azar un determinado número de individuos para reproducir al más apto; y en la selección por ruleta se crea una ruleta con los cromosomas de cada generación, aquellos cromosomas con mayor puntuación tendrán mayor probabilidad de ser seleccionados.

En cuanto al tipo de cruzamiento: En el cruzamiento de 1 punto, se cortan a los dos cromosomas en un punto específico, el nuevo individuo tendrá como cromosoma la primera parte de uno de los progenitores y la segunda parte del otro. En el cruzamiento de 2 puntos se corta a los dos cromosomas en dos puntos, un progenitor aporta dos partes (la inicial y la final) y el otro aporta solo una parte (la media). En el cruzamiento uniforme, en cambio, los padres aportan aleatoriamente cada uno de los genes a su descendiente.

Finalmente, como funciones de aptitud se utilizan las recomendadas por Aparisi y García-Díaz (Aparisi & García-Díaz, 2007) quienes aplicaron el concepto de diseño óptimo sobre un punto, que consiste en especificar el ARL_0 deseado y un vector de cambios d^* que se desea detectar. Por lo que se obtiene la combinación de parámetros del gráfico que minimiza el ARL fuera de control $ARL(d = d^*)$ restringido al valor ARL_0 deseado.

Estas funciones de aptitud son:

1. $\max - w_1 |ARL_0 - ARL(d = 0)| - w_2 ARL(d = d^*)$ si $(ARL_0 - ARL(d = 0)) \leq 0$,
2. $\max - 2w_1 |ARL_0 - ARL(d = 0)| - w_2 ARL(d = d^*)$ si $(ARL_0 - ARL(d = 0)) > 0$,

En donde ARL_0 representa el ARL deseado, $ARL(d = 0)$ representa el ARL bajo control que se obtiene con los parámetros del gráfico evaluados, $ARL(d = d^*)$ es el valor del ARL para detectar el cambio d^* , en el cual se optimiza el gráfico; y w_1, w_2 representan las ponderaciones de cada objetivo a optimizar en cada una de las ecuaciones. En nuestro programa estos pesos han sido

establecidos como 1 y 5 respectivamente siguiendo la recomendación de Aparisi y García-Díaz en trabajos anteriores.

2.7. ESTADO DEL ARTE

Como se mencionó en el capítulo 1 del presente proyecto muchos son los trabajos e investigaciones realizados sobre la carta de control T^2 de Hotelling y en esta sección mencionaremos las más recientes.

Vanhatalo, E. y Kulahci M. estudiaron el efecto de la autocorrelación en el gráfico T^2 de Hotelling (Vanhatalo & Kulahci, 2015), afirmando que en muchos casos la data que recolectamos de los procesos se encuentra correlacionada en el tiempo debido a factores como la frecuencia de muestreo y la propia dinámica del proceso; y, esta autocorrelación puede llegar a afectar la probabilidad de falsa alarma y la habilidad del gráfico de control para detectar cambios. En su estudio probaron que el ignorar las autocorrelaciones puede llevar a conclusiones erróneas, a que el ARL bajo control difiera significativamente del nominal y a una pobre habilidad para detectar cambios.

En el estudio llevado a cabo por Mostajeran A., Iranpanah N. y Noorossana R. se habla sobre el supuesto de normalidad que rige a la mayoría de los gráficos de control (Mostajeran , Iranpanah, & Noorossana R., 2016). Un investigador debería esperar resultados equivocados si este supuesto llegara a no cumplirse y que, para evitar esto, es necesario evaluarlo antes de aplicar un gráfico de control determinado. Sin embargo, hay ocasiones en las que este supuesto se pasa por alto o simplemente es muy difícil de verificar. Ante esta problemática los autores proponen un algoritmo basado en la técnica de re-muestreo Bootstrap para calcular los límites del control de una gráfica de control T^2 de Hotelling que al compararlo bajo el criterio del ARL bajo control (ARL_0) y el ARL fuera de control (ARL_1) con el gráfico T^2 de Hotelling tradicional presenta un mejor desempeño.

Una gráfica de control combinada es propuesta por Prado, Sifuentes y Monarrez para manejar las debilidades de la gráfica T^2 de Hotelling (Prado, Sifuentes, & Monarrez, 2015) . Estas debilidades son su insensibilidad a cambios pequeños en el vector de medias del proceso, su imposibilidad para identificar la(s) variable(s) que generan una señal fuera de control y su pérdida de sensibilidad cuando aumenta el numero de variables a controlar. La nueva gráfica de control propuesta es una combinación de la gráfica T^2 de Hotelling, la gráfica M (propuesta de Hayter en 1994) y una nueva gráfica basada en componentes principales llamada TMY desarrollada por los autores cuyas principales bondades son: 1) que permite identificar cualquier tipo y magnitud de cambios en el vector de medias del proceso y 2) al momento de presentarse una señal fuera de control, la gráfica combinada es capaz de indicar si es provocada por las variables y/o sus correlaciones. Los autores concluyen y prueban que la nueva gráfica combinada presenta mejor desempeño que cada una de las gráficas de forma individual y la propone como una buena estrategia de monitoreo en ambientes de manufactura.

Siguiendo con los trabajos de modificación de la gráfica de control T^2 de Hotelling, tenemos el trabajo de Ahsan, Mashuri, Kuswanto, Prastyo y Khusna quienes proponen un gráfico de control para monitorizar un proceso con características mixtas es decir un proceso cuyas características de calidad son tanto variables como atributos (Ahsan , Mashuri, Kuswanto, Prastyo, & Khusna, 2018). Usualmente este tipo de características son monitorizadas con procedimientos separados sin embargo su trabajo desarrolla un nuevo concepto de la carta de control T^2 de Hotelling basándose en el análisis de componentes principales mixtos (PCA Mix) que es una técnica que combina el análisis de componentes principales con el análisis de correspondencia, haciendo posible el control simultáneo de estos dos tipos de variables de calidad. Para estimar los límites de control de esta nueva carta los autores utilizaron un método no paramétrico como lo es la estimación de la densidad por Kernel. El desempeño de este gráfico de control fue probado a través del criterio del ARL y demostró excelentes resultados para detectar cambios en el

proceso al hacer uso de un apropiado número de componentes principales, además que esta investigación abre el debate sobre las diferencias en controlar por separado los atributos de las variables o controlarlas de forma simultánea a través de un estadístico que permita captar la información de ambas.

Otros trabajos de investigación que podemos mencionar relacionados con el control estadístico de calidad y cartas de control, son: “*The Variable-Dimension Approach in Multivariate SPC*” (Epprecht, Aparisi, & Ruiz, *The Variable-Dimension Approach in Multivariate SPC*, 2018) donde los autores indican que, al haber variables de calidad que son más costosas de medir que otras, es óptimo medir únicamente aquellas variables no costosas y, cuando estas arrojen una señal fuera de control, acudir a la medición de las variables más costosas. “*Optimum variable-dimension EWMA chart for multivariate statistical process control*” (Epprecht , Aparisi, & Ruiz, 2018), trabajo en el cual los autores llevan a cabo una mejora a la carta de control VDT², que es una carta utilizada cuando en un proceso existen variables fáciles y económicas de medir y otras complicadas y costosas. La carta VDT² prioriza el uso de las variables fáciles y económicas y únicamente hace uso de las variables costosas cuando las primeras dan una señal fuera de control. En este trabajo, se busca mejorar la capacidad del gráfico VDT² para detectar cambios pequeños al construir una versión EWMA del mismo; algo similar a lo que se realizó en este proyecto. Finalmente, en el trabajo “*Reducing sampling costs in multivariate SPC with a double-dimension T²*” (Epprecht, Aparisi, Ruiz , & Veiga, 2013) se construye, también, un gráfico de control que utiliza las variables de calidad más fáciles y menos costosas de medir, demostrando reducción en los costos de muestreo y, en algunos casos, mejor desempeño que la T² utilizando todas las variables de calidad.

CAPÍTULO III

3. METODOLOGÍA DEL DISEÑO

En este capítulo se explicará la teoría del gráfico de control multivariante con efecto memoria propuesto que se trata de una mejora al gráfico control T^2 de Hotelling. La idea es desarrollar un $EWMA(T^2)$ dando como resultado un gráfico de control que combina el aspecto multivariante del T^2 con el efecto memoria y capacidad para detectar cambios pequeños del gráfico $EWMA$.

3.1. GRÁFICO DE CONTROL EWMA(T^2)

Como se mencionó en el capítulo anterior, el estadístico T^2 , utilizado para un conjunto de variables (X_1, X_2, \dots, X_p) que siguen una distribución normal multivariada, se aproxima a una distribución chi-cuadrado con p grados de libertad y está definido por la siguiente ecuación:

$$T^2 = n(\bar{X} - \hat{\mu})^T S^{-1} (\bar{X} - \hat{\mu}) \quad (1)$$

La gráfica de control basada en este estadístico da una señal fuera de control cuando $T^2 \geq LCS$, donde LCS es el límite de control superior de la gráfica para un nivel de significancia dado (α).

La propuesta de este proyecto es construir una versión $EWMA$ del gráfico de control T^2 de Hotelling al cual llamaremos $EWMA(T^2)$. El estadístico por graficarse en esta carta de control sería:

$$EWMA(T^2)_t = r * (T^2)_t + (1 - r) * EWMA(T^2)_{t-1} \quad \text{para } t = 1, 2, 3, \dots \quad (2)$$

En donde $(T^2)_t$ es el valor del estadístico T^2 calculado usando la ecuación (1) con los valores de la muestra t , r es la constante de suavizamiento que cumple con la restricción de $0 < r \leq 1$ (con la particularidad de que si $r = 1$, el gráfico de control $EWMA(T^2)$ resultaría en el mismo gráfico T^2). Y, $EWMA(T^2)_0$ el valor inicial que se necesitaría para calcular $EWMA(T^2)_1$, es el valor esperado de T^2 cuando se encuentra bajo control que estaría dado por:

$$EWMA(T^2)_0 = E(T^2) = p \quad (3)$$

Ya que al tener T^2 una distribución chi-cuadrado con p grados de libertad, su valor esperado será justamente el valor de los grados de libertad.

Finalmente, el gráfico de control $EWMA(T^2)$ tiene dos límites de control, LCS y LCI ; siendo este último igual a 0.

3.2. DESEMPEÑO DEL GRÁFICO DE CONTROL EWMA(T²)

La medida más usada para evaluar el desempeño de los gráficos de control es sin duda el ARL, que se define como el número esperado de muestras hasta que el gráfico presenta una señal fuera de control. Mientras más grande sea el ARL cuando el proceso se encuentre bajo control mejor será el gráfico puesto que se reduciría la aparición temprana de falsas alarmas y; por otro lado, mientras más pequeño sea el ARL cuando el proceso se encuentra fuera de control, mejor.

El objetivo ideal sería entonces maximizar el ARL bajo control y minimizar el ARL fuera de control, pero en la práctica un buen diseño de un gráfico de control es aquel que arroja un ARL fuera de control tan pequeño como sea posible mientras se mantiene el ARL bajo control en niveles aceptables (usualmente elevados). El ARL fuera de control, además, no es único sino más bien está en función del valor del cambio en los parámetros del proceso. El

procedimiento más usual para optimizar el diseño de un gráfico de control es especificar un valor aceptable para el ARL bajo control y, con este valor como restricción, minimizar el ARL fuera de control para un cambio dado y que ha sido especificado por el usuario.

En los gráficos tipo Shewhart donde las muestras son independientes, el número de muestras hasta que se dé una señal de fuera de control sigue una distribución geométrica con parámetro q , siendo q la probabilidad de que cualquier muestra esté fuera de control y el ARL el valor esperado de esta distribución dado por la expresión: $ARL = 1/q$.

Para el caso de los gráficos EWMA, la forma recursiva de calcular el valor del estadístico EWMA introduce correlación entre las muestras lo cual hace que la fórmula del ARL expuesta en el párrafo anterior no sea válida. Además, la probabilidad de que haya una señal fuera de control tampoco es constante ya que está en dependencia del valor anterior del estadístico EWMA. En este caso el ARL debe ser calculado haciendo ecuaciones integrales, cadenas de Markov o modelos de simulación. En este proyecto se utilizó las cadenas de Markov cuya teoría fue expuesta en el capítulo 2, basados en el buen desempeño de estas en el trabajo de García Bustos (García Bustos S. , 2014).

Hay que tener en cuenta también que en gráficos de control con efecto memoria (como lo es el gráfico EWMA), donde hay un valor inicial para comenzar con los cálculos recursivos, el ARL fuera de control (al cual denotaremos como ARL_1) puede ser diferente para el caso en que el proceso está fuera de control desde un comienzo (ARL_1 zero state) o para el caso en que el proceso comienza bajo control y el cambio sucede después (ARL_1 steady state).

Para el cálculo del ARL, dividamos el intervalo entre el LCS y LCI en m grupos, y denotemos los límites superior e inferior del i -ésimo intervalo como U_i y L_i para los $i = 1, 2, \dots, m$. También definamos a M_i como el punto medio del i -

ésimo intervalo. El ancho de cada intervalo estaría dado con la siguiente expresión:

$$h = \frac{UCL - LCL}{m}$$

Puesto que $LCL = 0$, la expresión de la amplitud de cada intervalo quedaría definida como:

$$h = \frac{UCL}{m}$$

Y las expresiones para el límite superior, límite inferior y punto intermedio del intervalo i -ésimo serían las siguientes:

$$U_i = i * h$$

$$L_i = (i - 1) * h$$

$$M_i = (2i - 1) * (h/2)$$

Definamos ahora una cadena de Markov con m estados transitorios tal que el estado i ($i = 1, 2, \dots, m$) corresponde al evento en que el estadístico $EWMA(T^2)$ se encuentra en el i -ésimo subintervalo y, un estado absorbente, sería el caso de que el estadístico $EWMA(T^2)$ esté por debajo del LCI o por encima del LCS.

Definamos también la matriz de transición Q con dimensión $m \times m$, que está formada por las probabilidades q_{ij} de que el estadístico $EWMA(T^2)_t$ esté dentro del j -ésimo subintervalo dado que $EWMA(T^2)_{t-1}$ está dentro del i -ésimo subintervalo entre LCI y LCS. Para efecto de simplificar los cálculos se supondrá que cuando el estadístico $EWMA(T^2)_{t-1}$ está en el i -ésimo intervalo, su valor coincide exactamente con el punto medio de ese intervalo. Esta aproximación ha sido utilizada por muchos autores (García Bustos, Aparisi, & Epprecht, 2014) y se ha verificado su precisión siempre que se trabaje con un número elevado de estados, de tal manera que la amplitud de los intervalos

sea mínima. En el Gráfico 2 puede apreciarse de forma visual esta probabilidad q_{ij} .

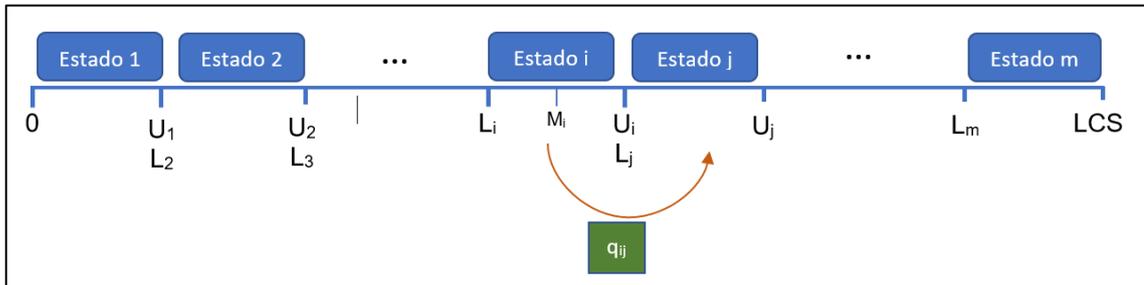


Gráfico 2: Diagrama de probabilidades de transición q_{ij}

Fuente: Elaboración propia

Con todo lo anterior expuesto entonces, q_{ij} quedaría expresado de la siguiente manera:

$$q_{ij} = P(L_j < EWMA(T^2)_t < U_j \mid EWMA(T^2)_{t-1} = M_i)$$

Haciendo uso de la ecuación (2) podemos expresar esta ecuación de la siguiente manera:

$$q_{ij} = P(L_j < r * (T^2)_t + (1 - r) * EWMA(T^2)_{t-1} < U_j \mid EWMA(T^2)_{t-1} = M_i)$$

Hasta que finalmente llegaremos a la siguiente expresión:

$$q_{ij} = P\left(\frac{L_j - (1 - r)M_i}{r} < T^2 < \frac{U_j - (1 - r)M_i}{r}\right)$$

Y, al tener T^2 una distribución chi-cuadrado con p grados de libertad la probabilidad q_{ij} se la calculará como la probabilidad de que una chi-cuadrado con p grados de libertad esté entre esos dos valores.

Es necesario diferenciar la matriz de transición para los casos en que el proceso esté bajo control (Q_0) y para los casos en los que el proceso esté fuera

de control (Q_1). Para el caso de Q_0 , las probabilidades de transición se calcularán de la forma antes expuesta; mientras que para el caso de Q_1 las probabilidades de transición q_{ij} se calcularán con una distribución chi-cuadrado descentralizada con p grados de libertad y parámetro de descentralización λ .

Finalmente, por las probabilidades de la cadena de Markov se tiene que:

$$ARL(d = 0) = v'(I - Q_0)^{-1}u$$

En donde $ARL(d = 0)$ es el valor del ARL bajo control, u es un vector de unos con dimensión $m \times 1$, I es una matriz identidad con dimensiones $m \times m$ y v es el vector de probabilidad de estados iniciales con dimensión $m \times 1$ y que tiene todos los elementos iguales a cero excepto el que corresponde al subintervalo que contiene el valor de $EWMA(T^2)_0$ dado por la ecuación (3) el cual es igual a uno. El intervalo que contiene el valor de $EWMA(T^2)_0$ tiene como índice el menor entero que es igual o mayor que p/h .

El valor del $ARL(d = d^*)$ en zero state, es decir el ARL fuera de control para un cambio de d^* cuando el proceso está fuera de control en el momento de la monitorización, está dado por:

$$ZARL(d = d^*) = v'(I - Q_1)^{-1}u$$

Y el valor del $ARL(d = d^*)$ steady state, es decir el ARL fuera de control cuando el proceso está en control al momento de la monitorización y el cambio se produce después, está dado por:

$$SARL(d = d^*) = w'(I - Q_1)^{-1}u$$

En donde w es igual a:

$$w = \frac{v'(I - Q_0)^{-1}}{ARL(d = 0)}$$

3.3. OPTIMIZACIÓN DEL GRÁFICO DE CONTROL EWMA(T^2)

Para encontrar los parámetros óptimos del gráfico de control propuesto se hizo uso de Algoritmos Genéticos (GA). Esta técnica fue programada en Microsoft Visual Studio diseñando una interfaz gráfica, amigable al usuario, que le permita ingresar los datos de entrada y obtener los parámetros optimizados del gráfico los cuales son: el límite de control superior (LCS) y la constante de suavizado (r), además de los valores óptimos de ARL fuera de control (zero state y steady state). Como referencia para la construcción de este optimizador se tomaron en cuenta los trabajos realizados por García-Bustos (García Bustos S. , 2014) y García-Bustos et al. (García Bustos, Aparisi, & Epprecht , 2014)

El problema de optimización puede expresarse de la siguiente forma:

DATOS DE ENTRADA

ARL bajo control deseado: ARL_0

Número de variables (p)

Distancia de Mahalanobis para la cual se minimiza el ARL

OBJETIVO

Minimizar: $SARL(d = d^*)$

RESTRICCIÓN

$ARL(d = 0) = ARL_0$

VARIABLES DE DECISIÓN

Límite de control superior (LCS)

Constante de suavizado (r)

3.4. INTERFAZ GRÁFICA DEL OPTIMIZADOR

Como se mencionó anteriormente el optimizador fue realizado en Microsoft Visual Studio con una interfaz gráfica en la cual el usuario ingresará: el número de variables (p), la distancia de Mahalanobis para la cual se desea optimizar los parámetros y el ARL bajo control deseado. Cabe mencionar que utilizamos la distancia de Mahalanobis porque es una distancia entre dos variables aleatorias multidimensionales que toma en cuenta la correlación entre las variables aleatorias (Salas Plata & Portillo, 2008).

Una vez ingresada toda esta información, el usuario dará clic en “START” y el programa comenzará a ejecutarse hasta que finalmente mostrará en pantalla las soluciones óptimas. Específicamente el programa enseña lo siguiente: LCS (límite de control superior) óptimo, la constante de suavizamiento óptima (r), el ARL bajo control que es aproximadamente igual al ARL ingresado por el usuario, el ARL fuera de control steady state y, finalmente el ARL fuera de control zero state.

En el gráfico a continuación podemos observar el programa (llamado “Aplicación EWMA T2”, realizando la optimización para: 10 variables, distancia de Mahalanobis de 2 y un ARL bajo control deseado de 500, En el recuadro a la derecha se observan los parámetros óptimos que son: límite de control superior de 12,7, constante de suavizamiento de 0,09, ARL bajo control de 504,23, ARL fuera de control steady state de 12,05 y un ARL fuera de control zero state de 12,12.

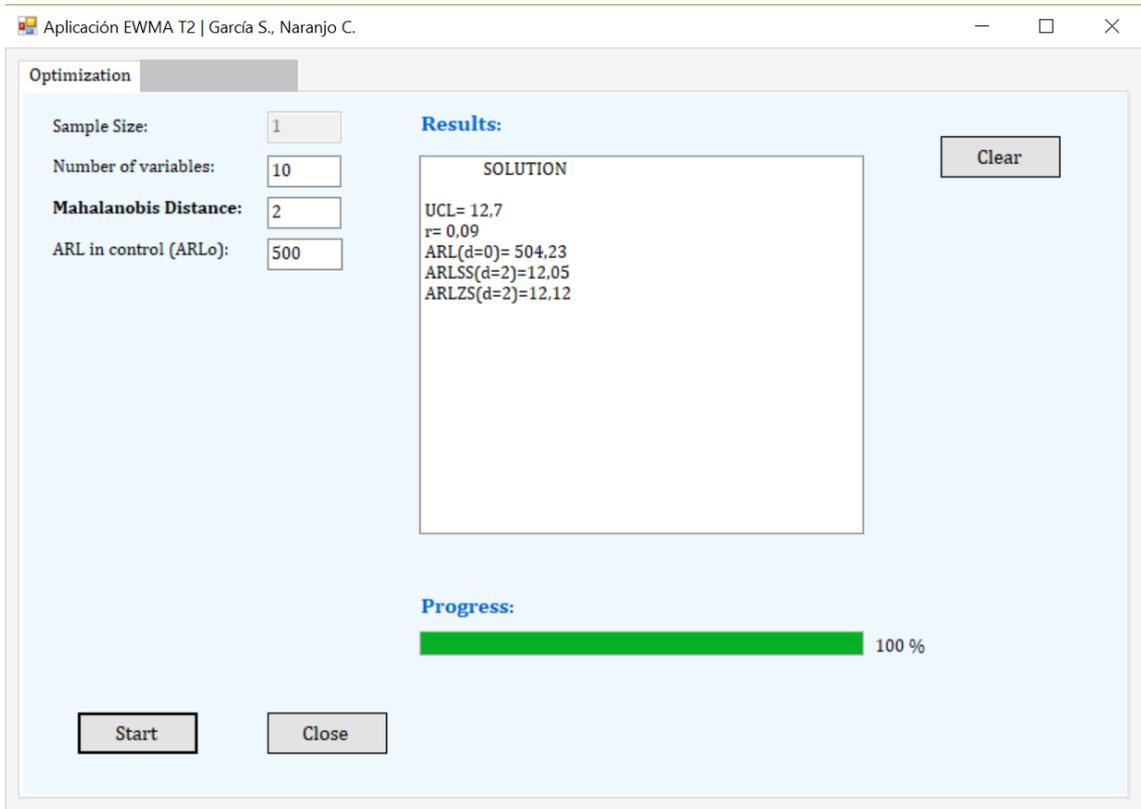


Gráfico 3: Captura de pantalla del programa “Aplicación EWMA T²”

Fuente: Elaboración propia

3.5. EJEMPLO DE APLICACIÓN

Para reforzar la explicación del uso del software utilizaremos un ejemplo de aplicación tomado del trabajo de Plaza en donde diseñó una carta control para la combinación lineal de variables normales (Plaza Rodas, 2017). El ejemplo se trata de un proceso de producción de galletas en el cual se definen tres variables de calidad a controlar que son: peso de la galleta horneada, volumen de la galleta horneada y peso crudo maquinado, con sus respectivas unidades que son gramos (g) y centímetros cúbicos (cm³). Estas variables deben de cumplir con las especificaciones dadas por el departamento de calidad de la industria, ya que cualquier incremento o disminución podría dar como resultado problemas en el proceso.

La matriz de correlación de las tres variables mencionadas es detallada a continuación. Nótese su correlación positiva lo que nos lleva a concluir que: a

mayor peso del material crudo, mayor será el peso de la galleta horneada y por ende mayor su volumen.

$$\begin{array}{l} \text{Peso Crudo} \\ \text{Peso Horneado} \\ \text{Volumen} \end{array} \begin{pmatrix} 1,000 & 0,900 & 0,700 \\ 0,900 & 1,000 & 0,800 \\ 0,700 & 0,800 & 1,000 \end{pmatrix}$$

El objetivo es obtener los parámetros optimizados de la carta de control EWMA(T^2) para detectar rápidamente los cambios en μ_1 , μ_2 y μ_3 en una desviación estándar con la condición de que el $ARL_0=400$. Se trabajará bajo el supuesto de que las variables están estandarizadas por lo que el vector de medias cuando el proceso está bajo control es el $\vec{0}$ con la matriz de correlación ya indicada.

Para resolver este problema de aplicación el usuario tendrá que colocar en el software lo siguiente:

- Como número de variables: 3
- Como ARL bajo control deseado: 400
- Como distancia de Mahalanobis la resultante al tener un desplazamiento de 1 unidad en cada una de las medias de las tres variables y con la matriz de correlación dada: 1,176471

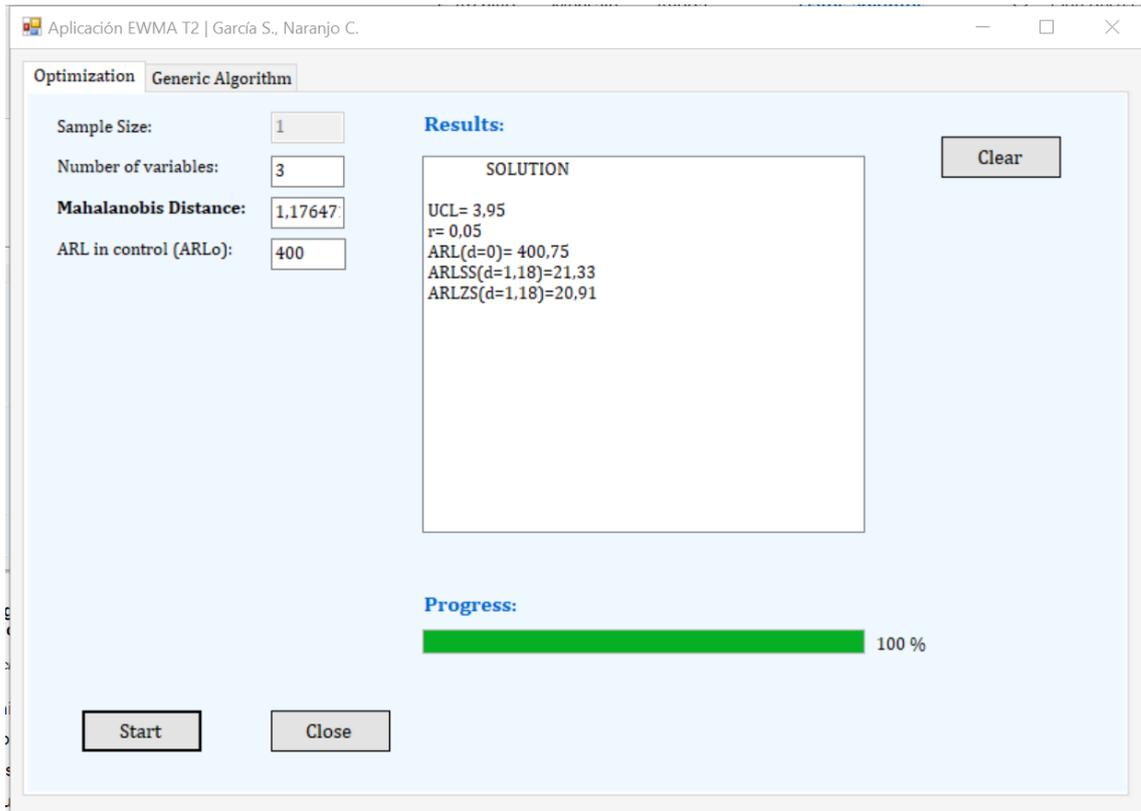


Gráfico 4: Salida del software para el gráfico de control EWMA(T^2) optimizado para el ejemplo de aplicación propuesto

Fuente: Elaboración propia

Interpretando la salida del programa, tenemos que el límite de control superior optimizado es de 3,95. Bajo control, el gráfico mostrará una señal de falsa alarma en la muestra número 400; y, cuando exista un incremento de una desviación estándar en cada una de las medias de las variables de calidad analizadas, el gráfico lo detectará en promedio cuando se hayan tomado 21 muestras de tamaño 1.

CAPÍTULO IV

4. RESULTADOS

En este capítulo se presentará el desempeño del nuevo gráfico de control propuesto EWMA(T^2) y se compararan sus rendimientos con el gráfico T^2 de Hotelling y con el gráfico MEWMA (Multivariate Exponentially Moving Average Chart). Como se mencionó en el capítulo anterior (en el apartado 3.4) para construir estos resultados se desarrolló en Microsoft Visual Studio un programa que permita encontrar los parámetros óptimos de la carta de control propuesta mediante Algoritmos Genéticos (GA).

Los valores del ARL fuera de control de la gráfica T^2 de Hotelling fueron obtenidos a través del software libre R y los valores del ARL fuera de control de la gráfica MEWMA fueron tomados del libro de Montgomery "*Introduction to Statistical Quality Control*". (Montgomery, 2006).

También en este capítulo se llevó a cabo un análisis de sensibilidad de la carta de control propuesta para analizar el comportamiento de una gráfica de control optimizada ante diferentes cambios.

4.1. DESEMPEÑO DEL GRÁFICO EWMA(T^2)

El desempeño del gráfico EWMA(T^2) se medirá a través del valor del ARL fuera de control steady state y se comparará con el valor del ARL fuera de control de los gráficos T^2 de Hotelling y MEWMA.

Se construyeron diversos escenarios para analizar el comportamiento del gráfico EWMA(T^2) en diferentes combinaciones de: ARL bajo control deseado,

Número de variables y cambios. En total se analizaron 30 escenarios que se resumen en la tabla a continuación:

| # escenario | ARL deseado (ARL ₀) | # de variables (p) | Cambio (d) | # escenario | ARL deseado (ARL ₀) | # de variables (p) | Cambio (d) |
|-------------|---------------------------------|--------------------|------------|-------------|---------------------------------|--------------------|------------|
| 1 | 200 | 2 | 0,5 | 16 | 200 | 10 | 0,5 |
| 2 | 200 | 2 | 1,0 | 17 | 200 | 10 | 1,0 |
| 3 | 200 | 2 | 1,5 | 18 | 200 | 10 | 1,5 |
| 4 | 200 | 2 | 2,0 | 19 | 200 | 10 | 2,0 |
| 5 | 200 | 2 | 3,0 | 20 | 200 | 10 | 3,0 |
| 6 | 200 | 4 | 0,5 | 21 | 500 | 4 | 0,5 |
| 7 | 200 | 4 | 1,0 | 22 | 500 | 4 | 1,0 |
| 8 | 200 | 4 | 1,5 | 23 | 500 | 4 | 1,5 |
| 9 | 200 | 4 | 2,0 | 24 | 500 | 4 | 2,0 |
| 10 | 200 | 4 | 3,0 | 25 | 500 | 4 | 3,0 |
| 11 | 200 | 6 | 0,5 | 26 | 500 | 10 | 0,5 |
| 12 | 200 | 6 | 1,0 | 27 | 500 | 10 | 1,0 |
| 13 | 200 | 6 | 1,5 | 28 | 500 | 10 | 1,5 |
| 14 | 200 | 6 | 2,0 | 29 | 500 | 10 | 2,0 |
| 15 | 200 | 6 | 3,0 | 30 | 500 | 10 | 3,0 |

Tabla 2: Escenarios de análisis del desempeño de la gráfica propuesta EWMA(T²)

Fuente: Elaboración propia

Presentamos los resultados para los escenarios del 1 al 5: 2 variables ($p=2$) y ARL bajo control deseado de 200. Como podemos apreciar en la **Tabla 3**, el gráfico de control propuesto EWMA(T²) presenta una mejora considerable cuando lo comparamos con el gráfico T² de Hotelling, mejora que se evidencia sobre todo en los cambios pequeños. Por ejemplo, para un cambio de 0,50 la carta de control EWMA(T²) presenta un ARL fuera de control zero state de 67,38 mientras que la carta de control T² presenta un ARL fuera de control de 115,53. Esto quiere decir que la carta de control propuesta, en este escenario, detecta una señal fuera de control en 41,68% menos pasos que la T² de Hotelling.

Sin embargo, al comparar la carta de control propuesta con la carta de control MEWMA, vemos que esta última tiene un mejor performance en todos los escenarios, pero la diferencia porcentual en el ARL fuera de control entre estas

dos gráficas comienza a disminuir a medida que el cambio (distancia de Mahalanobis) va en aumento. Por ejemplo, para un cambio de 3,00 la carta de control EWMA(T^2) requiere 1,80% más pasos para detectar una señal fuera de control que la carta de control MEWMA; mientras que para un cambio de 0,50 requiere 153,20% más pasos.

Es interesante también observar como a medida que el cambio analizado aumenta, lo hace también la constante de suavizamiento r , esto quiere decir que mientras más grande es el cambio menor peso se les da a las observaciones pasadas.

| p=2; ARL ₀ = 200 | | | | | | | | |
|-----------------------------|---------------|------|-------------|-----------|-------|-------------|------------------------|------------------------|
| Cambio | EWMA(T^2) | | | T^2 | MEWMA | | Diferencia % con T^2 | Diferencia % con MEWMA |
| | LCS | r | ARLSS(d=d*) | ARL(d=d*) | r | ARLSS(d=d*) | | |
| 0,50 | 2,52 | 0,04 | 67,38 | 115,53 | 0,05 | 26,61 | -41,68% | 153,20% |
| 1,00 | 2,85 | 0,07 | 18,94 | 41,92 | 0,05 | 11,23 | -54,82% | 68,63% |
| 1,50 | 3,23 | 0,11 | 7,81 | 15,78 | 0,10 | 6,11 | -50,49% | 27,82% |
| 2,00 | 5,81 | 0,42 | 4,70 | 6,88 | 0,40 | 3,53 | -31,58% | 33,25% |
| 3,00 | 7,01 | 0,57 | 1,93 | 2,16 | 0,50 | 1,90 | -10,41% | 1,80% |

Tabla 3: Desempeño de la carta de control EWMA(T^2) con ARL₀=200 y p=2.

Fuente: Elaboración propia

En la **Tabla 4** a continuación, presentamos los resultados para cuando se desea controlar 4 variables ($p=4$) con un ARL bajo control deseado de 200. Se probaron los mismos cambios que en la tabla anterior y se puede evidenciar una reducción en el ARL fuera de control en la carta EWMA(T^2) con respecto a la T^2 de Hotelling, es decir que la carta de control propuesta tiene mejor desempeño al momento de detectar señales fuera de control.

Al comparar la gráfica de control EWMA(T^2) con la carta de control MEWMA, esta última sigue teniendo mejor rendimiento.

| p=4; ARL ₀ = 200 | | | | | | | | |
|-----------------------------|-----------------------|------|-------------|----------------|-------|-------------|---------------------------------|------------------------|
| Cambio | EWMA(T ²) | | | T ² | MEWMA | | Diferencia % con T ² | Diferencia % con MEWMA |
| | LCS | r | ARLSS(d=d*) | ARL(d=d*) | r | ARLSS(d=d*) | | |
| 0,50 | 4,73 | 0,04 | 87,48 | 138,15 | 0,05 | 32,29 | -36,68% | 170,91% |
| 1,00 | 4,73 | 0,04 | 25,33 | 60,96 | 0,05 | 13,48 | -58,45% | 87,88% |
| 1,50 | 6,02 | 0,14 | 11,18 | 24,62 | 0,10 | 7,22 | -54,57% | 54,91% |
| 2,00 | 6,02 | 0,14 | 5,64 | 10,63 | 0,10 | 5,19 | -46,97% | 8,59% |
| 3,00 | 7,41 | 0,27 | 2,37 | 2,93 | 0,30 | 2,50 | -19,11% | -5,16% |

Tabla 4: Desempeño de la carta de controla EWMA(T²) con ARL₀=200 y p=4.

Fuente: Elaboración propia

El resto de los escenarios para el ARL bajo control deseado de 200 se encuentran en el apartado “Apéndices” del presente proyecto.

Analizaremos a continuación los escenarios del 26 al 30, que corresponden a los escenarios de 10 variables ($p=10$) y ARL bajo control deseado de 500 (véase **Tabla 5**). Al igual que los dos escenarios anteriores se aprecia una notable mejora en el ARL fuera de control de la carta EWMA(T²) con respecto a la T² de Hotelling. Por ejemplo, para un cambio de 1,50, EWMA(T²) da una alerta en 26,61 pasos mientras que la carta T² da la alerta en 91,72; esto evidencia una mejora de aproximadamente el 71%. Igual para el cambio de 2,00 en donde la nueva carta propuesta da una señal de fuera de control en 69,02% menos pasos que a T² de Hotelling.

| p=10; ARL ₀ = 500 | | | | | | | | |
|------------------------------|-----------------------|------|-------------|----------------|-------|-------------|---------------------------------|------------------------|
| Cambio | EWMA(T ²) | | | T ² | MEWMA | | Diferencia % con T ² | Diferencia % con MEWMA |
| | LCS | r | ARLSS(d=d*) | ARL(d=d*) | r | ARLSS(d=d*) | | |
| 0,50 | 11,51 | 0,04 | 261,71 | 391,80 | 0,03 | 55,94 | -33,20% | 367,84% |
| 1,00 | 11,77 | 0,05 | 70,85 | 208,91 | 0,09 | 19,29 | -66,09% | 267,29% |
| 1,50 | 12,90 | 0,10 | 26,61 | 91,72 | 0,16 | 10,01 | -70,99% | 165,84% |
| 2,00 | 13,29 | 0,12 | 11,89 | 38,39 | 0,24 | 6,25 | -69,02% | 90,28% |
| 3,00 | 13,29 | 0,12 | 4,40 | 7,96 | 0,42 | 3,24 | -44,73% | 35,74% |

Tabla 5: Desempeño de la carta de controla EWMA(T²) con ARL₀=500 y p=10.

Fuente: Elaboración propia

En general, podemos observar que la gráfica de control creada presenta un mejor desempeño que la gráfica de control T^2 de Hotelling, demostrando la potencia del gráfico EWMA para detectar cambios pequeños. Las mejoras conseguidas en este proyecto son similares a las conseguidas por los autores Epprecht, Aparisi y Ruiz en su trabajo "*Optimum variable-dimension EWMA chart for multivariate statistical process control*" (Epprecht, Aparisi, & Ruiz, 2018) donde también utilizaron el gráfico EWMA para potenciar el rendimiento del gráfico VDT².

Por otro lado, al comparar la carta de control EWMA(T^2) con la MEWMA, esta última sigue presentando mejores resultados como se vio también en los dos ejemplos anteriores.

El resumen del desempeño de los escenarios 21 al 25 se encuentra en el apartado "Apéndice" del proyecto.

4.2. ANÁLISIS DE SENSIBILIDAD DEL GRÁFICO EWMA(T^2)

Luego de comprobar el buen desempeño del gráfico EWMA(T^2) frente al gráfico T^2 de Hotelling, en esta parte del proyecto analizaremos la sensibilidad de este. Como se mencionó anteriormente el nuevo gráfico de control propuesto tiene como objetivo determinar los parámetros que detectan de forma rápida un cambio dado (d^*) y es por ello que tendrá un buen desempeño para este cambio. Lo que plantea la interrogante de qué pasaría si con esos mismos parámetros óptimos probamos el desempeño del gráfico, pero para otros cambios. En esto nos ayudará el análisis de sensibilidad, para comparar los distintos valores del ARL fuera de control para diferentes cambios de dos gráficos de control EWMA(T^2) optimizados para d_1^* y d_2^* respectivamente.

Para llevar a cabo este análisis hemos planteado cuatro casos, para cada uno de ellos hallaremos sus parámetros óptimos y luego compararemos su

desempeño en diferentes cambios. Los casos fueron planteados de tal manera que podamos realizar tres comparaciones: 1) entre dos gráficos de control optimizados para cambios pequeños, 2) entre dos gráficos de control optimizados para cambios grandes y 3) entre un gráfico de control optimizado para cambios pequeños y uno optimizado para cambios grandes. Estos casos son:

| Caso | Cambio | p | ARL ₀ |
|------|--------|---|------------------|
| 1 | 0,50 | 2 | 500 |
| 2 | 0,80 | 2 | 500 |
| 3 | 3,00 | 2 | 500 |
| 4 | 3,50 | 2 | 500 |

Tabla 6: Casos definidos para el análisis de sensibilidad

Fuente: Elaboración propia

A continuación, haremos la comparación del desempeño entre el caso 1 y el caso 3, es decir entre una carta de control optimizada para cambios pequeños y otra optimizada para cambios grandes. Como podemos ver en la **Tabla 7**, el gráfico de control optimizado para el cambio 0,50 resulta tener mejor desempeño en casi todos los cambios en comparación al gráfico de control optimizado para el cambio de 3,00. El único cambio para el que el gráfico de control del caso 3 resulta ser mejor es aquel para el cual fue construido y optimizado. También, es importante mencionar que ambas cartas de control EWMA(T²) resultan tener mejor desempeño que la carta T² de Hotelling.

| Cambio | Caso 1 | Caso 3 | T2 |
|--------|--------------------|--------------------|--------|
| | r=0,05 UCL=2,84 | r=0,58 UCL=8,18 | |
| 0,50 | 138,97 | 244,78 | 265,74 |
| 1,00 | 26,80 | 66,52 | 84,95 |
| 1,50 | 10,36 | 19,56 | 28,33 |
| 2,00 | 5,66 | 7,49 | 11,00 |
| 3,00 | 2,68 | 2,33 | 2,81 |

Tabla 7: ARL fuera de control para el gráfico EWMA(T²) optimizados para los cambios 0,50 y 3,00.

Fuente: Elaboración propia

Comparamos ahora dos cartas de control EWMA(T^2) optimizadas para cambios pequeños (0,50 y 0,80 respectivamente) (véase **Tabla 8**). Podemos evidenciar que la carta de control optimizada para el cambio de 0,50 es mejor que la optimizada para el cambio 0,80 pero únicamente para las distancias de 0,50 y de 1,00. A partir de la distancia de 1,50 la carta de control del caso 2 comienza a tener un mejor desempeño.

| Cambio | Caso 1 | Caso 2 | T2 |
|--------|--------------------|--------------------|--------|
| | r=0,05 UCL=2,84 | r=0,07 UCL=3,09 | |
| 0,50 | 138,97 | 150,53 | 265,74 |
| 1,00 | 26,80 | 27,94 | 84,95 |
| 1,50 | 10,36 | 10,20 | 28,33 |
| 2,00 | 5,66 | 5,46 | 11,00 |
| 3,00 | 2,68 | 2,55 | 2,81 |

Tabla 8: ARL fuera de control para el gráfico EWMA(T^2) optimizados para los cambios 0,50 y 0,80.

Fuente: Elaboración propia

Finalmente compararemos los casos 3 y 4 que son cartas de control EWMA(T^2) optimizadas para cambios grandes. En la **Tabla 9** podemos evidenciar como la carta de control optimizada para el cambio de 3,00 tiene un mejor desempeño que la carta de control optimizada para el cambio de 3,50.

| Cambio | Caso 3 | Caso 4 | T2 |
|--------|--------------------|--------------------|--------|
| | r=0,58 UCL=8,18 | r=0,66 UCL=8,97 | |
| 0,50 | 244,78 | 251,23 | 265,74 |
| 1,00 | 66,52 | 71,25 | 84,95 |
| 1,50 | 19,56 | 21,46 | 28,33 |
| 2,00 | 7,49 | 8,13 | 11,00 |
| 3,00 | 2,33 | 2,40 | 2,81 |

Tabla 9: ARL fuera de control para el gráfico EWMA(T^2) optimizados para los cambios 3,00 y 3,50.

Fuente: Elaboración propia

Se evidencia, entonces, que la carta de control EWMA(T^2) sí es sensible a los cambios. En general, cuando optimizamos una carta de control EWMA(T^2)

para un cambio d^* dado, esta va a tener un buen desempeño para detectar cambios iguales o superiores a d^* , pero al momento de ser necesario detectar cambios menores al cambio d^* la carta de control comienza a desmejorar su desempeño.

CAPÍTULO V

5. CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

- Se desarrolló un nuevo gráfico de control para variables normales multivariantes al que se llamó EWMA(T^2) y es una mejora del gráfico T^2 que combina el aspecto multivariante de T^2 de Hotelling con la capacidad para detectar cambios pequeños del EWMA.
- Se efectuó un programa en Visual Studio 2017 que permite calcular los parámetros óptimos de esta nueva gráfica de control propuesta (límite de control superior LCS y constante de suavizamiento r).
- Se demostró, usando el criterio de ARL, que el gráfico de control EWMA(T^2) tiene mejores rendimientos que el gráfico de control T^2 de Hotelling. Sin embargo, no superó en rendimiento al gráfico de control MEWMA.
- Se observó que a medida que el cambio analizado aumenta, aumenta también la constante de suavizamiento r , con lo que se concluye que a mayor cambio menor es el peso que se le da a las observaciones pasadas.
- Se evidenció, mediante el análisis de sensibilidad, que la carta de control EWMA(T^2) es sensible a los cambios. En general, si la carta es diseñada y optimizada para un cambio d^* , presentará un buen desempeño para los cambios iguales o superiores a d^* .

5.2. Recomendaciones

- Probar el rendimiento de este nuevo gráfico de control con datos reales de un proceso y compararlo con el rendimiento del gráfico T^2 de Hotelling para afianzar las conclusiones obtenidas en este proyecto.
- Se recomienda también seguir con esta línea de investigación y combinar los gráficos MEWMA y T^2 , para mejor el rendimiento no sólo del gráfico T^2 de Hotelling sino también del mismo gráfico MEWMA.

BIBLIOGRAFÍA

- Ahsan , M., Mashuri, M., Kuswanto, H., Prastyo, D. D., & Khusna, H. (2018). Multivariate control chart based on PCA mix for variable and attribute quality characteristics. *Production & Manufacturing Research*, 364-384.
- Aparisi, F., & García-Díaz, J. (2007). Design and Optimization of EWMA Control Charts for In-control, Indifferent, and Out-of-control Regions. *Computers & Operations Research* , 2096-2108.
- Aparisi, F., García Bustos, S., & Epprechth, E. (2013). Optimum Multiple and Multivariate Poisson Statistical Control Charts. *Quality and Reliability Engineering International*.
- Baba, S., Jabir , H., & Abdul-Salam, S. (2017). One-sided cumulative sum control chart for monitoring shift in the scale parameter delta (δ), of the new Weibull-Pareto distribution. *International Journal of Probability and Statistics*, Vol. 6, No. 4, pp.76–81, DOI: 10.5923/j.ijps.20170604.02.
- Brook , D., & Evans, D. (1972). An Approach to the Probability Distribution of Cusum Run Length. *Biometrika*, 539-549.
- Carr, J. (2014). An introduction to genetic algorithms. *Senior Project*, 1, 40.
- Contreras, A., Marroquin, E., & Zertuche, F. (2011). Monitoreo y Control de un Proceso Normal Multivariado. *Conciencia Tecnológica No. 43*,, 29-35.
- Epprecht , E. K., Aparisi, F., & Ruiz, O. (2018). Optimum variable-dimension EWMA chart for multivariate statistical process control. *Quality Engineering*, 268-282.
- Epprecht, E. K., Aparisi, F., & Ruiz, O. (2018). The Variable-Dimension Approach in Multivariate SPC. *Frontiers in Statistical Quality Control 12*, 143-155. doi:https://doi.org/10.1007/978-3-319-75295-2_8
- Epprecht, E. K., Aparisi, F., Ruiz , O., & Veiga, Á. (2013). Reducing sampling costs in multivariate SPC with a double-dimension T2 control chart. *International Journal of Production Economics*, 90-104.
- Gagniuc, P. A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. John Wiley & Sons.
- García Bustos, S. (2014). Control Multivariante Estadístico de Variables Discretas tipo Poisson. *Tesis doctoral*.

- García Bustos, S., Aparisi, F., & Epprecht, E. (2014). Optimal EWMA of Linear Combination of Poisson Variables for Multivariate Statistical Process Control. *International Journal of Production Research*.
- García-Bustos, S., Mite, M., & Vera, F. (2016). Control charts with variable dimension for linear combination of Poisson variables. *Quality and Reliability Engineering International*, DOI:10.1002/qre.1910.
- Holland, J. H. (1975). Adaptation in Natural and Artificial Systems. *University of Michigan Press, Ann Arbor*.
- Hotelling, H. (1947). Multivariable Quality Control—Illustrated By The Air Testing Of Sample Bombsights. En C. Eisenhart, M. Hastay, & W. Wallis, *Techniques of Statistical Analysis* (págs. 111-184). New York: McGraw Hill.
- Lu, X., Xie, M., & Goh, T. (1998). Control chart for multivariate attribute processes. *International Journal of Production Research*(36), 3477-3489.
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *National Institute of Science of India*, 49-55.
- Montgomery, D. C. (2006). *Introduction to Statistical Quality Control* (6 ed.). USA: Wiley.
- Mostajeran, A., Iranpanah, N., & Noorossana R. (2016). A New Bootstrap Based Algorithm for Hotelling's T² Multivariate Control Chart. *Journal of Sciences*, 269-278.
- Pérez R., F. O. (2007). *Introducción a las series de tiempo. Métodos paramétricos*. Medellín: Universidad de Medellín.
- Plaza Rodas, A. E. (2017). *Diseño de gráficos de control multivariantes basado en la combinación lineal de variables normales*. Guayaquil: ESPOL.
- Prado, E. M., Sifuentes, M. C., & Monarrez, M. P. (2015). Una gráfica de control combinada para monitorear y controlar procesos multivariados. *CULCyT*, 20.
- Salas Plata, J., & Portillo, M. (2008). *P. Ch. Mahalanobis y las aplicaciones de su distancia estadística*. CULCyT: Cultura Científica y Tecnológica.
- Seber, G. A. (1984). *Multivariate Observations*. Wiley Series in probability and mathematical statistics.
- Umit, F., & Cigdem, A. (2001). Multivariate Quality Control: A Historical Perspective. *Yildiz Technical University*, 54-65.

Vanhatalo , E., & Kulahci, M. (2015). The effect of autocorrelation on the Hotelling T2 control chart. *Quality and Reliability Engineering International*, 1779-1796.

Zertuche Luis, F., Cantú Sifuentes, M., & Piña Monarrez, M. M. (2007). Sistema de control multivariado no paramétrico de procesos. *CULCyT*.

APÉNDICES

APÉNDICE A: Desempeño de la carta de control EWMA(T^2)

| p=6; ARL ₀ = 200 | | | | | | | | |
|-----------------------------|------------|------|-------------|-----------|-------|-------------|------------------------|------------------------|
| Cambio | EWMA T^2 | | | T^2 | MEWMA | | Diferencia % con T^2 | Diferencia % con MEWMA |
| | LCS | r | ARLSS(d=d*) | ARL(d=d*) | r | ARLSS(d=d*) | | |
| 0,50 | 7,06 | 0,05 | 99,64 | 149,46 | 0,05 | 36,39 | -33,34% | 173,80% |
| 1,00 | 7,26 | 0,06 | 33,48 | 74,32 | 0,05 | 15,08 | -54,95% | 122,03% |
| 1,50 | 7,72 | 0,09 | 13,36 | 32,13 | 0,10 | 8,01 | -58,43% | 66,78% |
| 2,00 | 7,72 | 0,09 | 6,93 | 14,12 | 0,10 | 5,74 | -50,92% | 20,76% |
| 3,00 | 8,55 | 0,15 | 2,88 | 3,69 | 0,20 | 3,03 | -22,00% | -5,11% |

Tabla 10: Desempeño de la carta de controla EWMA(T^2) con ARL₀=200 y p=6.

Fuente: Elaboración propia

| p=10; ARL ₀ = 200 | | | | | | | | |
|------------------------------|------------|------|-------------|-----------|-------|-------------|------------------------|------------------------|
| Cambio | EWMA T^2 | | | T^2 | MEWMA | | Diferencia % con T^2 | Diferencia % con MEWMA |
| | LCS | r | ARLSS(d=d*) | ARL(d=d*) | r | ARLSS(d=d*) | | |
| 0,50 | 11,13 | 0,04 | 112,19 | 161,34 | 0,05 | 42,49 | -30,47% | 164,03% |
| 1,00 | 12,00 | 0,08 | 45,16 | 92,48 | 0,10 | 15,98 | -51,17% | 182,60% |
| 1,50 | 12,00 | 0,08 | 17,67 | 44,53 | 0,10 | 9,23 | -60,33% | 91,40% |
| 2,00 | 12,71 | 0,12 | 8,94 | 20,59 | 0,10 | 6,57 | -56,57% | 36,10% |
| 3,00 | 13,04 | 0,14 | 3,57 | 5,21 | 0,10 | 4,28 | -31,38% | -16,48% |

Tabla 11: Desempeño de la carta de controla EWMA(T^2) con ARL₀=200 y p=10.

Fuente: Elaboración propia

| p=4; ARL ₀ = 500 | | | | | | | | |
|-----------------------------|------------|------|-------------|-----------|-------|-------------|------------------------|------------------------|
| Cambio | EWMA T^2 | | | T^2 | MEWMA | | Diferencia % con T^2 | Diferencia % con MEWMA |
| | LCS | r | ARLSS(d=d*) | ARL(d=d*) | r | ARLSS(d=d*) | | |
| 0,50 | 5,15 | 0,05 | 189,26 | 327,14 | 0,04 | 42,22 | -42,15% | 348,27% |
| 1,00 | 5,48 | 0,07 | 42,44 | 129,75 | 0,11 | 14,60 | -67,29% | 190,67% |
| 1,50 | 5,78 | 0,09 | 14,63 | 46,78 | 0,18 | 7,65 | -68,73% | 91,18% |
| 2,00 | 6,98 | 0,18 | 7,25 | 18,06 | 0,28 | 4,82 | -59,84% | 50,46% |
| 3,00 | 5,48 | 0,07 | 3,26 | 4,05 | 0,52 | 2,55 | -19,49% | 27,78% |

Tabla 12: Desempeño de la carta de controla EWMA(T^2) con ARL₀=500 y p=4.

Fuente: Elaboración propia