

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ciencias Naturales y Matemáticas

Diseño de un gráfico de control multivariante para variables normales
usando el esquema de muestreo Generalized Multiple Dependent
State

PROYECTO INTEGRADOR

Previo la obtención del Título de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

Presentado por:
Joseph Jonathan León Cando

GUAYAQUIL - ECUADOR

Año: 2019

DEDICATORIA

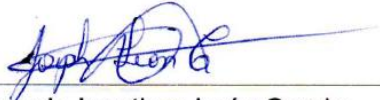
A Dios y a mi familia.

AGRADECIMIENTOS

Agradezco a profesores y compañeros que fueron parte directa o indirectamente en mi proceso de formación académica y que aportaron de alguna forma para culminarlo exitosamente. A la Escuela Superior Politécnica del Litoral por su exigencia que permitió siempre dar lo mejor de mí.

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, me corresponde conforme al reglamento de propiedad intelectual de la institución; *Joseph Jonathan León Cando* doy mi consentimiento para que la ESPOI realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”



Joseph Jonathan León Cando

EVALUADORES



PhD. Sandra García

PROFESOR DE LA MATERIA



PhD. Maria Nela Pastuzaca

PROFESOR TUTOR

RESUMEN

La calidad en los productos de una empresa, ya sean estos bienes o servicios, es un factor muy importante que incide directamente en la satisfacción de los clientes. El controlar de manera eficiente la calidad ayuda además a evitar costos por reprocesos de productos deficientes o por eliminación de productos defectuosos. Una de las herramientas más utilizadas en el control estadístico de procesos, a fin de monitorear la calidad con que se elabora un producto es por medio de gráficos de control, esta herramienta permite monitorear simultáneamente varias de las características del producto y detectar cuando estas características están fuera de los valores normales. En este proyecto se ha desarrollado dos nuevos gráficos de control multivariante cuyos parámetros de control han sido optimizados utilizando algoritmos genéticos. Estos gráficos se basan en una variación del gráfico T^2 de Hotelling usando los esquemas de muestreo llamados Multiple dependent state sampling (MDS) y Generalized multiple dependent state sampling (GMDS). Los nuevos gráficos fueron comparados con el T^2 de Hotelling utilizando el ARL fuera de control, observándose que ambos gráficos son más eficientes para detectar un proceso fuera de control. Finalmente se utilizó uno de los nuevos gráficos en la resolución de un problema real en una compañía.

Palabras Clave: Algoritmo genético, gráfico de control, multiple dependent state, generalized multiple dependent state, T^2 Hotelling.

ABSTRACT

The quality of the products in an organization, whether goods or services, is a very important factor that influence directly in the customer satisfaction. The efficient control of the quality help to avoid costs of reprocessing deficient products and costs of discard defective products. The use of control charts is a way to monitor the quality which a product is manufactured, this tool let to monitor simultaneously some product's quality characteristics and detect when these characteristics are out of optimal values. Two new multivariate control charts have been developed in this project whose control params has been optimized using genetic algorithms. These charts are based on a variation of Hotelling T^2 using the schemes called Multiple dependent state sampling (MDS) and Generalized multiple dependent state sampling (GMDS). The new control charts were compared with Hotelling T^2 using out-of-control ARL and they showed more efficiency to detect an out of control process. Finally, one of the new control charts was used to solve a real problem in a company.

Keywords: Genetic algorithm, control chart, multiple dependent state, generalized multiple dependent state, Hotelling T^2

ÍNDICE GENERAL

RESUMEN	6
ABSTRACT	7
ÍNDICE GENERAL	8
ABREVIATURAS	10
ÍNDICE DE FIGURAS	11
ÍNDICE DE TABLAS	12
CAPÍTULO 1	13
1. INTRODUCCIÓN	13
1.1 Descripción del problema	14
1.2 Justificación del problema.....	14
1.3 Objetivos	15
1.4 Marco teórico.....	16
CAPÍTULO 2.....	27
2. Metodología	27
2.1 Esquema MDS	28
2.2 Esquema GMDS.....	28
2.3 Problema de optimización.....	29
CAPÍTULO 3.....	32
3. Resultados	32
3.1 Rendimiento del gráfico MDS T^2	32
3.2 Rendimiento del gráfico GMDS T^2	35
3.3 Análisis de sensibilidad.....	37
3.4 Implementación en una empresa.....	42
CAPITULO 4.....	46

4.	Conclusiones y recomendaciones	46
4.1	Conclusiones	46
4.2	Recomendaciones	47
BIBLIOGRAFÍA		¡Error! Marcador no definido.

ABREVIATURAS

MDS	Multiple dependent state sampling
GMDS	Generalized multiple dependent state sampling
ARL	Longitud promedio de corridas
ARL ₀	ARL bajo control deseado
ARL ₁	ARL fuera de control
ARL(d=0)	ARL bajo control obtenido por el software
ARL(d=d*)	ARL fuera de control obtenido por el software para un cambio d *
LCL	Límite de control inferior
LCS	Límite de control superior
LCS ₁	Límite de control superior interno
LCS ₂	Límite de control superior externo

ÍNDICE DE FIGURAS

Gráfico 1.1 Pasos en la ejecución de un algoritmo genético.	24
Grafico 3.1 Respuesta del software para la implementación del grafico MDS en una empresa.....	45
Grafico 3.2 Respuesta del software para la implementación del grafico GMDS en una empresa.....	46

ÍNDICE DE TABLAS

Tabla 3.1 Escenarios contruidos para el análisis del desempeño para los gráficos de control propuestos.....	33
Tabla 3.2 Desempeño del gráfico MDS T^2 para cada uno de los escenarios	34
Tabla 3.3 Desempeño del gráfico GMDS T^2 para cada uno de los escenarios.....	36
Tabla 3.4 Casos a considerar en el análisis de sensibilidad.....	37
Tabla 3.5 Parámetros óptimos para los casos a estudiar en el análisis de sensibilidad	38
Tabla 3.6 ARL fuera de control del gráfico MDS T^2 optimizado para un cambio 0.5	39
Tabla 3.7 ARL fuera de control del gráfico MDS T^2 optimizado para un cambio 0.8	39
Tabla 3.8 ARL fuera de control del gráfico MDS T^2 optimizado para un cambio 3	39
Tabla 3.9 ARL fuera de control del gráfico MDS T^2 optimizado para un cambio 3.5	40
Tabla 3.10 Parámetros óptimos para los casos a estudiar en el análisis de sensibilidad	40
Tabla 3.11 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 0.5.....	41
Tabla 3.12 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 0.8.....	41
Tabla 3.13 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 3	42
Tabla 3.14 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 3.5.....	42

CAPÍTULO 1

1. INTRODUCCIÓN

En este proyecto se presenta una herramienta para la mejora de la calidad de los productos de una empresa, ya sean estos productos bienes o servicios. La calidad se ve afectada comúnmente por temas relacionados a la materia prima, la maquinaria y/o la mano de obra; con lo cual, el control de cada uno de estos aspectos resulta importante para garantizar una buena calidad del producto final. Una maquina mal calibrada, piezas muy desgastadas de los equipos, una materia prima que no cumple los estándares de calidad, un personal poco capacitado son posibles causas que influyen de manera negativa en la calidad y que deberían ser detectadas de forma oportuna.

Dentro de una organización existen varios procesos que intervienen en la elaboración de un producto. En estos es muy común utilizar herramientas para vigilar las características que están relacionadas a su calidad. Con la herramienta que se presenta en este proyecto se pretende realizar un aporte a la mejora de la calidad. Para esto, la nueva técnica realiza un monitoreo simultaneo de las características mencionadas, y de forma eficiente detecta cuando ha ocurrido un cambio en el proceso que afecte a estas características. Cabe resaltar la importancia de realizar un monitoreo simultaneo de estas características o variables, ya que, si estas están relacionadas y son analizadas de forma independiente, los resultados podrían no reflejar lo que realmente está ocurriendo en el proceso.

Por otro lado, la falta de control en la calidad del producto que ofrece una compañía, con frecuencia genera costos que disminuyen la rentabilidad y eficiencia de esta. Con la metodología propuesta se pretende hacer una detección temprana cuando surja un problema en los procesos, para así realizar las debidas correcciones y ajustes evitando costos adicionales y perdidas por elaboración de productos defectuosos que impliquen un reproceso o su eliminación.

Para conseguir lo antes mencionado, se propone una nueva técnica para el control estadístico de procesos basada en un gráfico de control que utiliza información anterior para la detección de anomalías considerando a las características de calidad como variables continuas que siguen una distribución normal.

1.1 Descripción del problema

La empresa en su giro de negocio cuenta con múltiples procesos, el control de cada uno de estos permite a la compañía conocer los parámetros correctos de funcionamiento, determinar el estado en que se encuentran y detectar cuando un proceso es deficiente o no está funcionando correctamente.

No controlar un proceso implica no identificar a tiempo cuando es necesaria una corrección o mejora de este, ni determinar si el producto elaborado cumple con los estándares de calidad propuestos por la empresa o exigidos por alguna normativa.

Por otro lado, la deficiencia o mal funcionamiento de un proceso conlleva a las empresas a elaborar un producto de baja calidad, lo que compromete la satisfacción de los clientes, induciendo a los mismos a la búsqueda de estos productos en la competencia y generando así una potencial pérdida de ventas. Es por esto por lo que una detección temprana ayudaría a la empresa a generar ahorros y mantener la calidad de su producto.

1.2 Justificación del problema

Un proceso deficiente puede generar para la compañía gastos por desperdicios de materiales o por reprocesos de los productos afectando así a los costos de producción, a la rentabilidad del negocio y a su vez a la competitividad.

Parra (2016) indica lo siguiente “se hace énfasis en la calidad, porque los costos asociados con la calidad son considerables. Los costos de calidad varían de 15 a 20% de los ingresos de ventas de muchas organizaciones, y con un buen programa de calidad, se pueden obtener ahorros sustanciales e ingresos más elevados”.

En base a lo anterior mencionado, resalta la importancia de contar con una herramienta que realice un monitoreo de los procesos e indique cuando se produzca una falla para así corregirlo y evitar los costos que pudiera generarse.

Una herramienta que nos permite realizar lo anterior dicho son los gráficos de control, que son una de las 7 herramientas de calidad definidas por Ishikawa. Estos gráficos permiten conocer de manera rápida y directa si un proceso está funcionando correctamente y debido a su sencillez pueden ser fácilmente interpretados por alguna persona sin mucho conocimiento estadístico.

1.3 Objetivos

1.3.1 Objetivo General

Desarrollar un eficiente gráfico de control multivariante, que permita identificar cuando un proceso necesita ser corregido o mejorado a partir de un pequeño tamaño de muestra, utilizando para esto las características de calidad consideradas más importantes dentro del proceso.

1.3.2 Objetivos Específicos

- Desarrollar un gráfico de control que mejore la eficiencia de los existentes en la literatura mediante la utilización de nuevos parámetros de control.
- Comparar el rendimiento del gráfico de control propuesto con el gráfico de control de Hotelling.

- Desarrollar un software que permita calcular los parámetros del gráfico de control basado en un ARL bajo control, minimizando el ARL fuera de control dado un cambio específico en el proceso.
- Elaborar un análisis de sensibilidad del gráfico de control propuesto.

1.4 Marco teórico

1.4.1 Contraste de hipótesis

Un contraste de hipótesis, dentro del contexto estadístico, es un procedimiento para determinar si un parámetro supuesto en una población cumple o no con lo observado en una muestra de dicha población, es decir, el contraste servirá como método que analiza si el supuesto debe ser rechazado o no. Los elementos que intervienen en un contraste de hipótesis son los siguientes (Wackerly et al. 2009):

- **Hipótesis nula (H_0):** es la hipótesis que se requiere contrastar.
- **Hipótesis alternativa (H_1):** hipótesis que se enfrenta a H_0 , y que es en la mayoría de casos la hipótesis que se desea apoyar con los datos de la muestra.
- **Estadístico de prueba:** es el estadístico obtenido de la muestra donde su distribución de probabilidad está asociada a la hipótesis del contraste.
- **Región de rechazo:** Es la región formada por los valores que son más improbables bajo la hipótesis planteada.

Errores en un contraste de hipótesis

Considerando que no se conoce el valor real de los parámetros de la población, sino que fueron planteados supuestos y que el criterio utilizado en un contraste de hipótesis tiene como resultado el rechazar o no el supuesto planteado, podemos notar que se pueden generar los siguientes errores:

- Error tipo I: Es el error que ocurre cuando se rechaza la hipótesis nula H_0 dado que es verdadera. Este error también es llamado falso positivo y tiene una probabilidad de ocurrencia α .
- Error tipo II: Es el error que ocurre cuando no se rechaza la hipótesis nula H_0 dado que es falsa. Este error también es llamado falso negativo y tiene una probabilidad de ocurrencia β .

1.4.2 Gráficos de control

El gráfico de control es una herramienta que nos permite detectar variaciones en el comportamiento de los procesos, evaluar la estabilidad, determinar un ajuste de ser necesario y confirmar la mejora en el mismo (Besterfield, 2009).

1.4.2.1 Gráfico de control por variables

En los gráficos de control por variables se hacen uso de estadísticos obtenidos a partir de datos cuantitativos. En este tipo de gráfico se puede llegar a medir la característica de calidad a estudiar debido a que estas características son susceptibles a medición, razón por la cual son capaces de “avisarnos” de posibles problemas de calidad incluso antes de que éstos sean ya relevantes.

1.4.2.2 Gráfico de control por atributos

Este tipo de gráfico se usa para contrastar características cualitativas. Los gráficos de control por atributos tienen la ventaja de sintetizar de forma rápida toda la información referida a diferentes aspectos de calidad de un producto, ya que permiten clasificar éste como aceptable o inaceptable; además, no suelen necesitar de sistemas de medición muy complejos y son más fácilmente entendibles por los no especialistas.

1.4.3 ARL: Medida de rendimiento de los gráficos de control

El ARL (Average Run Length) es el número de puntos que se deben graficar en una carta de control antes de que un punto indique una salida de control (Montgomery, 2006). El ARL, para cualquier gráfico de control Shewhart, está dado por la media de una distribución geométrica y se calcula de la siguiente forma:

$$ARL = \frac{1}{p} \quad (1.1)$$

Siendo p la probabilidad de que un punto se encuentre por fuera de los límites de control.

1.4.3.1 ARL bajo control (ARL₀)

Es el número promedio de muestras hasta que el estadístico tome un valor fuera de los límites de control, considerando que el proceso se encuentra bajo control. Está dado como $1/\alpha$ donde α es la probabilidad de que el gráfico emita una falsa alarma (error tipo I).

1.4.3.2 ARL fuera de control (ARL₁)

Es el número promedio de muestras hasta que el estadístico tome un valor fuera de los límites de control, considerando que el proceso se encuentra fuera de control. Está dado como $1/(1 - \beta)$ donde β es la probabilidad de que el gráfico no detecte un cambio (error tipo II).

Lo que se desea en un gráfico de control es que el valor de ARL₀ sea alto y el valor de ARL₁ sea pequeño, para así detectar de forma rápida cambios que afecten al proceso.

1.4.4 Gráfico de control multivariado T² de Hotelling

Harold Hotelling (1947) definió el estadístico T^2 para un grupo de variables aleatorias provenientes de una distribución normal multivariada de la siguiente forma:

$$T^2 = n(\bar{X} - \hat{\mu})^T S^{-1}(\bar{X} - \hat{\mu}) \quad (1.2)$$

Siendo \bar{X} el vector de medias muestrales. Dado que en la práctica muy a menudo no se conocen los valores de los parámetros, Hotelling uso S y $\hat{\mu}$ como estimadores de la matriz de varianzas y covarianzas y del vector de medias.

Los límites de control para este gráfico estarán dados de la siguiente manera: límite de control superior $LCS = \chi_{\alpha,p}^2$ y como límite de control inferior $LCI = 0$, donde α el nivel de significancia y p la cantidad de atributos de calidad que serán monitorizados.

Al existir un cambio en alguna de las medias de las variables del proceso, el vector μ_1 se aleja del vector de medias μ_0 que se encuentra bajo control. La distancia entre μ_0 y μ_1 esta dada por la distancia de Mahalanobis (1936) d , con lo que 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 fórmula matemática:

$$\lambda = nd^2 = n(\mu_1 - \mu_0)^T \Sigma^{-1}(\mu_1 - \mu_0) \quad (1.3)$$

Siendo n la cantidad de observaciones en la muestra y Σ la matriz de varianzas y covarianzas de X cuando el proceso se encuentra bajo control.

El ARL definido para este tipo de gráfico está dado de la siguiente forma:

$$ARL = \frac{1}{1 - P(T^2 < LCS)} \quad (1.4)$$

1.4.5 Multiple dependent state sampling (MDS)

MDS es una técnica de muestreo ampliamente usada en el control estadísticos de procesos. Este esquema fue presentado en sus inicios por Wortham and Baker (1976) y utilizado en el monitoreo de la media de un proceso univariado bajo una distribución normal por Aslam et al. (2014). La técnica consiste en decidir si una muestra es aceptada o rechazada no sólo basándose en la información extraída de la muestra actual, sino que el esquema considera información obtenida de muestras anteriores. La finalidad de este método es utilizar la información de muestras secuencialmente anteriores con el objetivo de reducir el número de muestras necesarias para detectar un cambio y por ende los gastos que implican tomarlas. Existen varios investigadores que han aplicado esta técnica en sus estudios incluyendo entre estos Balamurali y Jun (2006), Mogg and Wortham (1970).

El procedimiento de Multiple dependent state sampling está denotado por MDS (LCS_1, LCS_2, m) y considera que el estadístico no puede tomar valores menores que cero. Los pasos son los siguientes:

- Para un lote t se selecciona una muestra de tamaño n y se calcula el estadístico C .
- Se acepta la muestra si $0 < C < LCS_1$. Se rechaza la muestra en caso de que $C > LCS_2$. Caso contrario para $LCS_1 < C < LCS_2$, la muestra es aceptada si las m muestras anteriores fueron aceptadas, si alguna fue rechazada entonces la muestra actual es rechazada.
- Se incrementa t y se continua con el siguiente lote desde el paso 1.

Donde:

C: es el estadístico que se desea monitorear.

LCS₁: es el límite de control superior interno que indica el valor máximo que puede tomar el estadístico para ser aceptado sin condición.

LCS₂: es el límite de control superior externo que indica el valor máximo que puede tomar el estadístico para no ser rechazado.

m: es el número de muestras anteriores a considerar para aceptar o rechazar.

El estadístico a usar dependerá de la característica de calidad que se requiera monitorear. Para este caso se ha considerado que el estadístico no puede tomar valores menores que cero, es decir, el límite de control inferior $LCI = 0$.

1.4.6 Generalized multiple dependent state sampling (GMDS)

GMDS es una técnica de muestreo la cual es una generalización de MDS. Fue propuesta en el diseño de gráficos de control para una distribución Poisson multivariante por Ali Raza, Aslam (2018), su objetivo es mejorar un gráfico de control multivariante utilizando dicho estadístico. En ese estudio se propuso un gráfico de control para el conteo de defectos que utiliza una distribución Poisson multivariante (PM) y la técnica MDS y otro gráfico igual utilizando GMDS. Estos gráficos fueron comparados con los gráficos de control existentes, para medir sus rendimientos y lograr la comparación se utilizó el ARL_1 que es una medida ampliamente usada en la comparación de gráficos de control.

Los resultados revelaron que ambos gráficos de control propuestos en el trabajo de Ali Raza y Aslam (2018) detectan de mejor manera los cambios abruptos en comparación a los gráficos de control PM existentes en ese momento. Por otro lado, se resalta la superioridad de el grafico GMDS sobre los otros gráficos de control al ser este más sensible.

El esquema GMDS consigue mejorar el rendimiento de los gráficos de control tipo Shewhart agregando un nuevo parámetro al esquema MDS, es decir, ambas técnicas incluyen un límite control superior externo LCS₂, un límite de

control superior interno LSC_1 y utilizan información pasada para terminar si el estadístico C de la actual muestra se encuentra dentro de los límites de control. La técnica GMDS consiste en lanzar una señal fuera de control cuando para al menos k de m anteriores muestras el estadístico C toma valores entre:

$$LCS_1 < C < LCS_2 \quad (1.5)$$

Donde los valores de k y m son seleccionados por especialistas en el área de control de calidad según la sensibilidad que el proceso requiera para ser monitoreado.

1.4.7 Algoritmos genéticos (GA)

Los algoritmos genéticos son métodos de búsqueda heurística inspirados en la teoría de evolución natural de Charles Darwin, son utilizados en la optimización ya que proporcionan soluciones más robustas comparadas con otras técnicas de búsqueda como programación lineal, heurística, búsqueda en profundidad y praxis. Estos algoritmos pueden encontrar solución a un sin número de problemas aplicados a diferentes áreas.

Charles Darwin en su libro “El origen de las especies” del año 1859 destaca “la supervivencia del más apto”, es así que estos algoritmos simulan el proceso de selección natural seleccionando los individuos más fuertes y aptos con el objetivo de reproducirse y dejar descendencia. Basándose en esto los GA son capaces de generar soluciones con valores óptimos para un problema.

En los años 50, Alex S. Fraser realizó algunas investigaciones referentes a la evolución de sistemas biológicos en una computadora digital, esto sirvió como base para el nacimiento de los algoritmos genéticos. Pero no fue hasta 1975 donde John Henry Holland asentó los principios básicos de GA, seguido

luego por varios investigadores como Goldberg (1989), Davis (1991), Michalewicz (1992), Reeves (1993).

Los GA utilizan técnicas adaptivas que permiten encontrar soluciones para problemas de búsqueda y optimización, sus principios se encuentran en el proceso genético de los seres vivos, son capaces de encontrar soluciones que a través de iteraciones evolucionan hacia valores más óptimos del problema (Holland, 1992).

Sin embargo, los GA tienen algunas limitaciones:

- Para garantizar un máximo o mínimo global, el algoritmo requerirá más interacciones si la función a optimizar cuenta con muchos máximos o mínimos locales.
- No se puede garantizar encontrar una solución óptima cuando esta presenta varios puntos cercanos en la función objetivo, en estos casos el algoritmo encontrará algún punto de ellos y no precisamente el óptimo.

La representación de utilizada en los GA es la siguiente:

- Individuos: son las posibles soluciones del problema.
- Genes: es la agrupación de parámetros.
- Cromosoma: se refiere al conjunto formado por genes los cuales utilizan codificación binaria que presenta mejores resultados que otros métodos de codificación.
- Población: conjunto de posibles soluciones

La metodología utilizada por los GA es expuesta a continuación:

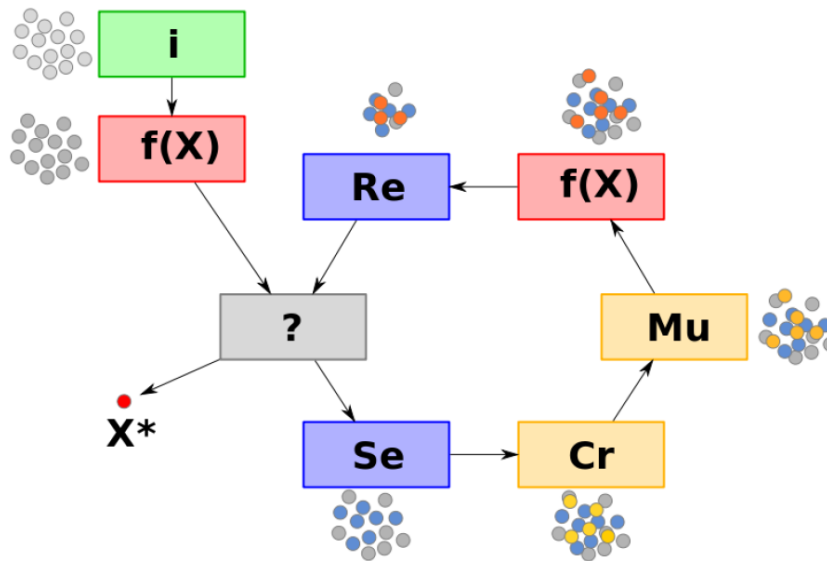


Gráfico 1.1 Pasos en la ejecución de un algoritmo genético.

Fuente: Wikipedia. Algoritmo Genético, 2019.

- 1) Iniciación(i): En este paso se inicia creando agrupaciones de cromosomas, es decir, grupos de individuos que van a ser aleatorios.

Los siguientes pasos se ejecutarán siempre que el número de iteraciones sea menor al número de iteraciones definido en el algoritmo genético o a su vez cuando no existe ningún cambio en el grupo de individuos:

- 2) Evaluación $f(X)$: en esta fase, con el fin de conocer que tan buena es la solución que ofrece cada individuo, el algoritmo evaluará cada cromosoma utilizando una función de aptitud.
- 3) Selección (Se): se procederá a seleccionar los individuos de acuerdo con el puntaje obtenido en la fase anterior, los individuos mejores puntuados tendrán mayores probabilidades de resultar seleccionados. Para esto existen varias técnicas de selección, entre las más usadas podemos mencionar: Por ruleta, por torneo o por rango.
- 4) Cruzamiento (Cr): En esta etapa dos cromosomas generan dos descendientes los cuales contienen propiedades heredadas de los

cromosomas padres, esta fase representa la reproducción de los individuos. Existen varios tipos de cruzamiento entre estos están: Uniforme, aritmético, un punto y dos puntos.

- 5) Mutación (Mu): Consiste en realizar al alzar modificaciones del cromosoma de los individuos, esto permitirá considerar nuevas opciones y a su vez analizar espacios no cubiertos de la población objetivo.
- 6) Reemplazo (Re): Una vez culminado los pasos anteriores, se elegirán a los individuos mejores de un grupo, para así crear la próxima generación de cromosomas.

Es importante recordar que, existen variaciones en las operaciones de genética dentro de los individuos, esto provocará que se genere una población distinta para cada variación.

1.4.8 Estado del arte

Son muchos los investigadores que han aportado con sus ideas y estudios en el área del control multivariado de procesos entre ellos podemos mencionar la revisión realizada en la aplicación de técnicas de monitoreo estadístico multivariado de procesos en los campos no industriales discutido en Bersimis, Sgora and Psarakis (2016), donde se realizó una extensa discusión y se resalta la necesidad de un estudio más profundo de las fórmulas matemáticas ya existentes y de la mejora y extensión de las técnicas establecidas. Además, afirma que existen aún varios desafíos que tiene el monitoreo de procesos, entre estos se encuentran la autocorrelación, donde los datos se encuentran relacionados con el tiempo, la alta calidad de los procesos, que es necesaria en la supervisión de la salud y la correcta interpretación de las señales fuera de control, ya que en muchos de los casos se ve la necesidad no solo de usar técnicas de monitoreo estadísticos de procesos, sino de combinarlas con técnicas estadísticas avanzadas.

Vanhatalo, E. y Kulahci M, en efecto corroboraron la importancia de la autocorrelación, en este caso para un gráfico de control de Hotelling (Vanhatalo & Kulahci, 2015). Ellos llegaron a la conclusión que la mayoría de veces que los datos son recaudados de los procesos, estos se encuentran correlacionados con el tiempo y con la propia dinámica del proceso. Los investigadores afirman que es posible que la autocorrelación llegue a afectar la probabilidad de que el gráfico de control indique una falsa alarma, incluso llegaría a afectar su capacidad de detectar cambios. En la investigación lograron determinar que el no considerar las autocorrelaciones, puede generar conclusiones incorrectas dado que el ARL bajo control puede diferir significativamente del nominal.

Un gráfico de control de Hotelling con tamaño de muestras variables (Aparisi, 1996) fue propuesto para mejorar el gráfico de control de Hotelling en la detección de pequeños cambios en los procesos. En este estudio se utilizaron las técnicas empleadas en el caso de gráficos de control univariados para tamaño de muestras variables. La investigación encuentra una mejora en los ARL fuera de control para el caso de cambios pequeños, sin embargo, cuando los cambios son grandes y el número de variables incrementa, el ARL pierde ligeramente su capacidad de detección. Aparisi además del ARL fuera de control también consideró el número de piezas promedio a ser inspeccionadas como medida de comparación. Esta medida es un parámetro importante en el control de calidad debido a que, si el número de piezas aumenta, los costos implicados en el control de calidad también aumentan. Considerando esta medida se encontró de nuevo que el esquema propuesto es mucho mejor para cambios pequeños debido a que son necesarios muy pocas muestras hasta el gráfico emita una señal fuera de control.

Consideremos también el estudio de un gráfico de control multivariado de medias móviles exponenciales ponderadas (Cynthia A. Lowry, William H.

Woodall , Charles W. Champ & Steven E. Rigdon, 1992), esta versión multivariada del gráfico fue realizada, como su nombre lo indica, a partir de un gráfico de control univariado de medias móviles exponenciales ponderadas (EWMA). En esta investigación adicionalmente se presentan guías para implementar procedimientos multivariados e indica como realizar una correcta interpretación de las señales fuera de control. Por medio del ARL, se realizó una comparación de rendimientos del gráfico propuesto encontrándose similitudes al gráfico de control multivariado de sumas acumuladas (CUSUM) en cuanto a detectar un cambio en las medias de una distribución normal multivariante. Sin embargo, el procedimiento utilizando MEWMA tiene un mejor rendimiento que el procedimiento CUSUM multivariado solo cuando el proceso inicialmente se encuentra fuera de control. De igual forma que en los gráficos de T^2 Hotelling y CUSUM multivariado, el rendimiento del gráfico medido por medio del ARL depende del vector de medias y matriz de covarianzas solo a través del valor del parámetro de no centralidad. Los autores recomiendan que se debe utilizar siempre el gráfico de control de T^2 Hotelling en conjunto con los gráficos CUSUM y EWMA multivariado para evitar problemas que se pueden originar en ciertos escenarios.

CAPÍTULO 2

2. Metodología

En esta sección se detalla el procedimiento que se siguió para la elaboración del gráfico de control basado en el esquema de muestreo MDS y el gráfico de control basado en el esquema GMDS asumiendo que las variables de interés siguen una distribución normal multivariante p -variada.

Sea $X_i (i = 1, 2, \dots, p)$ variables aleatorias que siguen conjuntamente una distribución normal multivariante p -variada con un vector de media μ y una matriz

de covarianza Σ y que luego de una estandarización de estas variables se obtiene $Z_i \sim N(\vec{0}, R)$, donde R es la matriz de correlación de las variables X_i y considerando el tamaño de la muestra igual a uno, entonces el estadístico a utilizar en los gráficos de control propuesto es el T^2 de Hotelling que está dado por:

$$T^2 = n(Z - \vec{0})^T R^{-1}(\bar{X} - \vec{0})$$

La distribución de T^2 se aproxima a una distribución Chi- cuadrada χ_p^2 donde p son los grados de libertad (Seber, 1984). Los límites de control LCS_1 y LCS_2 para cada gráfico serán hallados a partir del algoritmo genético y dado que la distribución Chi-cuadrada no puede tomar valores negativos entonces $LCL = 0$.

2.1 Esquema MDS

En el esquema MDS cuando el estadístico C de la muestra actual se ubica en $C > LCS_2$ el proceso estará fuera de control. Si el estadístico C se encuentra $C < LCS_1$ el proceso se encuentra dentro de control. Cuando $LCS_1 < C < LCS_2$ dicha muestra será aceptada si los valores de los estadísticos de las m anteriores muestras se encuentran dentro de $0 < C < LCS_1$. La probabilidad que el proceso esté dentro de control estará dada por:

$$P_0^1 = P(\chi^2 < LCS_1) + P(LCS_1 < \chi^2 < LCS_2) * P(\chi^2 < LCS_1)^m \quad (2.1)$$

El cálculo del ARL bajo control se lo realizará de la siguiente forma:

$$ARL_0 = \frac{1}{1 - P_0^1} \quad (2.2)$$

2.2 Esquema GMDS

En este esquema es agregado un parámetro adicional k a los parámetros existentes en el esquema MDS. Entonces si el estadístico C de la muestra

actual se ubica $C > LCS_2$ el proceso estará fuera de control. Si el estadístico C se encuentra $C < LCS_1$ el proceso se encuentra dentro de control. Cuando $LCS_1 < C < LCS_2$ el proceso será declarado bajo control cuando al menos k de las m muestras anteriores el estadístico C se encuentra dentro de $0 < C < LCS_1$, caso contrario se dirá que el proceso está fuera de control. Los valores de k y m serán encontrados por medio del algoritmo genético.

Para este esquema el proceso se encontrará bajo control con la siguiente probabilidad:

$$P_0^2 = P_{C2} + P_{C1} * \sum_{j=k}^m \binom{m}{j} P_{C2}^j (1 - P_{C2})^{m-j} \quad (2.3)$$

Donde:

$$P_{C2} = P(\chi^2 < LCS_1)$$

$$P_{C1} = P(LCS_1 < \chi^2 < LCS_2)$$

Así el ARL bajo control quedaría definido de la siguiente forma:

$$ARL_0 = \frac{1}{1 - P_0^2} \quad (2.4)$$

2.3 Problema de optimización

Los parámetros de los gráficos de control propuestos en este proyecto fueron optimizados utilizando algoritmos genéticos. Estos algoritmos fueron implementados mediante el software Visual Basic Community, en el cual se desarrolló una interfaz de usuario que permite ingresar los parámetros del algoritmo y como respuesta nos devuelve ya optimizados los parámetros de los gráficos de control.

Los parámetros de ajustes para el algoritmo genético fueron los siguientes:

- **Métodos de selección:** torneo y ruleta
- **Tipo de cruzamiento:** un punto, dos puntos y uniforme
- **Probabilidad de ocurrencia:** 0.99
- **Probabilidad de mutación:** 0.03
- **Funciones de aptitud:** las funciones utilizadas por Aparisi y García-Díaz (Aparisi & García-Díaz, 2007).

El método de selección por torneo consiste en seleccionar aleatoriamente una específica cantidad de individuos con el objetivo de reproducir al individuo más apto entre estos.

Por otro lado, el método de selección por ruleta se realiza creando una ruleta de cromosomas para cada región, siendo los cromosomas con mayor probabilidad de ser escogidos los que tienen la puntuación más alta.

EL tipo de cruzamiento de 1 punto se realiza a través del corte de dos cromosomas en un determinado punto, esto generará un individuo nuevo el cual su cromosoma estará conformado por la primera parte de un padre y la segunda parte por el otro padre. En cambio, en el tipo de cruzamiento de 2 puntos, el cromosoma es dividido en dos partes, pero ahora un progenitor aporta su cromosoma inicial y su cromosoma final, es decir, aporta dos partes, mientras que el segundo progenitor aporta con la parte media de su cromosoma, es decir, solo aporta una parte.

En el cruzamiento uniforme los progenitores aportan a sus hijos sus genes al azar (Holland, 1992).

Las funciones de aptitud utilizadas para este algoritmo, fueron las recomendadas por Aparisi y García-Díaz (Aparisi & García-Díaz, 2007) que aplican el concepto de diseño óptimo en un punto. El método consistió en

proporcionar como parámetro el ARL dentro de control que se desea obtener y un vector de cambios d^* que se requiere detectar.

Es así que se logra obtener el conjunto de parámetros del diseño del gráfico de control que produce el mínimo ARL fuera de control $ARL(d = d^*)$ sujeto a un valor de ARL dentro de control deseado. Las funciones de actitud utilizadas son las siguientes:

$$\begin{aligned} \max & -w_1 |ARL_0 - ARL(d = 0)| - w_2 ARL(d = d^*) \\ \text{si } & (ARL_0 - ARL(d = 0)) \leq 0 \quad (1.6) \end{aligned}$$

$$\begin{aligned} \max & -2w_1 |ARL_0 - ARL(d = 0)| - w_2 ARL(d = d^*) \\ \text{si } & (ARL_0 - ARL(d = 0)) > 0 \quad (1.7) \end{aligned}$$

En donde:

ARL_0 : es el ARL deseado

$ARL(d = 0)$: es el ARL bajo control calculado a partir de los parámetros del gráfico evaluado.

$ARL(d = d^*)$: representa el ARL que detecta el cambio d^* para el cual se optima el gráfico.

w_1, w_2 : son los pesos de cada objetivo a optimizar para cada una de las ecuaciones. En el algoritmo genético hemos fijado estas ponderaciones en 1 y 5 como lo indica los trabajos de Aparisi y García-Díaz.

El problema por resolver está definido de la siguiente forma:

Dado:

ARL_0 : Número promedio de muestras hasta que el estadístico toma un valor fuera de los límites, cuando el proceso se encuentra dentro de control.

p : Número de variables a controlar.

d : Distancia de Mahalanobis de las p variables a controlar calculado a partir de un cambio en dichas variables.

Encontrar:

LCS_1 : Límite de control superior interno del gráfico.

LCS_2 : Límite de control superior externo del gráfico.

m : Número de muestras anteriores a analizar.

k : Número mínimo de muestras debajo de LCS_1 para aceptar la muestra actual (esquema GMDS)

ARL_1 : ARL fuera de control, valor a ser minimizado sujeto al ARL_0 .

CAPÍTULO 3

3. Resultados

En esta sección del estudio se compara los rendimientos de los gráficos propuestos con el gráfico T^2 de Hotelling, además los resultados son usados para calcular un gráfico de control óptimo para el monitoreo de un proceso de producción de harina de pescado dentro de una fábrica.

3.1 Rendimiento del gráfico MDS T^2

El rendimiento del gráfico MDS es comparado con el gráfico T^2 de Hotelling mediante los valores de ARL fuera de control que son obtenidos a partir de diferentes escenarios.

Los escenarios fueron construidos considerando combinaciones de ARL bajo control deseados, número de variables a monitorear y valores de los cambios a optimizar. Los escenarios que fueron utilizados se encuentran en la siguiente tabla:

Tabla 3.1 Escenarios contruidos para el análisis del desempeño para los gráficos de control propuestos.

Fuente: Christian Naranjo (2019)

# Escenario	ARL deseado (ARL ₀)	# de variables (p)	Distancia de Mahalanobis(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

La tabla 3.2 presenta los resultados de optimizar el gráfico de control que utiliza el esquema MDS para cada uno de los escenarios contruidos, además se muestra en la columna Diferencia T² el porcentaje de mejora comparado con el gráfico de Hotelling. En la tabla se observa que, para todos los escenarios contruidos, el gráfico control MDS obtiene un mejor desempeño que el grafico T² de Hotelling, mostrando mejor rendimiento cuando el gráfico es optimizado para cambios grandes, es el caso del escenario 24 donde el desempeño del nuevo gráfico es mejor en un 60.74%. Además, se evidencia que los ARL fuera de control más pequeños se dan cuando la distancia de Mahalanobis es alrededor de 2, así lo observamos en los escenarios 3, 9, 14, 19, 24, 30 donde los rendimientos del gráfico MDS

son mejores que el gráfico de Hotelling en más del 45%. Por otro lado, notamos rendimientos un poco más bajos cuando los cambios son pequeños, por ejemplo, el escenario 26 donde la mejora es solo del 7.01%. Adicional se evidencia que, cuando se aumenta la cantidad de variables el rendimiento del gráfico disminuye ligeramente, así se observa en los escenarios 1, 6 y 11 donde la mejora en el rendimiento ha ido disminuyendo ligeramente en 12.22%, 12.04% y 9.46% respectivamente. Es importante mencionar que a medida que los cambios se hacen más grandes, los ARL fuera de control se hacen mucho más pequeños.

Tabla 3.2 Desempeño del gráfico MDS T^2 para cada uno de los escenarios

Fuente: Elaboración propia

# Escenario	ARL ₀	# Var	Distancia Mahalanobis (d)	ARL ₁ T^2	MDS				
					LCS ₂	LCS ₁	m	ARL ₁	Diferencia T^2 %
1	200	2	0.5	115.53	17.12	6.39	3	101.41	-12.22%
2	200	2	1	41.92	16.4	6.68	4	26.22	-37.45%
3	200	2	1.5	15.78	18.78	7.03	6	7.51	-52.41%
4	200	2	2	6.88	16.64	6.4	3	3.35	-51.31%
5	200	2	3	2.16	17.79	6.38	3	1.36	-37.04%
6	200	4	0.5	138.15	30.67	10.23	4	121.51	-12.04%
7	200	4	1	60.96	26.39	10.72	6	40.25	-33.97%
8	200	4	1.5	24.62	22.02	9.07	3	13.83	-43.83%
9	200	4	2	10.63	18.47	10.74	5	5.12	-51.83%
10	200	4	3	2.93	26.64	9.47	2	1.68	-42.66%
11	200	6	0.5	149.46	37.53	13.74	5	135.32	-9.46%
12	200	6	1	74.32	23.76	12.71	2	57.79	-22.24%
13	200	6	1.5	32.13	19.19	15.59	6	23.37	-27.26%
14	200	6	2	14.12	25.13	12.64	2	7.67	-45.68%
15	200	6	3	3.69	22.74	12.8	2	2.01	-45.53%
16	200	10	0.5	161.34	37.1	19.94	6	148.5	-7.96%
17	200	10	1	92.48	34.95	19.96	6	68.9	-25.50%
18	200	10	1.5	44.53	37.12	19.94	6	25.95	-41.72%
19	200	10	2	20.59	33.02	20	6	10.02	-51.34%

20	200	10	3	5.21	34.33	19.37	4	2.55	-51.06%
21	500	4	0.5	327.14	29.45	11.83	6	284.66	-12.99%
22	500	4	1	129.75	24.82	11.65	5	83.12	-35.94%
23	500	4	1.5	46.78	20.06	11.92	5	23.06	-50.71%
24	500	4	2	18.06	26.52	11.84	6	7.09	-60.74%
25	500	4	3	4.05	19.63	12.2	6	2.06	-49.14%
26	500	10	0.5	391.8	35.18	20.43	3	364.33	-7.01%
27	500	10	1	208.91	34.1	19.87	2	166.3	-20.40%
28	500	10	1.5	91.72	36.73	20.39	3	55.18	-39.84%
29	500	10	2	38.39	37.34	20.38	3	18.61	-51.52%
30	500	10	3	7.96	39.35	20.36	3	3.37	-57.66%

3.2 Rendimiento del gráfico GMDS T²

En la tabla 3.3 se encuentran los resultados de realizar la optimización del gráfico de control T² utilizando el esquema GMDS para los escenarios construidos. En la columna Diferencia T² se muestra en porcentaje la mejora del ARL fuera de control del nuevo gráfico comparado con el gráfico T². Se logra observar que para todos los escenarios construidos existe una mejora. Los mejores desempeños se obtienen cuando el gráfico es optimizado para cambios grandes tal como se muestra en los escenarios 29 y 30 donde el porcentaje de mejora son del 76% y 77% respectivamente. Adicional se evidencia que al aumentar el cambio d , los ARL fuera de control se hace mucho más pequeños, esto se observa en los escenarios 24 y 25 donde los ARL son 5 y 1 respectivamente. Analizando el otro extremo, tenemos que los desempeños más bajos se dan cuando el gráfico es optimizado para cambios pequeños, esto lo podemos observar en los escenarios 11 y 16 donde existe mejora, pero esta llega al 12% y 8% respectivamente. Además, los resultados muestran que al aumentar el número de variables los rendimientos disminuyen ligeramente en los cambios pequeños, lo mismo que fue evidenciado por Aparisi en su investigación en 1996, sin embargo, para algunos cambios grandes los rendimientos aumentan cuando aumenta el número de variables

como es el caso de los escenarios 5, 10 y 15 donde el rendimiento fue en aumento (45%, 61%, 64%).

Tabla 3.3 Desempeño del gráfico GMDS T² para cada uno de los escenarios

Fuente: Elaboración propia

# Escenario	ARL0	# Var	Distancia Mahalanobis (d)	ARL1 T ²	GMDS					
					LCS ₂	LCS ₁	m	k	ARL1	Diferencia T ² %
1	200	2	0.5	115.53	17.91	4	6	4	89.05	-22.92%
2	200	2	1	41.92	16.64	4.67	4	3	22.48	-46.37%
3	200	2	1.5	15.78	14.26	2.32	6	2	5.07	-67.87%
4	200	2	2	6.88	19.18	3.27	4	2	2.35	-65.84%
5	200	2	3	2.16	17.85	2.13	4	1	1.18	-45.37%
6	200	4	0.5	138.15	20.88	7.04	6	4	112.37	-18.66%
7	200	4	1	60.96	21.72	4.69	6	2	32.91	-46.01%
8	200	4	1.5	24.62	20.41	4.71	6	2	8.47	-65.60%
9	200	4	2	10.63	16.92	4.92	6	2	3.1	-70.84%
10	200	4	3	2.93	19.61	4.73	6	2	1.15	-60.75%
11	200	6	0.5	149.46	26.23	8.66	4	2	131.65	-11.92%
12	200	6	1	74.32	26.11	8.32	6	3	40	-46.18%
13	200	6	1.5	32.13	23.4	9.19	2	1	19.82	-38.31%
14	200	6	2	14.12	20.57	8.68	6	3	4.04	-71.39%
15	200	6	3	3.69	26.92	7.69	5	2	1.32	-64.23%
16	200	10	0.5	161.34	37.1	19.94	6	6	148.5	-7.96%
17	200	10	1	92.48	39.76	14.89	6	4	55.25	-40.26%
18	200	10	1.5	44.53	36.53	14.9	6	4	17.36	-61.02%
19	200	10	2	20.59	27.77	15.36	6	4	6.59	-67.99%
20	200	10	3	5.21	25.79	15.84	5	3	1.99	-61.80%
21	500	4	0.5	327.14	30.71	5.11	6	2	264.11	-19.27%
22	500	4	1	129.75	25.97	5.13	6	2	64.08	-50.61%
23	500	4	1.5	46.78	22.21	4	6	1	18.07	-61.37%
24	500	4	2	18.06	21.76	4.01	6	1	5.29	-70.71%
25	500	4	3	4.05	19.28	4.14	6	1	1.4	-65.43%
26	500	10	0.5	391.8	39.5	15.78	6	4	333.78	-14.81%
27	500	10	1	208.91	39.23	13.89	6	3	115.33	-44.79%

28	500	10	1.5	91.72	39.49	12.16	6	2	32.14	-64.96%
29	500	10	2	38.39	32.4	12.29	6	2	9.33	-75.70%
30	500	10	3	7.96	34.61	15.84	6	4	1.85	-76.76%

3.3 Análisis de sensibilidad

Los gráficos de control propuestos en este proyecto tienen como finalidad detectar de forma eficiente un cambio específico (d^*), dado esto, su desempeño cuando ocurra dicho cambio será muy bueno. Sin embargo, con el análisis de sensibilidad se pretende estudiar el desempeño de estos gráficos cuando el cambio que ocurrió fue diferente al cambio para los que fueron optimizados. De igual forma que en el análisis de rendimientos, el análisis de sensibilidad utiliza el ARL fuera de control para realizar la comparación en este caso de los desempeños de los gráficos de control para distintos cambios.

Se han escogido cuatro casos para realizar el análisis de sensibilidad, para cada caso se calculará los parámetros óptimos mediante el software desarrollado y luego se comparará su rendimiento para cambios diferentes para los que el gráfico fue optimizado. Los casos fueron seleccionados de tal forma que se puede realizar comparaciones entre cartas de control optimizadas para pequeños cambios y cartas de control optimizadas para grandes cambios. Estos casos son siguientes:

Tabla 3.4 Casos a considerar en el análisis de sensibilidad

Fuente: Christian Naranjo, 2019

Caso	Distancia Mahalanobis	p	ARL ₀
1	0,50	2	500
2	0,80	2	500
3	3,00	2	500
4	3,50	2	500

3.3.1 Gráfico MDS T²

En la tabla 3.5 se han calculado mediante el software los parámetros óptimos para el gráfico GMDS T^2 en los casos mencionados anteriormente.

Tabla 3.5 Parámetros óptimos para los casos a estudiar en el análisis de sensibilidad

Fuente: Elaboración propia

Caso	Distancia		ARL ₀	m	LCS ₂	LCS ₁	ARL ₁
	Mahalanobis	p					
1	0.5	2	500	4	16.93	7.67	224.48
2	0.8	2	500	4	17.87	7.63	92.14
3	3	2	500	2	19.48	6.92	1.52
4	3.5	2	500	6	18.29	8.01	1.25

Las tablas siguientes corresponden al análisis de sensibilidad de los gráficos MDS T^2 en la tabla 3.5. En estas tablas se observa que para todos los casos propuestos el nuevo gráfico tiene un mejor rendimiento que el gráfico de Hotelling sin importar que el cambio ocurrido sea grande o pequeño. Observando las tablas 3.6 y 3.7 donde el gráfico de control fue optimizado para cambios pequeños notamos que en ambos casos los rendimientos son buenos y muy parecidos. En estos gráficos los mejores rendimientos se obtienen cuando los cambios ocurridos son mayores a uno, donde el desempeño del nuevo gráfico supera al de Hotelling en más del 41%, es decir, a pesar de ser optimizado para cambios pequeños el gráfico tiene buen rendimiento cuando ocurre cambios grandes. En las tablas 3.8 y 3.9 los gráficos han sido optimizados para cambios grandes, sin embargo, notamos que cuando ocurren cambios pequeños el gráfico propuesto de todas formas tiene un mejor desempeño comparado al T^2 de Hotelling, así lo evidenciamos con un cambio de 0.5 donde la mejora es de 11% en el caso 3 y 18% en el caso 4.

Según lo expuesto anteriormente, se recomienda utilizar el gráfico MDS T^2 optimizando sus parámetros para distancias de Mahalanobis pequeñas dado su buen desempeño.

Tabla 3.6 ARL fuera de control del gráfico MDS T^2 optimizado para un cambio 0.5

Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T^2	Caso 1	
		ARL ₁ MDS T^2	Diferencia T^2
0.5	265.74	224.48	-15.53%
1	84.95	50.02	-41.12%
1.5	28.33	12.58	-55.59%
2	11	4.47	-59.36%
3	2.81	1.53	-45.55%

Tabla 3.7 ARL fuera de control del gráfico MDS T^2 optimizado para un cambio 0.8

Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T^2	Caso 2	
		ARL ₁ MDS T^2	Diferencia T^2
0.5	265.74	224.68	-15.45%
1	84.95	49.88	-41.28%
1.5	28.33	12.51	-55.84%
2	11	4.44	-59.64%
3	2.81	1.53	-45.55%

Tabla 3.8 ARL fuera de control del gráfico MDS T^2 optimizado para un cambio 3

Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T^2	Caso 3	
		ARL ₁ MDS T^2	Diferencia T^2
0.5	265.74	237.84	-10.50%
1	84.95	56.63	-33.34%
1.5	28.33	14.55	-48.64%
2	11	5.01	-54.45%
3	2.81	1.52	-45.91%

Tabla 3.9 ARL fuera de control del gráfico MDS T² optimizado para un cambio 3.5

Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T ²	Caso 4	
		ARL ₁ MDS T ²	Diferencia T ²
0.5	265.74	218.46	-17.79%
1	84.95	46.89	-44.80%
1.5	28.33	11.76	-58.49%
2	11	4.33	-60.64%
3	2.81	1.58	-43.77%

3.3.2 Gráfico GMDS T²

En la tabla 3.10 se han encontrado los parámetros óptimos para el gráfico GMDS T² en los casos mencionados anteriormente.

Tabla 3.10 Parámetros óptimos para los casos a estudiar en el análisis de sensibilidad

Fuente: Elaboración propia

Caso	Distancia Mahalanobis	p	ARL ₀	k	m	LCS ₂	LCS ₁	ARL ₁
1	0.5	2	500	4	6	19.12	4.5	198.54
2	0.8	2	500	3	6	16.1	3.52	71.01
3	3	2	500	2	4	18.09	3.77	1.17
4	3.5	2	500	1	2	17.41	4.2	1.13

Las siguientes tablas muestran el desempeño de cada gráfico de control para diferentes cambios. En estos se observa un comportamiento similar al gráfico MDS T², pero con mejores desempeños. Claramente notamos que para todos los casos propuestos el nuevo gráfico tiene un mejor rendimiento que el gráfico de Hotelling sin importar que el cambio ocurrido sea grande o pequeño. Observando las tablas 4.11 y 4.12 donde el gráfico de control fue optimizado para cambios pequeños notamos que en ambos casos los rendimientos son buenos y muy parecidos. En estos gráficos los mejores rendimientos se obtienen cuando los cambios ocurridos son mayores a uno,

donde el desempeño del nuevo gráfico supera al de Hotelling en más del 60%, es decir, a pesar de ser optimizado para cambios pequeños el gráfico tiene buen rendimiento cuando ocurren cambios grandes. En las tablas 3.13 y 3.14 los gráficos han sido optimizados para cambios grandes, sin embargo, notamos que cuando ocurren cambios pequeños el gráfico propuesto de todas formas obtiene un mejor desempeño comparado al T^2 de Hotelling, así lo evidenciamos con un cambio de 0.5 donde la mejora es de 17% en el caso 3 y 8% en el caso 4.

Según lo expuesto anteriormente, se recomienda utilizar el gráfico GMDS T^2 optimizando sus parámetros para distancias de Mahalanobis pequeñas dado su buen desempeño.

Tabla 3.11 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 0.5
Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T^2	Caso 1	
		ARL ₁ GDMS T^2	Diferencia T^2
0.5	265.74	198.54	-25.29%
1	84.95	32.48	-61.77%
1.5	28.33	6.62	-76.63%
2	11	2.39	-78.27%
3	2.81	1.17	-58.36%

Tabla 3.12 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 0.8
Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T^2	Caso 2	
		ARL ₁ GDMS T^2	Diferencia T^2
0.5	265.74	204.1	-23.20%
1	84.95	33.74	-60.28%
1.5	28.33	6.59	-76.74%
2	11	2.27	-79.36%

3	2.81	1.11	-60.50%
---	------	------	---------

Tabla 3.13 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 3

Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T^2	Caso 3	
		ARL ₁ GDMS T^2	Diferencia T^2
0.5	265.74	221.22	-16.75%
1	84.95	42.23	-50.29%
1.5	28.33	8.95	-68.41%
2	11	2.97	-73.00%
3	2.81	1.17	-58.36%

Tabla 3.14 ARL fuera de control del gráfico GMDS T^2 optimizado para un cambio 3.5

Fuente: Elaboración propia

Distancia Mahalanobis	ARL ₁ T^2	Caso 4	
		ARL ₁ GDMS T^2	Diferencia T^2
0.5	265.74	245.26	-7.71%
1	84.95	57.69	-32.09%
1.5	28.33	14.11	-50.19%
2	11	4.69	-57.36%
3	2.81	1.41	-49.82%

3.4 Implementación en una empresa

En la producción de harina de pescado algunas de las variables que se controlan y que intervienen en la calidad de la harina son los porcentajes de:

- Proteína
- Ceniza
- Grasa
- Acidez

La cantidad de proteína está relacionada directamente con el precio al que se venderá la harina, entre mayor sea el porcentaje de proteína, mayor será el

precio de venta. La ceniza está vinculada a la humedad y esta está relacionada con la presencia de hongos, entre más humedad exista, mayor será la probabilidad de aparición de hongos en la harina y que esta se eche a perder, sin embargo, una menor humedad incrementa la presencia de ceniza lo cual hace que el camarón, uno de los consumidores de este tipo de harina, tenga baja digestibilidad de este producto y por ende requiera más alimento para poder crecer. En cuanto al porcentaje de grasa, este también requiere un monitoreo ya que, a mayor grado de grasa en el producto, más expuesto estará a la rancidez y con esto a la pérdida de valor nutricional de la harina.

Lo anterior expuesto explica la importancia de estas variables en el proceso que por ende deben cumplir con las especificaciones técnicas estipuladas por el área de calidad de la empresa.

Cabe recalcar que las variables mencionadas están correlacionadas entre si ya que forman parte de la composición química de la harina de pescado, y la disminución de alguna variable significará el aumento en cierto grado de las demás.

Se realizó una prueba multivariante de Shapiro para corroborar la normalidad multivariante de las variables dando como resultado la eliminación de la variable ceniza que, a pesar de aplicar las transformaciones conocidas para conseguir la normalidad, esta variable al ser incluida dentro de los parámetros en la prueba de Shapiro provocaba que se rechazara el supuesto de normal multivariante, es así que finalmente las variables a monitorear en nuestra aplicación serán la proteína, grasa, y acidez.

A partir de los datos obtenidos de una muestra, calculamos la matriz de correlación S de las variables seleccionadas. Podemos observar que las correlaciones son todas negativas (a excepción de la matriz diagonal), este resultado era esperado dado el tipo de relación que guardan las variables.

$$\begin{matrix} \text{Proteína} \\ \text{Grase} \\ \text{Ácidez} \end{matrix} \begin{pmatrix} 1 & -0.56 & -0.11 \\ -0.56 & 1 & -0.51 \\ -0.11 & -0.56 & 1 \end{pmatrix}$$

Como siguiente paso se requiere hallar los parámetros que optimizan nuestro gráfico de control. Estos parámetros estarán optimizados para cuando exista un cambio de una desviación estándar en cada una de las medias de las variables y sujeto a que el $ARL_0 = 400$. Adicional, debemos considerar el supuesto que las variables se encuentran estandarizadas. Dado esto, el vector de medias cuando el proceso se encuentra dentro de control será $\bar{X} = \vec{0}$ y el vector de medias del cambio será $\hat{u} = \vec{1}$. Con estos datos y la matriz de correlación se puede calcular la distancia de Mahalanobis que en este caso estará dado por 15.37152.

Para encontrar los parámetros óptimos del gráfico se ingresa en el software desarrollado los siguientes parámetros:

- Número de variables: 3
- ARL_0 : 400
- Distancia de Mahalanobis: 15.37152
- Tamaño de la muestra: 1

Al finalizar los cálculos, el software retorna los siguientes valores: límite de control superior externo UCL , límite de control superior interno UCL_1 , número de muestras anteriores a analizar m , ARL bajo control ARL_0 y ARL fuera de control ARL_1

MDS

En el gráfico 3.1 se muestra el uso del software para el cálculo de los parámetros de la carta de control T^2 que utiliza el esquema de muestreo MDS. La respuesta del software nos indica que los óptimos límites de control superior

externo e interno son 14.41 y 11.7 respectivamente, la cantidad de muestras anteriores a analizar en el caso de que el estadístico este entre 14.41 y 11.7 son 2, el número promedio de muestras de tamaño 1 hasta obtener una señal fuera de control estando el proceso bajo control es de 400 y el número de muestras hasta detectar un cambio de una desviación estándar en cada una de las medias es de 1.

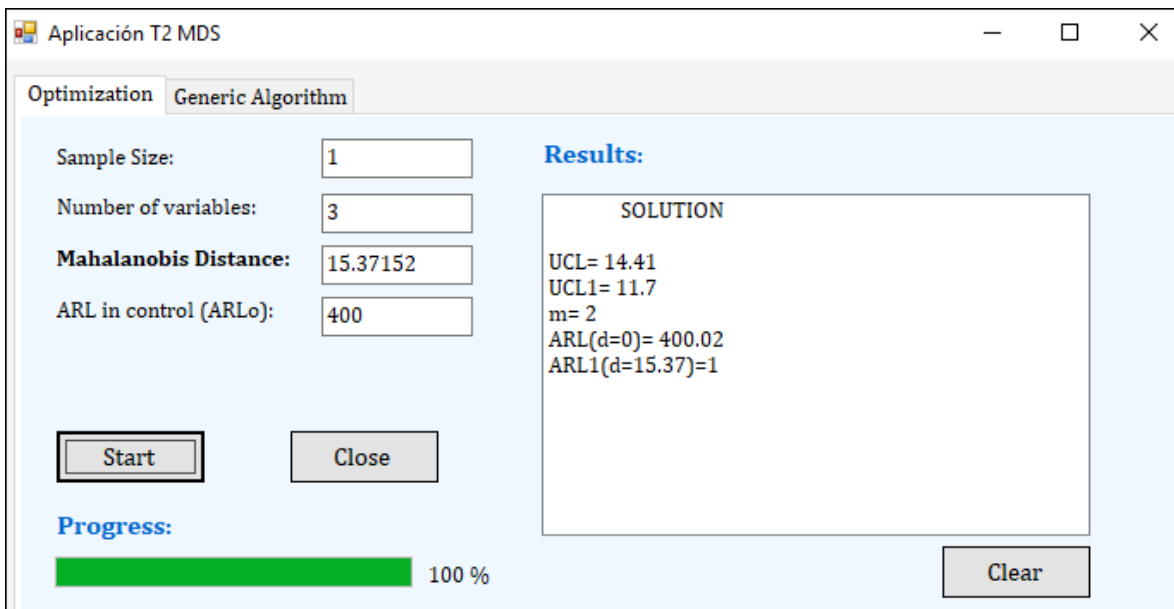


Gráfico 3.1 Respuesta del software para la implementación del gráfico MDS en una empresa.

Fuente: Elaboración propia

GMDS

En el gráfico 3.2 se muestra el uso del software para el cálculo de los parámetros de la carta de control T^2 que utiliza el esquema de muestreo GMDS. La respuesta del software nos indica que los óptimos límites de control superior externo e interno son 14.4 y 5.14 respectivamente, la cantidad de muestras anteriores a analizar en el caso de que el estadístico este entre 14.4 y 5.14 son 6 de las cuales al menos 2 deben ubicarse debajo del límite interno para que el gráfico no lance una alerta. El número promedio de muestras de tamaño 1 hasta obtener una señal fuera de control estando el proceso bajo

control es de 400 y el número de muestras hasta detectar un cambio de una desviación estándar en cada una de las medias es de 1.

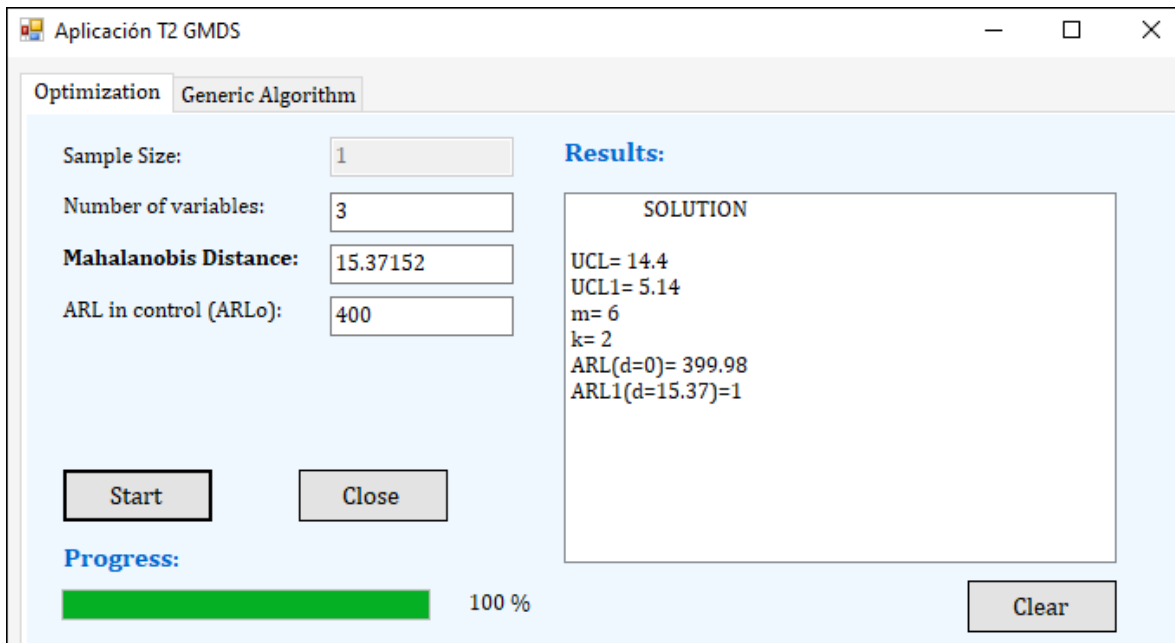


Grafico 3.2 Respuesta del software para la implementación del grafico GMDS en una empresa.

Fuente: Elaboración propia

CAPITULO 4

4. Conclusiones y recomendaciones

4.1 Conclusiones

- Se logró diseñar dos gráficos de control multivariante nuevos, uno utilizando el esquema de muestreo GMDS y otro usando el esquema MDS, los cuales incorporan parámetros adicionales a los tradicionales para lograr un mejor rendimiento.
- Se desarrolló un software que calcula y optimiza los parámetros de los gráficos control utilizando para esto algoritmos genéticos.

- Se comparó el rendimiento de las nuevas cartas de control con el gráfico de control T^2 de Hotelling resultando más eficientes tanto en la detección de cambios pequeños y como en la detección de cambios más grandes, obteniendo en las pruebas un máximo de eficiencia del 60.74% y 76.76% para el grafico MDS y GMDS respectivamente, siendo el mejor gráfico el GMDS T^2 .
- Se evidenció la eficiencia de los gráficos propuestos en el análisis de sensibilidad donde se demostró que estos gráficos poseen un buen desempeño tanto si ocurre un cambio pequeño o un cambio grande sin importar el cambio para el que haya sido optimizado el gráfico.
- Se utilizó los gráficos propuestos y el software implementado para el desarrollo de una carta de control optimizada a partir de los datos reales tomados de una empresa, consiguiendo con esto poder monitorear el proceso de producción de harina de pescado que ellos llevan.

4.2 Recomendaciones

- Se recomienda comparar los gráficos de control propuestos con otros gráficos más recientes que el T^2 de Hotelling para así confirmar los resultados de su eficiencia.
- Continuar con la línea de investigación con el fin de mejorar la detección de cambios pequeños, para esto se podría utilizar un esquema de muestreo repetitivo, que ha obtenido buenos resultados para varios autores.

Bibliografía

- Aparisi, F. (1996). *Hotelling's T' control chart with adaptive sample sizes*. INT. J. PROD. RES., 1996, VOL. 34, NO 10, 2853-2862.
- 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.
- Aslam, M. K.-H. (2014). *A Multiple Dependent State Control Chart Based on Double Control Limits*. Research Journal of Applied Sciences, Engineering and Technology.
- Balamurali S, Jun C-H. (2007). *Multiple dependent state sampling plans for lot acceptance based on measurement data*. European Journal of Operational Research.
- Besterfield, D. H. (2009). *Control de calidad*. Naucalpan de Juárez: Pearson Educación de México.
- Cynthia A. Lowry, William H. Woodall , Charles W. Champ & Steven E. Rigdon . (1992). *A Multivariate Exponentially Weighted Moving Average Control Chart*. Technometrics.
- Darwin, C. (2003). *El origen de las especies (Vol. No. 575)*. J. d. Rojas, Ed.
- Davis, L. (1991). *Handbook of Genetic Algorithms*. New York: Van Nostrand Reinhold.
- Fraser, A. S. (1957). "*Simulation of genetic systems by automatic digital computers. I. Introduction,*". Aust. J. Biol. Sci., vol. 10,.
- Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA.: Addison-Wesley.
- Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.
- Hotelling, H. (1947). *Multivariate quality control, illustrated by the air testing of simple bombsights*. . Techniques of statistical analysis, 111-184.
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *National Institute of Science of India*, 49-55.

- Michalewicz, Z. (1992). *Genetic Algorithms + Data Structures = Evolution Programs*. Berlin Heidelberg: Springer-Verlag .
- Mogg, J. M., and Wortham, A. W. (1970). *Dependent stage sampling inspection*. Int. J. Prod. Res., 8, 385.
- Montgomery, D. C. (2006). *Introduction to Statistical Quality Control* (6 ed.). USA: Wiley.
- Murphy, B. (1987). *Selecting Out of Control Variables With the T2 Multivariate Quality Control Procedure*. . The Statistician, 571-581.
- Parra, H. (2016). *Calidad, Productividad y Costos: Análisis de relaciones entre estos tres conceptos*. Revista de la Facultad de Ciencias Económicas y Empresariales., 46-65.
- Reeves, C. (1993). *Modern Heuristic Techniques for Combinatorial Problems*. Blackwell Scientific Publications.
- Seber, G. A. (1984). *Multivariate Observations*. Wiley Series in probability and mathematical statistics.
- W. Wortham & R. C. Baker. (1976). *Multiple deferred state sampling inspection*. International Journal of Production Research, 14:6, 719-731.
- Wackerly, D. D., Mendenhall III, W., & Scheaffer, R. L. (2009). *Estadística Matemática*. Santa Fe: Cengage Learning Editores.