

GRÁFICOS DE CONTROL PARA PROCESOS CON VARIAS VARIABLES CRÍTICAS QUE SE ESTUDIAN CONJUNTAMENTE (UNA REVISIÓN DEL ESTADO DEL ARTE)

Jaime Mosquera Restrepo

Profesor Escuela de Ingeniería Industrial y Estadística Universidad del Valle, Colombia.

Contacto: jaimores@pino.univalle.edu.co

Resumen

Las técnicas de control estadístico de procesos constituyen una herramienta de apoyo fundamental en la búsqueda del mejoramiento continuo y cercamiento de las condiciones de variabilidad de innumerables procesos de producción. Gráficos de control tipo Shewarth, los gráficos de control alternativos y el análisis de capacidad de proceso, han demostrado durante mucho tiempo su efectividad en el cumplimiento de este objetivo cuando se controla una sola variable de proceso. A pesar de ello, son muchos los procesos productivos para los cuales la calidad se define a través de la interacción de diferentes características cuantitativas, cuyo análisis y control conjuntos enfrentan a los departamentos de calidad a lo que se denomina control de procesos en el ámbito multivariante. En este trabajo, se presenta de manera resumida algunas alternativas posibles para el manejo de situaciones como las planteadas anteriormente, identificando ventajas y desventajas de cada una de ellas y sus procedimientos de aplicación.

Palabras Claves: Control Estadístico de Procesos, Gráficos de Control Multivariados, Descomposición MTY.

Abstract

The techniques of statistical process control constitute a fundamental tool in search of continuous improvement and reduction of variability conditions of a lot production process. Shewart control charts, alternative control charts and capability analysis, have been used sucesfully during long time your efficiency in the fulfillment of this objective when a single variable of the process is controlled. But, many productive processes exist for which the quality is defined for the interaction between different quantitative characteristics, whose analysis and joint control confront to the quality departments to which it is denominated like process control in the multivariate environment. This work present the summary of some possible alternatives for the treatment of situations like was posed previously, identifying advantages and disadvantages of everyone of these and your application procedures.

Keywords: Statistical Process Control, Multivariate Control Charts, Decomposition MTY.

1. Introducción

Con el objetivo de lograr una ventaja competitiva en los mercados a nivel regional, nacional e internacional, las industrias se ven enfrentadas a mejorar de manera continua la calidad de los productos y servicios que ofrecen, generándose de esta manera la necesidad de implementar dentro de sus metodologías de producción algún tipo de control que les permita responder a las exigencias de calidad presentes en los actuales mercados. Para ello disponen actualmente de amplios desarrollos en este campo, tal como son los diferentes tipos de gráficos de control, los cuales permiten hacer seguimiento de las condiciones de variabilidad y centramiento de alguna característica de calidad que surge en un proceso de producción. Pero, ¿Qué hacer cuando la calidad de un producto no se mide a través de una sola característica, sino por el contrario, está correspondiente a la interacción de múltiples características?

Una posible solución a un caso como el mencionado anteriormente, puede ser efectuar el control de proceso a través de la implementación de un gráfico de control para cada uno de las características en mención, contrayendo consigo una serie de desventajas, entre ellas:

1. el número de gráficos de control a implementar y vigilar puede llegar a ser relativamente grande, esto dependiendo del número de características que se deseen controlar, lo cual en la práctica se hace inmanejable de no contarse con herramientas computacionales adecuadas.
2. Al realizar el control de las características de manera independiente se esta obviando cualquier tipo de relación que pudiera existir entre las características de calidad. Adicionalmente, el control a través de gráficos individuales conlleva a una alteración del verdadero nivel de significancia con que se realiza la prueba. Esto se puede demostrar a través del siguiente desarrollo:

Si se controlan p variables independientes a través de gráficos tipo Shewart, asumiendo que las características siguen distribuciones normales y situando los límites de control de tal manera que se satisface que cada gráfico de control posee una probabilidad de error tipo I = α , la probabilidad del error tipo I para el proceso en conjunto, α^c , es:

$$(1 - \alpha^c) = (1 - \alpha_1) \cdot (1 - \alpha_2) \cdot (1 - \alpha_3) \cdot \dots \cdot (1 - \alpha_{p-1}) \cdot (1 - \alpha_p) = \prod_{i=1}^p (1 - \alpha_i) \quad (1)$$

Asumiendo que cada uno de los gráficos fue creado utilizando el mismo nivel de significancia, la expresión se reduce a:

$$(1 - \alpha^c) = (1 - \alpha)^p \Rightarrow \alpha^c = 1 - (1 - \alpha)^p \quad (2)$$

Dada la necesidad de resolver el problema planteado anteriormente, y contemplando las desventajas a las que conlleva la aplicación independiente de múltiples gráficos de control univariados, diferentes autores han propuesto alternativas de solución, entre las que se cuentan los diferentes tipos de gráficos de control multivariado: T^2 de Hotelling, MCUSUM, MEWMA, cuya sustento teórico y forma de aplicación se estudiarán en el presente artículo.

2. Presentación Metodológica de los Gráficos de Control Multivariado

El controlar múltiples características de manera simultánea a través de gráficos tipo Shewart individuales conlleva a la formación de una región de control rectangular (caso 2 características) o cúbica (caso mas de 2 características), delimitada por los diferentes límites de control definidos para cada una de las características. Una representación gráfica de esta situación se ilustra en la figura 1 para el caso de 2 variables.

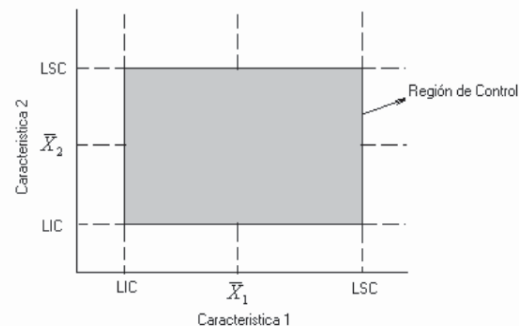


Figura 1. Región de control para dos variables mediante gráficos independientes

La zona oscura representa la región de control, es decir, una observación que al graficarla pertenezca a esta región será una observación de un proceso bajo control, mientras que una observación que no pertenece a esta región, daría un indicio de que el proceso está fuera de control. Sin embargo, esta región deja de lado la posible interacción o relación existente entre las características, bajo las cuales la región de control se vería modificada. Pensando en ello diferentes autores entre los que se encuentra Harold Hotelling (1947)¹, postulan que la región de

control para estos casos asume una forma elipsoidal, lo cual es fácil de deducir cuando se asume que las características siguen una distribución normal multivariante, tal como se ilustra en figura 2.

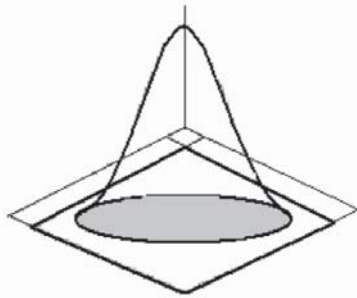
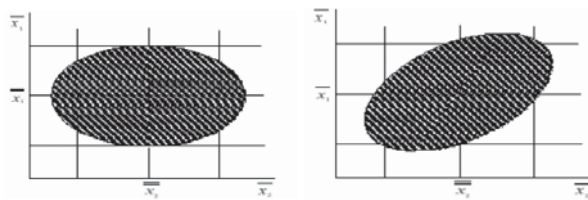


Figura 2. Elipse de control como proyección de la distribución normal multivariante

Obsérvese como, casi de manera natural, la elipse de control surge como la proyección sobre los ejes de las características de la distribución normal multivariante, cuya excentricidad dependerá del nivel de asociación entre las características y su inclinación de la direccionalidad de esta asociación. Tal como se ilustra en la Figura 3.



a. Correlación Positiva b. No correlación

Figura 3. Inclinación elipse de control según asociación entre características

La implementación de estas elipses de control presenta dos grandes inconvenientes: 1. Imposibilidad de ubicar el conjunto de observaciones en una secuencia temporal, lo cual limitaría la posibilidad de observar cambios graduales en el centramiento del proceso. 2. En el

caso en que el número de características a controlar sea relativamente grande ($p > 2$) la región de control dejará de ser una elipse para convertirse en un elipsoide, lo cual es muy difícil de llevar a una expresión gráfica.

Partiendo de estas dificultades, la tendencia en este campo se ha dado hacia la reducción de la información arrojada por el proceso de producción, a través de la construcción de un estadístico capaz de resumir en un escalar las condiciones de centramiento y variabilidad del proceso en las dimensiones o características observadas. Adicionando mediante algún procedimiento probabilístico o por simulación algún límite de control, similar al procedimiento en control de procesos para el caso univariante.

2.1 Gráfico T^2 de Hotelling.

El estadístico T^2 es un escalar que combina información de la variación y centramiento de las variables analizadas. Cuando se conocen los verdaderos parámetros de la distribución bajo control, es decir, conocidos el centramiento y la estructura de asociación entre las características, bajo el supuesto de normalidad multivariante, el estadístico T^2 sigue una distribución $\chi^2_{(p)}$. En el caso bivariado, este estadístico tiene la siguiente forma:

$$\chi^2 = n(\bar{\mathbf{X}} - \bar{\boldsymbol{\mu}})^T \Sigma^{-1} (\bar{\mathbf{X}} - \bar{\boldsymbol{\mu}}) \quad (3)$$

Donde $\bar{\boldsymbol{\mu}}_{(px1)}$ es el vector que contiene las medias poblacionales de las características bajo control, Σ_{pxp} es la matriz de varianzas y covarianzas cuando el proceso se encuentra bajo control, n es el tamaño muestral empleado y $\bar{\mathbf{X}}_{(px1)}$ es el vector que contiene los promedios muestrales observados de las p características analizadas.

De esta manera, es posible construir un gráfico de control con solo un límite de control superior (ya que este estadístico es no negativo), el cual se situaría para un nivel de significancia, en $LSC =$, valor que representa el punto porcentual de la distribución con p grados de libertad.

¹ Citado por Alfaro Navarro José Luís (2002). "Métodos Multivariantes en Control Estadístico de la Calidad" http://www.uclm.es/ab/fcee/paginas/D_trabajos/2-2002-2.pdf.

Es muy común que en la práctica se desconozcan los parámetros de la distribución del proceso bajo control (centramiento y estructura de asociación), surgiendo la necesidad de estimarlos a través de una muestra obtenida del mismo proceso, dando origen a la distribución T^2 de Hotelling. En este caso el estadístico asumiría la siguiente expresión:

$$T^2 = n(\vec{X} - \vec{A})^T S^{-1}(\vec{X} - \vec{A}) \quad (4)$$

Donde $\vec{A}$ es el vector que contiene las estimaciones preliminares de las medias poblacionales de las p características de calidad cuando el proceso está controlado y S es la matriz que contiene las estimaciones preliminares de la estructura de varianza y covarianzas del proceso bajo control.

Ryan (1989)² derivó la distribución probabilística de la estadística T^2 para dos fases del procedimiento, las cuales denomina Instalación y Operación. En la primera fase se estiman los parámetros de la distribución y se evalúan las observaciones frente a un límite de control preliminar el cual puede ubicarse según la siguiente expresión:

$$LSC = \frac{p(c-1)(n-1)}{cn-c-p+1} F_{\alpha/2; p, cn-c-p+1} \quad (5)$$

p: Número de características de calidad monitoreadas

c: Número de subgrupos utilizados en la estimación del vector de medias y la matriz de covarianzas

n: Tamaño de muestra por subgrupo

$F_{\alpha/2; p, cn-c-p+1}$: Percentil $\alpha/2$ de la distribución F con p grados de libertad en el numerador y $(cn-c-p+1)$ en el denominador.

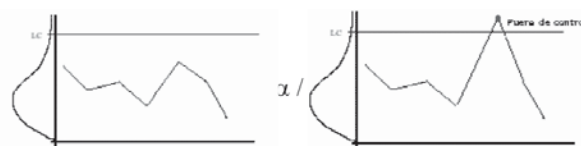
Aquellos subgrupos que en la fase de instalación sean clasificados como situaciones fuera de control, deben ser eliminados, estimando nuevamente los parámetros de la distribución con los subgrupos restantes y reevaluando frente al límite de control hasta tanto el gráfico muestre un proceso completamente bajo control.

Para la segunda fase (operación), se replantea el límite de control para monitorear observaciones futuras, utilizando la siguiente expresión:

$$LSC = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha/2; p, mn-m-p+1} \quad (6)$$

En este caso m representa el número de subgrupos con los se realizó la estimación definitiva del vector de medias y la matriz de covarianzas en la fase de instalación. Es decir, el subconjunto obtenido luego de filtrar los subgrupos que señalaron situaciones fuera de control en la fase de instalación.

De esta manera es posible determinar si el proceso se encuentra bajo control mediante la representación de los valores T^2 junto a su respectivo límite de control. Si todas las muestras arrojan valores T^2 inferiores a LSC, se dice que el proceso se encuentra bajo control estadístico, de lo contrario se dirá que el proceso está fuera de control, tal como lo muestra la figura 4.



Proceso Bajo Control Proceso Fuera de Control

Figura 4. Ejemplo de Gráfico T^2 de Hotelling.

2.2 Gráfico de Control MEWMA (Multivariado de Medias Móviles Exponencialmente Ponderadas)

El gráfico de control MEWMA, busca mejorar la eficiencia en la detección de desviaciones de moderada magnitud en el centramiento del proceso, a través de la inclusión de información pasada del mismo. Lo anterior lo logra mediante el uso de una serie de ponderadores que dan mayor o menor relevancia a la historia del proceso. Este gráfico puede entenderse como la extensión multivariante del gráfico EWMA ya que, Análogo al caso univariado, se asignan pesos r_i a las observaciones actuales, para cada variable de proceso. tal como se pudo observar en la siguiente expresión:

² Citado por Konrath Andrea. "Decomposição da Estatística do Gráfico de Controle Multivariado T^2 de Hotelling por Meio de um Algoritmo Computacional". Universidade Federale de Santa Catalina. 2002

$$\begin{aligned}\bar{\mathbf{Z}}_k &= \mathbf{R}\bar{\mathbf{X}}_k + (\mathbf{I} - \mathbf{R})\bar{\mathbf{Z}}_{k-1} \quad k=1,2,3,4,\dots \\ \bar{\mathbf{Z}}_0 &= \mathbf{0}\end{aligned}\quad (7)$$

Donde $\mathbf{R} = \text{diag}(r_1, r_2, \dots, r_p)$ representa una matriz diagonal de pesos exponenciales, Estos pesos cumplen con la condición $(0 < r_i < 1)$ $i = 1, 2, \dots, p$. Un peso cercano a 1 brindaría mayor ponderación a la información actual del proceso, mientras que un peso cercano a 0, pondera la información histórica. El subíndice k representa un contador muestral, que se incrementa con cada observación adicionada al gráfico de control. De esta manera, se considera que el proceso se encuentra fuera de control en el instante k cuando el gráfico de control MEWMA señala que:

$$T_k^2 = \bar{\mathbf{Z}}_k^T \Sigma_{\mathbf{Z}_k}^{-1} \bar{\mathbf{Z}}_k > L \quad (8)$$

$\Sigma_{\mathbf{Z}_k}$ es la matriz de covarianza del vector $\bar{\mathbf{Z}}_k$. "Lowry, (1992) ha demostrado que el elemento (i, j) de la matriz $\Sigma_{\mathbf{Z}_k}$ es³:

$$\Sigma_{(i,j)} = r_i r_j \frac{[1 - (1 - r_i)^k (1 - r_j)^k]}{r_i + r_j - r_i r_j} \sigma_{i,j} \quad i, j = 1, 2, \dots, p \quad (9)$$

En muchas ocasiones, no se dispone del suficiente criterio como para aplicar diferentes pesos exponenciales a las características de calidad estudiadas. En ese caso, se usan pesos iguales para todas las características, tal que $r_i = r$ para $i = 1, 2, \dots, p$. El vector MEWMA puede escribirse bajo el supuesto de pesos iguales de la siguiente manera:

$$\mathbf{R} = r\mathbf{I} \quad \bar{\mathbf{Z}}_k = r\bar{\mathbf{X}}_k + (1 - r)\bar{\mathbf{Z}}_{k-1} \quad (10)$$

Haciendo una particularización del caso general, es posible demostrar que cuando se emplean pesos iguales la matriz de covarianza del vector MEWMA es:

$$\Sigma_{\mathbf{Z}_k} = \frac{r}{2 - r} [1 - (1 - r)^{2k}] \Sigma_{\mathbf{X}} \quad (11)$$

Para valores grandes de k , la cantidad $[1 - (1 - r)^{2k}]$ se aproxima a la unidad, por lo que la matriz de

covarianza del vector MEWMA puede definirse de manera asintótica como:

$$\Sigma_{\mathbf{Z}_k} = \frac{r}{2 - r} \Sigma_{\mathbf{X}} \quad (12)$$

Lo anterior permite simplificar el estadístico T_k^2 puede de la siguiente manera:

$$\begin{aligned}T_k^2 &= \bar{\mathbf{Z}}_k^T \Sigma_{\mathbf{Z}_k}^{-1} \bar{\mathbf{Z}}_k \quad T_k^2 = \bar{\mathbf{Z}}_k^T \left[\frac{r}{2 - r} \Sigma_{\mathbf{X}} \right]^{-1} \bar{\mathbf{Z}}_k \\ T_k^2 &= \frac{2 - r}{r} \bar{\mathbf{Z}}_k^T \Sigma_{\mathbf{X}}^{-1} \bar{\mathbf{Z}}_k\end{aligned}\quad (13)$$

El gran limitante que presenta este tipo de gráfico consiste en que la distribución del estadístico no hace parte del conjunto de distribuciones comúnmente empleadas en el campo de la estadística, por lo cual la determinación del límite superior de control (L) se viene realizando mediante prácticas de simulación, intentando lograr un ARL del proceso bajo control específico. Alfaro Navarro en su artículo "Métodos Multivariantes en el Control Estadístico de la Calidad" proporciona una tabla guía para la selección de este límite cuando se usan pesos iguales, la cual depende del valor r seleccionado y del número de características monitoreadas. Estos valores fueron determinados de tal forma que se satisfaga una $ARL_{(0)} = 200$. La Tabla 1 recoge algunos de los valores para este límite de control.

r	Número de Variables			
	p=2	p=4	p=6	p=10
0.05	7.35	11.22	14.6	20.72
0.1	8.64	12.73	16.27	22.67
0.2	9.65	13.87	17.51	24.07
0.3	10.08	14.34	18.01	24.62
0.4	10.31	14.580	18.26	24.89

Tabla 1. Límites para el Gráfico de Control MEWMA

Fuente: Alfaro Navarro. "Métodos Multivariantes en el Control Estadístico de la Calidad".

Otra alternativa para la selección del límite de control y los factores ponderadores es brindada por Aparisi (2003)⁴, a través del desarrollo de un software que optimiza dicho parámetro para el

³ Citado por K. Linderman, T. Love. "Economic and economic statistical design for MEWMA control charts". Journal of Quality Technology. Vol. 32. No. 4. Octubre 2000.

gráfico de control mediante la aplicación de Algoritmos Genéticos.

2.3 Gráfico MCUSUM (Multivariado de Sumas Acumulativas)

El gráfico MCUSUM puede verse como la extensión de un gráfico CUSUM al caso multivariado, encontrándose en la práctica dos métodos para su implementación. El primer método consiste en reducir las observaciones multivariantes a un escalar y elaborar un gráfico CUSUM sobre el mismo. Esta reducción puede lograrse a través de diferentes métodos, entre ellos: Análisis de Componentes Principales, estadístico T^2 de Hotelling u otra forma de reducción adecuada. El segundo método consiste en elaborar un MCUSUM directamente para el conjunto de observaciones multivariantes. En este sentido, Pignatello y Runger (1990), proponen una de las metodologías con mejor comportamiento en términos de ARL, cuya metodología de aplicación resumir mediante el siguiente algoritmo:

1. Establecer o estimar mediante procedimientos estadísticos el vector de centramiento del proceso bajo control $\bar{\mu}_0$ y la matriz de varianzas y covarianzas Σ .
2. Determinar el valor de referencia " K_0 " que determina la máxima distancia a tolerar entre el vector de promedios observados y el vector de centramiento del proceso. Normalmente este valor se fija en la cantidad 0.5.
3. Recoger muestras de tamaño n del proceso, y calcular el vector de promedios observados $\bar{X}$.
4. Construir el vector de desviaciones y calcular su distancia estandarizada mediante las siguientes expresiones:

$$\bar{C}_0 = (\bar{x}_1 - \bar{\mu}_0) \quad \|\bar{C}_0\| = \sqrt{\bar{C}_0' \Sigma^{-1} \bar{C}_0} \quad (14)$$

5. Contrastar la distancia estandarizada versus el valor de referencia establecido, concluyendo de la siguiente manera:

Si $\|\bar{C}_0\| > K_0$ El proceso se encuentra fuera de control

6. En caso de que el proceso se encuentre fuera de control, deberán aplicarse los correctivos necesarios y reiniciar el procedimiento.

7. En el caso de que el proceso se encuentre bajo control, se adiciona otra observación al proceso y se calcula el vector de desviaciones acumuladas, acumulando desviaciones desde el último punto clasificado como señal fuera de control.

$$\bar{C}_t = \sum_{i=1}^t (\bar{x}_i - \bar{\mu}_0) \quad (15)$$

8. El valor de referencia o limite de control, debe recalcularse para cada observación adicionada en la suma de desviaciones, incrementándolo en K_0 unidades por cada nueva observación, es decir:

$$K_t = t * K_0 \quad (16)$$

9. Calcular la magnitud de la distancia acumulada

$$\|\bar{C}_t\| = \sqrt{\bar{C}_t' \Sigma^{-1} \bar{C}_t} \quad (17)$$

10. Si la magnitud de distancia acumulada supera el máximo permitido se emite una señal fuera de control, es decir:

Si $\|\bar{C}_t\| > K_t$ El proceso se encuentra fuera de control.

11. Mientras el procedimiento se encuentre bajo control se seguirán acumulando observaciones, hasta tanto no se emita una señal fuera de control, caso en el cual el procedimiento reiniciara. Tal como se ilustra en la Figura 5.

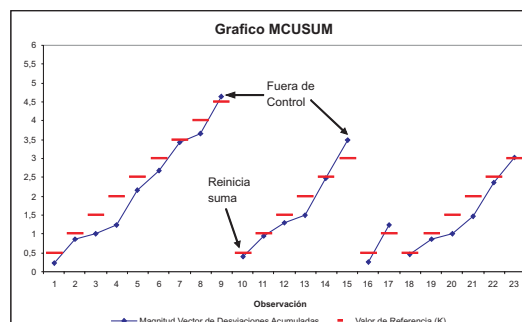


Figura 5. Ejemplo de Gráfico MCUSUM

2.4 Interpretación de Señales Fuera de Control

Los gráficos de control multivariantes solucionan el problema de evaluar el estado de control de un proceso. Sin embargo estos no arrojan información acerca de la fuente que origina estas situaciones

fuera de control, lo cual hace necesario acudir a alguna técnica complementaria que permita identificar estas fuentes de variación. Una primera aproximación al problema puede ser el acompañamiento de los gráficos de control multivariantes de sus respectivos gráficos de control individuales, lo cual conllevaría a identificar fuentes de variación cuando estas son originadas por un cambio en el centramiento o en las condiciones de variabilidad de alguna de las características observadas. Esta solución no es muy recomendable, puesto que una señal fuera de control puede tener su origen en un cambio en la estructura de asociación de las características monitoreadas, lo cual no sería detectado por los gráficos individuales.

Un segundo enfoque consistente en la realización de una serie de intervalos en donde se esperaría se encuentren las promedios muestrales de las características de un proceso bajo control, estos intervalos reciben el nombre de Intervalos de Bonferroni. La idea general de esta metodología consiste en construir p intervalos (uno para cada característica de calidad) para cada muestra que produce una señal fuera de control en el gráfico multivariado. Este intervalo para el subgrupo j -ésimo y la característica de calidad i -ésima tendrá la siguiente forma:

$$\bar{\bar{x}}_i - t_{\alpha/2, p, m(n-1)} S_{pi} \sqrt{\frac{m-1}{mn}} \leq \bar{x}_i^{(j)} \leq \bar{\bar{x}}_i + t_{\alpha/2, p, m(n-1)} S_{pi} \sqrt{\frac{m-1}{mn}} \quad (18)$$

Si el valor del promedio de la variable en la muestra no se encuentra dentro de este intervalo, entonces la característica puede juzgarse como una causa asignable de la señal fuera de control. De nuevo, esta metodología se ve limitada cuando el origen del cambio proviene de la estructura de asociación entre variables.

Otra alternativa fue propuesta por Mason, Tracy y Young (1995), (en adelante denominada como Descomposición MTY) quienes proponen descomponer el estadístico T^2 de Hotelling en componentes ortogonales, las cuales ofrecen información acerca de la causa asignable, luego que una observación ha sido clasificada como una situación fuera de control por el gráfico T^2 de Hotelling. Según estos autores el estadístico T^2 de Hotelling puede ser particionado en dos partes, la cuales son:

$$T^2 = T_{p-1}^2 + T_{p|1,2,\dots,p-1}^2 \quad (19)$$

El término T_{p-1}^2 es el T^2 de Hotelling evaluado con las primeras $p-1$ variables.

$$T_{p-1}^2 = (\bar{X}_i^{(p-1)} - \bar{\bar{X}}^{(p-1)})^T S_{XX}^{-1} (\bar{X}_i^{(p-1)} - \bar{\bar{X}}^{(p-1)}) \quad (20)$$

Para el cual $\bar{\bar{X}}^{(p-1)}$ equivale a la estimación del vector de medias muestrales del proceso bajo control y S_{XX} es la partición de la matriz de varianzas y covarianzas para las $p-1$ primeras variables, es decir:

$$S = \begin{bmatrix} S_{XX} & S_{xX} \\ S_{xX}^T & S_x^2 \end{bmatrix} \quad (21)$$

El estadístico $T_{p|1,2,\dots,p-1}^2$ equivale al cuadrado del valor correspondiente a la p -ésima variable observada, estandarizada por la media y desviación estándar de la distribución condicional de X_p dado $X_1, X_2, X_3, \dots, X_{p-1}$. Esto es:

$$T_{p|1,2,\dots,p-1} = \frac{X_{ip} - \bar{X}_{p|1,2,\dots,p-1}}{S_{p|1,2,\dots,p-1}}$$

Donde $\bar{X}_{p|1,2,\dots,p-1} = \bar{X}_p + b_p^T (X_i^{(p-1)} - \bar{X}^{(p-1)})$ y

$$S_{p|1,2,\dots,p-1} = s_x^2 - s_{xX}^T S_{XX}^{-1} s_{xX} \quad (22)$$

$\bar{X}_p$ es la media muestral de las m observaciones con las que finalmente se estableció el gráfico de control para la variable p , $b_p = S^{-1} s_{xX}$ es el vector conformado por los estimaciones de los parámetros de regresión de la variable p en función de las $p-1$ variables restantes. De la misma manera el estadístico T_{p-1}^2 puede descomponerse como:

$$T_{p-1}^2 = T_{p-2}^2 + T_{p-1|1,2,\dots,p-2}^2 \quad (23)$$

Aplicando el mismo concepto iterativamente, Mason, Tracy y Young (1995) demuestran que el estadístico T^2 de Hotelling puede descomponerse en p componentes ortogonales, de la siguiente manera:

$$T^2 = T_1^2 + T_{2|1}^2 + T_{3|1,2}^2 + \dots + T_{p-1|1,2,\dots,p-2}^2 + T_{p|1,2,\dots,p-1}^2 = T_1^2 + \sum_{i=1}^p T_{i+1|1,\dots,i}^2 \quad (24)$$

Para el cual T_1^2 representa el estadístico T^2 de Hotelling calculado solamente para la primera variable, es decir:

$$T_1^2 = \frac{(X_1 - \bar{X}_1)^2}{S_1^2} \quad (25)$$

Este último termino, denominado termino independiente al igual que los términos condicionales $T_{i+1||, \dots, i}^2$, siguen una distribución conocida, la cual se logra a través de una transformación de la distribución F .

$$T_1^2 \approx \left(\frac{m+1}{m} \right) F_{(1;n-1;\alpha)} \quad (26)$$

$$T_{i+1||, \dots, i}^2 \approx \left(\frac{(m+1)(m-1)}{m(m-j)} \right) F_{(1;n-1;\alpha)}$$

Los percentiles de esta distribución pueden jugar como límite de control para evaluar una causa asignable, de manera que si el término independiente supera el límite de control establecido, indicaría que la señal fuera de control se debe a un cambio presentado en el centramiento de la variable evaluada, mientras que un valor del termino condicional por encima del límite de control indicaría que se ha presentado un cambio en la estructura de asociación existente entre la variable independiente y el conjunto de variables condicionantes.

Como se mencionó anteriormente, el estadístico T^2 puede ser descompuesto en tantas componentes ortogonales como características de procesos se controlen, el problema radica en que la escogencia del término independiente puede darse de entre p opciones posibles, posteriormente el primer término condicional, puede estar condicionado por cualquiera de las $p-1$ variables restantes y así sucesivamente. Lo que genera un conjunto de $p!$ posibles maneras de realizar la descomposición, cada una de ellas con algún aporte sobre la explicación de la señal fuera de control. De tal manera que para un caso de 6 variables se tendrían 720 posibles maneras de descomponer el estadístico, lo cual es en gran medida tedioso de no contarse con la implementación computacional adecuada. Con el objetivo de reducir este procedimiento, Mason, Young y Tracy (1997), proponen un algoritmo que reduciría el número de cálculos necesarios.

Paso 1.

Una vez el gráfico de control T^2 emite una señal de fuera de control, se analizan los términos independientes T_i^2 , si alguno de estos términos es significativo para alguna(s) característica, se dirá que la variable esta fuera de control de manera individual y no será necesario analizar su relación con las demás características de proceso.

Paso 2.

Se calcula el estadístico T^2 de Hotelling únicamente para el conjunto de características que no fueron identificadas como causas asignables individuales en el paso 1. Si este estadístico T^2 no es significativo el procedimiento termina concluyendo que la señal fuera de control puede atribuirse a la(s) característica(s) identificadas en el paso 1, de lo contrario se ejecuta el paso 3.

Paso 3.

Se calculan todas los términos condicionales bivariados T_{ij}^2 con las variables restantes en el vector, si alguno de estos estadísticos es significativo, se dirá que la señal puede deberse a un cambio en la estructura de asociación del par de variables, entonces se eliminarán las variables y con el vector resultante nuevamente se evalúa el estadístico T^2 de Hotelling.

Paso 4.

De la misma manera se repite el procedimiento anexando una variable mas al término condicional hasta tanto el estadístico T^2 del vector resultante sea no significativo.

3. Situaciones a Considerar en la Elección del Gráfico de Control

La elección de un gráfico de control debe realizarse teniendo en cuenta las condiciones tanto de eficiencia estadística como operativas que presenta cada uno de ellos. Dentro de las ventajas estadísticas debe ser considerada la potencia de cada gráfico. En este sentido ha sido demostrado que el gráfico de control T^2 de Hotelling es más potente que los gráficos de control MEWMA y MCUSUM cuando la distancia ente el vector de centramiento y el vector observado es superior a 2 unidades de distancia ($d > 2$), para cambios de menor magnitud la potencia del gráfico T^2 se ve disminuida en relación con la de las otras alternativas. Con el objetivo de mejorar esta falencia en este tipo de gráfico, Aparisi (2001) propone algunas metodologías alternas que conllevarían a aumentar la potencia de este gráfico en la detección de moderados cambios en el proceso, uno de ellos es la implementación de un plan de muestreo variable, en el cual se juega de alguna manera con el tamaño de la muestra y la frecuencia de muestreo. Otra alternativa propuesta por este autor se basa en la inclusión de una serie de límites adicionales que delimitan las que él denomina zonas de atención, en

las cuales después de determinado número de observaciones consecutivas en estas zonas se emitiría una señal fuera de control. Si bien esta metodología puede mejorar las condiciones de potencia del gráfico, trae consigo un incremento en el número de falsas alarmas emitidas. Una alternativa diferente es planteada por Vargas, Yañez y González (2002) en la cual se propone estimar los parámetros en la fase de instalación mediante estimadores robustos.

Otra de las características que juegan en contra del gráfico T^2 es su alta sensibilidad ante condiciones de autocorrelación serial de las observaciones, ya que la presencia de esta patología incrementa el número de falsas alarmas emitidas por el gráfico de control. En este sentido los gráficos de control MEWMA y MCUSUM son más robustos ante esta situación, dado que su estructura autoregresiva (similar a un proceso AR(1)) le permite recoger esta patología sin que ello afecte de mayor manera su normal comportamiento.

A pesar de las falencias resaltadas para el gráfico T^2 de Hotelling, su operatividad y transparencia en la aplicación lo convierten en una excelente alternativa, puesto que, a diferencia de los estadísticos MCUSUM y MEWMA, su distribución de probabilidades es conocida, lo cual permite que el límite de control pueda fijarse mediante procedimientos probabilísticos, bajo el cumplimiento del supuesto de normalidad multivariante. Otro punto que juega a favor del gráfico T^2 es que este no requiere de la elección subjetiva de parámetros anticipados, como se realiza al escoger el factor ponderador r_i para el gráfico MEWMA y el factor de referencia K_0 para el MCUSUM, cuya elección requiere de un amplio conocimiento del proceso de producción por parte del analista. Adicionalmente, la posibilidad de descomposición que presenta el estadístico T^2 , brinda la oportunidad de identificar cual característica produce la señal fuera de control, lo cual en los gráficos MEWMA y MCUSUM no es posible.

4. Actualidad

Actualmente el estado del arte se encuentra en constante evolución. Diferentes autores se encuentran en la tarea de desarrollar investigaciones para mejorar las condiciones de eficiencia de este tipo de metodologías, complementándolas con herramientas computacionales que viabilicen su aplicación y la interpretación de las señales fuera de

control. En este sentido se han implementado rutinas computacionales que buscan identificar la causa de una señal fuera de control a través del entrenamiento de una red neuronal, (Aparisi, 2005). Sin embargo esta metodología tan solo ha sido probada para el caso en que se controlan dos variables. Otros autores han volcado su interés sobre la identificación de la distribución probabilística de los estadísticos MEWMA y MCUSUM, utilizando estimaciones de densidades a través de métodos no paramétricos, específicamente estimación Kernel.

Por último, actualmente se estudia acerca de los efectos de la multicolinealidad en la efectividad de los gráficos de control, ante lo cual se sugiere construir los estadísticos sobre las $j < p$ primeras componentes principales, de tal manera que esta patología no mal condicione la matriz de covarianzas; claro está, esta implementación soluciona el problema de multicolinealidad, pero trae consigo problemas de interpretabilidad de los resultados y la identificación de causa asignables, dado que el conjunto de características iniciales se transforma en un nuevo conjunto de combinaciones lineales cuya interpretación requiere de un amplio conocimiento del proceso y de un adecuado manejo de las componentes, lo cual es difícil de implementar en una línea de producción.

5. Referencias Bibliográficas

1. Alfaro Navarro José Luís (2002). Métodos multivariantes en control estadístico de la calidad. http://www.uclm.es/ab/fcee/paginas/D_trabajos/2-2002-2.pdf.
2. Aparisi Francisco (2003). Optimización de los Gráficos de control Estadístico EWMA y MEWMA Mediante Algoritmos Genéticos. <http://www.udl.es/usuarios/seio2003/treballs/04-1-1.pdf>.
3. Aparisi Francisco, Avendaño Gerardo, Sanz José. Interpreting T^2 control charts signals: Effectiveness of MYT decomposition vs. a neural network. (2005).
4. García Díaz Juan Carlos, Aparisi Francisco. Aumento de la potencia del gráfico de control T^2 de Hotelling utilizando señales adicionales de falta de control. *Estadística Española*, Vol. 43, No 148, 2001.
5. García Díaz Juan Carlos, Aparisi Francisco. Optimización de los gráficos de control estadístico de procesos EWMA y MEWMA mediante algoritmos genéticos. 27^º Congreso de

- Estadística e Investigación Operativa*. http://www.udl.es/usuarios/seio20037treballs/04_1_1.pdf.
6. Haro Cesar, Aparisi Francisco. Hotelling's T^2 control charts with variable sampling intervals. *International Journal of Production Research*. Vol. 39, No. 14, 2001. <http://tandf.co.uk/journals>.
 7. Konrath Andrea (2002). *Decomposição da Estatística do Gráfico de Controle Multivariado T^2 de Hotelling por Meio de um Algoritmo Computacional*. Tesis de Maestría. Universidade Federale de Santa Catalina.
 8. Linderman Kevin, Love Thomas. Economic and economical statistical design for MEWMA control charts. *Journal of Quality Technology*. Vol. 32. No. 4. Octubre 2000.
 9. Mason Robert, Tracy Nola, Young John. Decomposition of T^2 for multivariate Control Chart Interpretation. *Journal of Quality Technology*. Vol. 27. No. 2. Abril 1995.
 10. Pignatiello, J.J. & Runger, G.C. Comparisons of Multivariate Cusum Charts". *Journal of Quality Technology* 22, pp. 173 -186. 1990.
 11. Vargas José, Yañez Sergio, González Nelfi, Carta T^2 con base en los estimadores robustos de los parámetros. *Revista Colombiana de Estadística*. Vol. 26 No 2. p: 159 a 179. Diciembre 2002.