

UNA COMPARACIÓN DEL DESEMPEÑO DE LAS CARTAS DE CONTROL T^2 DE HOTELLING Y DE CLASIFICACIÓN POR RANGOS

A PERFORMANCE COMPARISON AMONG THE HOTELLING T^2 AND THE RANK CLASSIFICATION CONTROL CHARTS

F. Zertuche-Luis y M. Cantú-Sifuentes
Corporación Mexicana de Investigación en Materiales
División de Posgrado en Ingeniería Industrial, México
E- mails: sei01@terra.com.mx, mcantu@comimsa.com.mx

(Recibido: mayo de 2006; aceptado: diciembre de 2007)

Resumen

En el sector industrial es fundamental contar con métodos estadísticos que permitan monitorear eficientemente los procesos de producción, particularmente, cuando las características de calidad son multivariadas. Los procesos de producción son, generalmente, monitoreados mediante cartas de control, las cuales dan información gráfica de si el proceso se encuentra o no bajo control. Bajo esta situación en este trabajo se comparan, mediante simulación, el desempeño de dos tipos de cartas de control multivariadas: la clásica basada en el estadístico T^2 de Hotelling y una no-paramétrica basada en el concepto de profundidad de datos. El estudio de simulación fue motivado cuando al estudiar un proceso de la industria automotriz mediante ambos enfoques, las conclusiones a las que se llegan son diferentes. Se dan recomendaciones acerca del uso de una u otra carta de control según sea el comportamiento estadístico del proceso a monitorear.

Descriptores: Control estadístico del proceso, T^2 de Hotelling, normal multivariada, profundidad de datos, gráfica de clasificación por rangos.

Abstract

In the industrial sector it is fundamental to have statistical methods that allow efficient monitoring of production processes, particularly when the characteristics of quality are multivariate. The production processes are generally monitored by means of control charts, which give graphic information that tells if the process is, or not, under control. In this context, this article, compares, through simulation, the performance of two types of multivariate control charts: the classic chart based on Hotelling T^2 statistics, and a non-parametric chart based on the concept of data depth. We decided to perform the simulation when, studying a process within the automotive industry using both approaches, the conclusions arrived at were different. Recommendations are given about the use of one or the other control chart in accordance with the statistical behavior of the monitored process.

Keywords: Statistical process control, Hotelling T^2 , multivariate normality, data depth, r -chart.

Introducción

Los procesos industriales son generalmente monitoreados mediante el control estadístico de procesos (CEP). El CEP es la colección de métodos usados para reconocer *causas especiales de variación* y proporcionar medios para llevar un proceso a un *estado de control* y reducir la variación en torno a un valor objetivo dado. Es deseable que el proceso bajo análisis alcance este valor para cada producto. Sin embargo, en cada proceso hay una variabilidad aleatoria inherente independientemente de su diseño y de la precisión de la maquinaria usada para la producción. El objetivo del CEP es mantener la variabilidad, en torno al valor objetivo, lo más pequeña posible.

Existen al menos dos causas de variabilidad en cada proceso de producción. La causa común (o ruido blanco) es la variabilidad natural que cada proceso experimenta. Un proceso que opera solamente con este tipo de variabilidad se dice que está bajo control. Una segunda causa de variabilidad es la especial (o asignable), la cual es el resultado de factores que no son puramente aleatorios. Estos factores causan heterogeneidad en el proceso y como resultado afectan la calidad del producto. En este caso, se dice que el proceso está fuera de control. Este último tipo de variabilidad puede ser detectada mediante *cartas de control*, que proporcionan una descripción gráfica del proceso y dan a cualquier administrador, con o sin conocimiento estadístico, la información inmediata de si el proceso está o no bajo control.

Una carta de control es una representación gráfica de una o varias características del proceso bajo investigación, y es la herramienta más usada para identificar causas especiales de variabilidad en un proceso. Cuando un proceso tiene asociada una carta de control, se dice que está siendo controlado.

Cualquier tipo de carta de control debe de cumplir al menos las siguientes cuatro condiciones:

- 1) Una sola repuesta a la pregunta ¿está el proceso bajo control?,
- 2) Especificación del error tipo I global,
- 3) Se deben tomar en cuenta la relación entre las características a controlar, y
- 4) Se debe dar un procedimiento que permita responder la pregunta: Si el proceso está fuera de control ¿cuál es la causa? (Maravelakis et al., 2002).

La grafica de control comúnmente usada es la de Shewart y es útil para controlar una sola característica del producto. Cuando un producto tiene varias características a controlar, generalmente se construye una gráfica de Shewart para cada una de ellas. En este caso, hay dos suposiciones subyacentes: sus mediciones sucesivas no están correlacionadas y se distribuyen normalmente. Sin embargo, generalmente las características de un producto están correlacionadas de tal forma que cambios en una característica pueden afectar la media o la variabilidad de las restantes. Cuando se desean controlar varias características simultáneamente, la herramienta más usada es la carta de control multivariada propuesta por Hotelling basada en el estadístico T^2 de Hotelling (Pignatiello, 1993). Aquí la suposición fuerte es que las mediciones sucesivas de las características siguen una distribución normal multivariada. Sin embargo, la normalidad de las observaciones no siempre puede suponerse.

A fin de salvar estos inconvenientes, un enfoque es monitorear conjuntamente las características del producto sin la suposición de normalidad multivariada; esto es: usar cartas de control multivariadas no paramétricas. Por supuesto, éstas deberán cumplir con al menos las cuatro condiciones arriba listadas.

En el presente trabajo se estudia un proceso industrial real mediante técnicas multivariadas no paramétricas, y los resultados se comparan con los

obtenidos usando una carta de control de Hotelling. Las conclusiones motivan un estudio de simulación, cuyo objetivo es comparar el desempeño de ambas cartas. Este estudio permite dar recomendaciones acerca del uso de una u otra carta de control según sea el comportamiento estadístico del proceso a monitorear.

El proceso se ubica en el ramo automotriz y consiste en la elaboración de convertidores de torque, tales como el mostrado en la figura 1. El convertidor de torque es un dispositivo que se utiliza en los vehículos de transmisión automática para transferir la potencia que genera el motor a la caja de transmisión. Esto da por resultado el desplazamiento del vehículo. El convertidor de torque realiza una función similar a la del embrague en los vehículos de transmisión manual. Las funciones principales de este dispositivo son:

- 1) Transferir eficientemente la potencia que genera el motor,
- 2) Aumentar el Torque que es generado por el motor y
- 3) Transferir eficientemente las *rpm* del motor a la caja de transmisión.

El Convertidor de Torque es importante en el diseño de un vehículo, debido a que proporciona un mejor aprovechamiento de las fuerzas generadas por el motor para el desplazamiento eficiente del vehículo, generando un aumento de torque y un fácil cambio de velocidades en la caja de transmisión.

Los ingenieros de planta determinaron que la parte del convertidor que debería de analizarse fuera la bomba. Las características consideradas fueron: altura de pista a buje, diámetro interior de buje y diámetro exterior de buje. Las mediciones de tales características se pueden coleccionar en un vector tridimensional z donde la primera componente es la altura pista-buje, la segunda el diámetro interior y la tercera el diámetro exterior. Estas características están naturalmente correlacionadas y, por lo tanto, las gráficas de control univariadas no son adecuadas.

Análisis del proceso mediante la gráfica de control de Hotelling

Las gráficas de control multivariadas están diseñadas para monitorear el proceso de producción de un producto, el cual interesa controlar, con p variables de calidad posiblemente correlacionadas. La gráfica de control más usada es la llamada T^2 de Hotelling. Para su construcción son necesarias

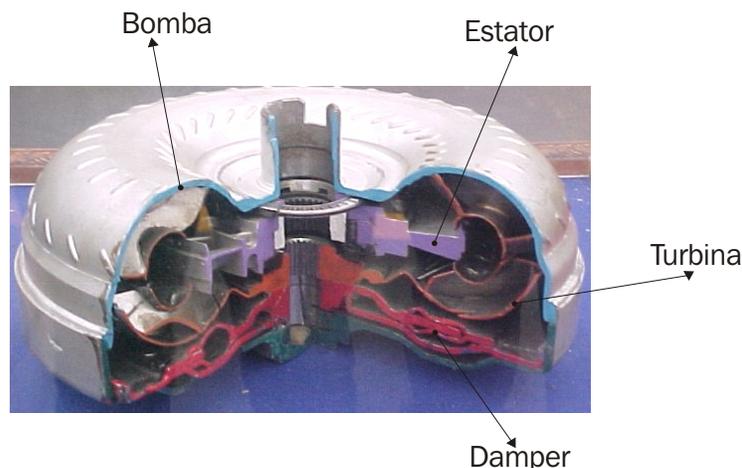


Figura 1. Convertidor de Torque

dos fases: En la Fase I se considera un conjunto de m datos históricos supuestamente bajo control y con distribución normal multivariada; dichos datos históricos nos servirán para estimar el vector de medias \bar{X} , y la matriz de varianzas – covarianzas S (Mason et al., 2002). Los estimadores resultan ser, respectivamente, $\bar{X} = (1/n)X'\bar{1}$, con X una matriz $n \times p$ de datos, donde $n = m$ es el tamaño de un subconjunto adecuado de los datos históricos, y $S = (1/(n-1))(X'HX)$, con $H = I - (1/n)Q\bar{1}\bar{1}'$ y $\bar{1}$ representado un vector de unos, conformable al producto actual.

Por otra parte, en la Fase II se calcula el estadístico T^2 de Hotelling para las mediciones actuales de las variables a controlar, X_1, X_2, \dots definido mediante la expresión:

$$T_i^2 = (X_i - \bar{X})' S^{-1} (X_i - \bar{X}) \quad (1)$$

Se puede demostrar que, bajo la suposición de normalidad, T_1^2, T_2^2, \dots conforman una muestra aleatoria proveniente de una distribución F (Mason et al., 2002). Este resultado se utiliza, dada $F_{(\alpha; p, n-p)}$, para fijar el límite de control en la Fase II, calculado como:

$$UCL = \frac{p(n-1)(n-1)}{n(n-p)} F_{(\alpha; p, n-p)} \quad (2)$$

donde

n es el tamaño de la muestra del conjunto de datos históricos y

$F_{(\alpha; p, n-p)}$ es el α -ésimo cuantil de una distribución F con p y $n-p$ grados de libertad.

Los valores de esta muestra aleatoria se grafican conjuntamente con el límite de control determinado para completar la construcción de una gráfica de control T^2 de Hotelling.

La metodología arriba descrita se aplicó a $m=127$ mediciones históricas de convertidores de torque y se determinó que

$$n=127, \bar{X}=(2.8363, 1.9151, 2.5199),$$

y		0.6654	0.0084	0.0122
S	10^6	0.0084	0.0257	0.0085
		0.0122	0.0085	0.0869

El límite de control se fijó en 17.7256. En la figura 2 se muestra la gráfica de control para 25 datos monitoreados.

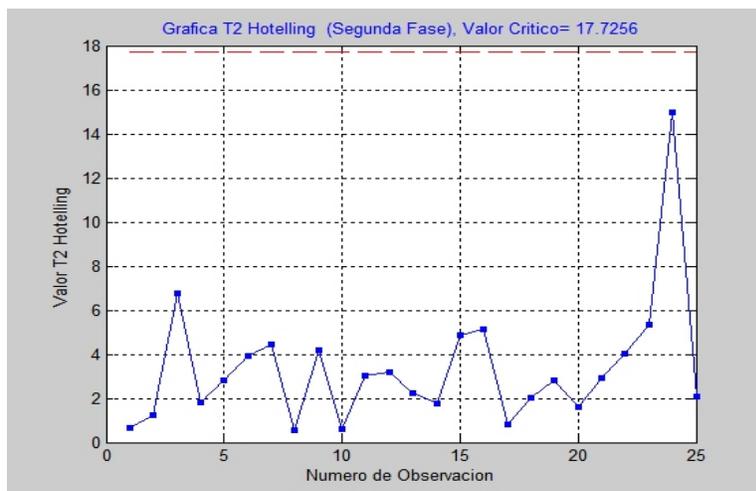


Figura 2. Gráfica de control T^2 de Hotelling para 25 observaciones de las tres variables de calidad de convertidores de torque

En la figura 2 se observa que ningún punto está fuera de control, y se concluye que el proceso está controlado con respecto a la distribución de referencia, en este caso no se determinó si la distribución de referencia en forma conjunta seguía una distribución normal multivariada, la cual es un supuesto para poder aplicar en forma correcta la metodología, en la siguiente sección se analiza la distribución de referencia para verificar la existencia de una distribución normal multivariada y asegurarse de que el análisis de la gráfica T^2 de Hotelling era el correcto.

Normal multivariada

El supuesto de normalidad es usual en el análisis estadístico clásico, tanto univariado como multivariado. En particular, en la construcción de gráficas de control de Hotelling el supuesto de normalidad multivariada es central.

Se han desarrollado una gran cantidad de métodos, tanto formales como informales (gráficos), para probar la normalidad para datos multivariados. De los formales, uno de los más potentes es el de Henze-Zirkler. (Mecklin, 2004). Por otra parte, los métodos gráficos permiten, de un golpe de vista, decidir si es razonable la suposición de normalidad. De éstos, quizá los más utilizados son las Gráficas cuantil-cuantil, o QQ-plot en el lenguaje inglés.

En esta sección se describen y se aplican a los datos del convertidor de torque el método de Henze-Zirkler y la construcción de la gráfica cuantil-cuantil.

Gráfica cuantil-cuantil

Como es sabido, siendo x_1, x_2, \dots, x_n una muestra de tamaño n , la gráfica cuantil-cuantil consiste en construir un diagrama de dispersión de los puntos $(\hat{F}(x_i), F(x_i))$ $i = 1, 2, \dots, n$, con

$$\hat{F}(x) = \frac{1}{n}$$

número de observaciones menores o iguales que x

y

$$F(x) = \int_{-\infty}^x f_x(x) dx$$

Por supuesto, $f_x(x)$ representa la densidad de la distribución de interés. Entre más los puntos del diagrama se parezcan a una línea recta a cuarenta y cinco grados, más evidencia se tiene que la muestra proviene de la densidad $f_x(x)$.

Cuando se desea ver si el supuesto de normalidad multivariada es aceptable, se puede construir una gráfica cuantil-cuantil usando el hecho de que bajo normalidad, las cantidades dadas por:

$$d_i^2 = (x_i - \bar{x})' S^{-1} (x_i - \bar{x}) \tag{3}$$

La cual se distribuye como una χ_n^2 (n), con n grados de libertad. Por esta razón, a esta gráfica se le conoce con el nombre de Gráfica cuantil-cuantil o gráfica Ji-cuadrada (Koziol, 1993). En la figura 3, se muestra la gráfica en papel normal para los datos del convertidor de torque.

En la figura 3 se puede observar poca evidencia que apoye el supuesto de normalidad multivariada en los datos del convertidor de torque. Si todos los datos se vieran como los primeros en la gráfica, el supuesto de normalidad sería razonable.

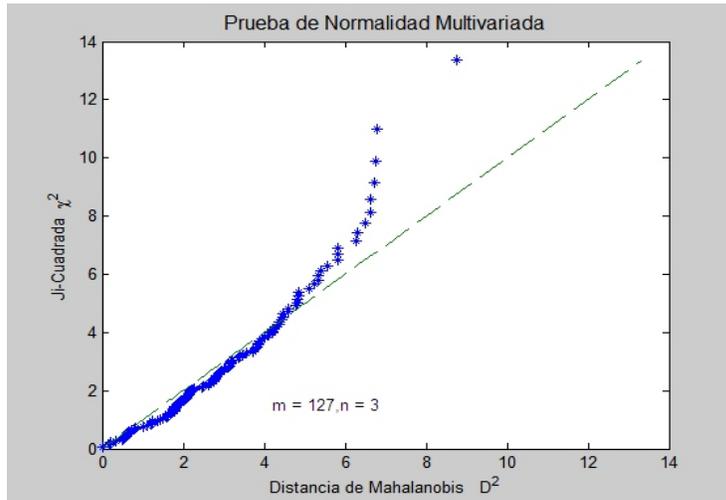


Figura 3. Gráfica cuantil-cuantil para los datos del convertidor de torque

Prueba Henze-Zirkler

Una prueba estadística para determinar si un conjunto de datos sigue una distribución normal multivariada, conocida por su gran potencia, es la prueba de Henze-Zirkler (Mecklin, 2004). Si x_1, x_2, \dots, x_n denota una muestra aleatoria de una distribución p variada, entonces, el estadístico de la prueba de Henze-Zirkler se define mediante la expresión (Henze et al., 1990):

$$T = \frac{1}{n} \sum_{j,k=1}^n \exp \left\{ -\frac{1}{2} |y_j - y_k|^2 \right\} \frac{1}{2(1 - 2^{-2})} \quad (4)$$

$$\sum_{j=1}^n \exp \left\{ -\frac{1}{2(1 - 2^{-2})} |y_j|^2 \right\} \frac{1}{n(1 - 2^{-2})} \quad (4)$$

Donde,

$$|y_j - y_k|^2 = (x_j - x_k)' S^{-1} (x_j - x_k), |y_j|^2 = (x_j - \bar{x})' S^{-1} (x_j - \bar{x})$$

$$(x_j - \bar{x})' S^{-1} (x_j - \bar{x}) \text{ y } \bar{x}, S$$

denotan el vector de medias muestrales y la matriz de varianzas-covarianzas respectivamente. El parámetro λ es un parámetro de suavizamiento, determinado por el tamaño de la muestra y definido de la forma siguiente:

$$\lambda(n) = \frac{1}{\sqrt{2}} \frac{2p - 1}{4} n^{1/(p-4)} \quad (5)$$

Si el conjunto de datos sigue una distribución normal multivariada, el estadístico de prueba T es aproximadamente distribuido lognormal con:

$$E[T] = \frac{1}{2} (1 - 2^{-2})^{p/2} \quad (6)$$

$$\frac{1}{2} \frac{p - 2}{1 - 2^{-2}} \frac{p(p - 2)^4}{2(1 - 2^{-2})^2}$$

$$Var[T] = \frac{2(1 - 4^{-2})^{p/2}}{2(1 - 2^{-2})^p} \frac{1}{2} \frac{2p^4}{(1 - 2^{-2})^2} \quad (7)$$

$$\frac{3p(p - 2)^8}{4(1 - 2^{-2})^4} \frac{4w^{p/2}}{1} \frac{3p^4}{2w} \frac{p(p - 2)^8}{2w^2}$$

donde $w = (1 + B^2)(1 + 3B^2)$. Estos resultados son usados para contrastar el siguiente juego de hipótesis:

H_0 : Los datos siguen una distribución Normal Multivariada vs.

Ha: Los datos no siguen una distribución Normal Multivariada.

Si, $T < c_1$, con c_1 el cuantil $(1 - \alpha)$ de una distribución lognormal con media $E[T]$ y varianza $var[T]$, se rechaza la hipótesis nula y se acepta en caso contrario.

Para los datos del convertidor de torque, y con $\alpha = 0.05$, se obtienen, para la prueba de Henze-Zirkler, los resultados mostrados en la tabla 1.

Según los resultados de la tabla 1, se rechaza la hipótesis nula. Esto es, no es razonable suponer que los datos pueden modelarse mediante una distribución normal multivariada; con lo que el enfoque de la T^2 de Hotelling para construir una gráfica de control para tales datos no es adecuado. Aún así, sigue siendo deseable el contar con una gráfica de control del proceso de producción de convertidores de torque. Se puede optar por un enfoque no paramétrico, en el cual la suposición de normalidad no es necesaria. En el apartado siguiente se describe y aplica un método no paramétrico basado en el concepto de profundidad de datos.

Gráficas de control basadas en la profundidad de datos

Entre los métodos no paramétricos usados en el análisis de datos se encuentran los métodos de profundidad de datos. Éstos empezaron prácticamente con el artículo de Liu (1992). Quizás una de las aplicaciones principales de esta metodología está en el control de calidad. El único requerimiento para usar este método, es contar con una *distribución de referencia* denotada como F , que describe una distribución k ($k \geq 1$) dimensional, en la práctica ésta suposición se traduce en contar

con una muestra Y_1, \dots, Y_m , llamada muestra de referencia de vectores k dimensionales (Parelius, 1999)

La noción de profundidad de datos se basa en el hecho de que cualquier densidad de probabilidades distingue entre puntos “centrales” y “periféricos”. Una función de profundidad asigna a cada $y \in R^k$ un *puntaje* no negativo, el cual puede interpretarse como su localización en la nube de datos. Las profundidades más grandes corresponden al centro de la distribución, mientras que las más pequeñas corresponden a regiones externas. Este sistema de *clasificación por rangos* es usado para determinar si una observación es diferente de aquellas que conforman el conjunto original. Las funciones de profundidad deben satisfacer un número de propiedades (Zuo et al., 2000): invariante afin, monotonicidad, maximalidad al centro, desvanecimiento al infinito. Existen funciones que cumplen con estas propiedades, entre otras, la profundidad de Mahalanobis, la profundidad simplicial y la profundidad de Tukey. De éstas, la más usada, quizá debido a su facilidad de cómputo e interpretación, es la profundidad de Mahalanobis denotada por $MD_F(y)$, y definida como (Liu et al., 1993):

$$MD_F(y) = \frac{1}{[1 + (y - \mu_F)' \Sigma_F^{-1} (y - \mu_F)]} \quad (8)$$

Donde,

μ_F es el vector de medias y

Σ_F^{-1} es la matriz de varianzas-covarianzas de la distribución de referencia F .

Si los parámetros de la distribución de referencia no son conocidos, la versión muestral de la

Tabla 1. Resultados de la prueba Henze-Zirkler

Tamaño de muestra	Variables	Estadístico H-Z (T)	Valor crítico (c_1)	Valor de p $P(T < 2.8183)$
127	3	2.8183	2.6644	0.0211

profundidad de Mahalanobis dada por la ecuación (9) es:

$$MD_{F_m}(y) = \frac{1}{[1 (y - \bar{Y})' S^{-1} (y - \bar{Y})]} \quad (9)$$

Donde

\bar{Y} es el vector de medias muestrales de los datos

Y_1, \dots, Y_m y S es la matriz de varianzas-covarianzas de la muestra de referencia F_m .

Para construir una gráfica de control basada en la profundidad de Mahalanobis, llamada comúnmente gráfica de clasificación por rangos (o r-chart en el idioma inglés), se realiza el siguiente procedimiento:

1. Se calculan el vector de medias, la matriz de varianzas-covarianzas y la profundidad de cada uno de los datos en F_m , $MD_{F_m}(y_i) \ i=1,2,\dots,m$.

2. Se obtienen los estadísticos de orden de las $MD_{F_m}(y_i) \ i=1,2,\dots,m$, y se denotan como $y_{[1]}, y_{[2]}, \dots, y_{[m]}$.

3. Sean X_1, X_2, \dots las nuevas observaciones, las cuales se suponen siguen una distribución continua G . Para cada X_i se obtiene su profundidad, aplicando la ecuación (9), utilizando el vector de medias y la matriz de varianzas-covarianzas, de F_m , las cuales se denotan mediante $MD_{F_m}(x_i) \ i=1,2,\dots$

4. El estadístico de clasificación por rangos, $r()$, para cada nueva observación se calcula mediante:

$$r(x_i) \quad (10)$$

$$\# \frac{y_j | MD_{F_m}(y_j) \ MD_{F_m}(x_i), j \ 1, \dots, m}{m \ 1}$$

5. Se grafican los estadísticos de clasificación por rangos de cada X_i contra el tiempo, con un límite central de control CL=0.5 y un límite inferior de control LCL= . Donde es llamada la proporción de alarma.

Liu (1995) demostró que el estadístico de clasificación por rangos sigue una distribución uniforme en el intervalo [0,1]. La gráfica de clasificación por rangos, llamada también gráfica r , contrasta sucesivamente las hipótesis:

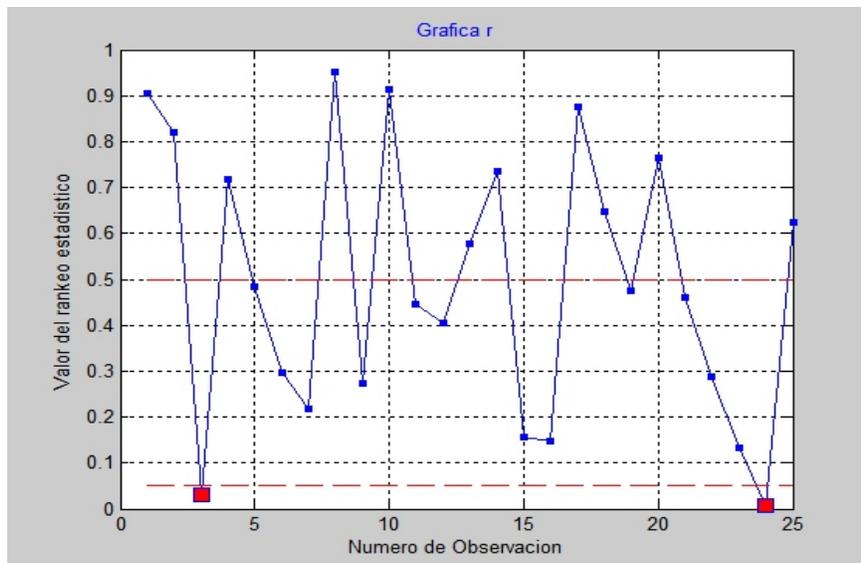


Figura 4. Gráfica de clasificación por rangos, para los datos del convertidor de torque

Ho: La nueva observación tiene la misma distribución que la distribución de referencia vs.

Ha: Existe un cambio en la ubicación o dispersión de la distribución de la nueva observación con respecto a la distribución de referencia.

Si $r(x_i)$ se grafica por debajo del LCL, entonces la observación se declara fuera de control. Esto es una señal de alerta para los ingenieros del proceso. En la figura 4 se muestra la gráfica de clasificación por rangos para los datos del convertidor de torque.

Comparación entre las gráficas T^2 de Hotelling y de clasificación por rangos

Conforme a los resultados en los apartados anteriores, es claro que las conclusiones que se pueden extraer de un análisis de datos dependen de qué atención se ponga a la verificación de los supuestos del modelo usado. En este caso, el uso del enfoque de Hotelling en la construcción de gráficas de control sin verificar la normalidad de las variables, conduce a la conclusión de que no hay artículos (convertidores de torque) fuera del límite de control. Sin embargo, un enfoque no paramétrico indica que dos artículos están fuera de control. Esto conduce a las siguientes preguntas: Si los datos hubiesen sido normales ¿Cuál de los dos

enfoques proporciona resultados más adecuados? y ¿cuál es el desempeño relativo de ambas gráficas a medida que los datos se alejan de la normalidad? Estas preguntas se pueden responder mediante un estudio de simulación. La exposición de los resultados de un estudio tal es el objeto del siguiente apartado.

Los métodos T^2 de Hotelling y no paramétrico fueron aplicados a dos diferentes conjuntos de datos. El primer conjunto de datos fue obtenido mediante un generador normal multivariado con

$$\bar{X} = (10.0021, 10.4723, 10.5201, 10.2671, 10.3410)$$

y

S=

3.6176	0.4220	0.2377	0.1340	0.6456
0.4220	3.4991	0.0363	0.1754	0.4368
0.2377	0.0363	3.8034	0.0958	0.2313
0.1340	0.1754	0.0958	3.7951	0.2406
0.6456	0.4368	0.2313	0.2406	3.1382

Las cinco variables del segundo conjunto de datos se generaron como sigue: La primera sigue una distribución uniforme en el intervalo [10,12]. Las cuatro restantes se generaron como una com-

Tabla 2. ARL para el primer conjunto (normal)

Cambio en la medida de las variables	Pequeño (1)	Medio (2)	Grande (3)
ARL para las gráficas de rangos	8.0401	2.7301	0.0222
ARL para las gráficas T^2 de Hotelling	5.2051	1.9632	0.0179
Eficiencia relativa (ARL T^2 / ARL rangos)	0.647	0.719	0.806

Tabla 3. ARL para el segundo conjunto (no normal)

Cambio en la media de la 1a. variable original	Pequeño	Mediano	Grande
ARL para las gráficas de rangos	15.026	5.7708	3.4179
ARL para las gráficas T^2 de Hotelling	500	37.6589	5.5748
Eficiencia relativa (ARL T^2 / ARL rangos)	33.275	6.525	1.631

binación de la primera y una variable uniforme $[0,1]$. Para cada conjunto de datos, se generó una distribución de referencia F de 300 observaciones. Después se generaron 500 nuevas observaciones para cada conjunto de datos, con diferentes cambios en los valores de las medias de las variables. Estas observaciones constituyen las distribuciones de monitoreo G . Las nuevas observaciones fueron monitoreadas mediante la gráfica T^2 de Hotelling, y la carta de clasificación por rangos, con un $\alpha=0.05$. Para comparar la eficiencia relativa entre los dos métodos se usó la longitud promedio de corridas (ARL). Ésta se define como el promedio de puntos dentro de control antes de observar un punto fuera de control. El ARL fue calculado con 1,000 simulaciones para cada uno de los cambios en la distribución de monitoreo: pequeño, medio y grande. En las tablas 2 y 3 se presentan el ARL para cada caso y para cada tipo de gráfica.

Cuando las variables son generadas de una distribución normal, los resultados de la eficiencia relativa en la tabla 2 indican que el ARL de la gráfica T^2 de Hotelling es más pequeño que el ARL del método no-paramétrico. Sin embargo, cuando el cambio en la media de las variables incrementa, el ARL de ambos métodos tiende a ser igual. Por otra parte, cuando las variables son generadas de una distribución no-normal, los resultados de la eficiencia relativa en la tabla 3 indican que el desempeño del método no paramétrico es mejor que la T^2 de Hotelling.

Conclusiones y trabajo futuro

1. Si los datos del convertidor de torque, no normales, según la Prueba de Henze-Zirkler, se monitorean mediante la Gráfica T^2 de Hotelling se encuentra que ninguno de ellos se sale de control. Sin embargo, cuando se monitorean mediante la gráfica clasificación por rangos se encuentra que hay datos fuera de control. Entonces, se esperaría que los indicadores de calidad sean mejores cuando se calculan mediante el método de Hotelling, y los ingenieros de planta estarían cometiendo un error de estimación acerca del proceso.

2. Los resultados de la simulación muestran que el desempeño del método no paramétrico es mejor que el T^2 de Hotelling cuando los datos son provenientes de una distribución no-normal multivariada. Por otro lado, el desempeño de la Gráfica T^2 de Hotelling es más eficiente cuando los datos provienen de una distribución normal multivariada, a menos que el cambio en la media sea grande, pues en tal caso ambas gráficas parecen tener un desempeño similar.

Después de una amplia revisión bibliográfica no se encontró una metodología para calcular indicadores de calidad para cada uno de los dos enfoques. Por otra parte, en el método no paramétrico no se encontró un método práctico que conduzca a determinar la componente de más influencia en los puntos fuera de control.

Esto marca pautas para un trabajo futuro: proponer indicadores de calidad para cada uno de los enfoques; y para el caso de la gráfica de clasificación por rangos encontrar una metodología para determinar la variable de más influencia.

Referencias

- Henze N. and Zirkler B. A class of invariant consistent test for multivariate normality. *Commun. Statist-Theor. Meth.*, 19(10):3595-3618. 1990.
- Kozioł J. Probability plots for assessing multivariate normality. *The Statistician* 42. Pp161-174. 1993.
- Liu R. Data depth and multivariate rank test. *Statistics and Related Methods*. Pp. 279-294. 1992.
- Liu R. A quality index based on data depth and multivariate rank test. *Journal American Statistics Association*. 91:266-277. 1993.
- Liu R. Control chart for multivariate processes. *Journal American Statistics Association*. 90: 209-125. 1995.
- Maravelakis P.E. and Panaretos J. Effect of estimation of the process parameters on the control limits of the univariate control chart for process dispersion. *Communications in Statistics*, 31(3): 443-462. 2002.

Mecklin-Christopher J. An appraisal and bibliography of test for multivariate normality. *International Statistical Review*, 72(1):123-138. 2004.

Parelius-Jesse M. Multivariate analysis by data depth: Descriptive statistics, graphics and inference. *The Annals of Statistics*, 27(3):783-858. 1999.

Pignatiello J.J. On constructing T^2 control chart for on-line process monitoring. *IIE Transactions*, 31:529-536. 1993.

Zuo Y. and Serfling R. Structural properties and converge results for contours of sample statistical depth functions. *The annals of statistics*, 28(2):483-499, 2000.

Bibliografía sugerida

Mason L.R. and Young C.J. *Multivariate statistical process control with industrial applications*. Ed. ASA SIAM. 2002.

Semblanza de los autores

Federico Zertuche-Luis. Realizó la licenciatura en ingeniería industrial en el Instituto Tecnológico de Saltillo en 1992. El I.T.E.S.M (Campus Monterrey) en 1994, le otorgó el grado de maestro en ciencias en ingeniería industrial. Actualmente es candidato a doctor en ingeniería industrial y de manufactura en el Programa Interinstitucional en Ciencia y Tecnología, impartido en la Corporación Mexicana en Investigación de Materiales. Desde 1996, se desempeña como profesor investigador en el Instituto Tecnológico de Saltillo, dentro del área de Posgrado en Ingeniería Industrial. A partir de 2003, es miembro del sistema de investigación del estado de Coahuila. Como consultor ha participado en la implementación de la metodología Seis Sigma en las siguientes plantas: Mabe (Saltillo), Siemens VDO, FEMSA, Lear (Saltillo), Lear (Ramos Arizpe), Linamar (Saltillo), Linamar (Torreón), Sanmina (Apodaca).

Mario Cantú-Sifuentes. Obtuvo su licenciatura en ingeniería industrial en el Instituto Tecnológico de Saltillo en 1976. Asimismo, el grado de maestría en ciencias en estadística en la Universidad Autónoma Agraria Antonio Narro y el grado de doctor en ciencias en estadística en el Colegio de Posgraduados. De 1977 a 1983, realizó diversos trabajos en la iniciativa privada y dependencias gubernamentales. Desde 1985, se desempeña como profesor investigador el Departamento de Estadística de la Universidad Autónoma Agraria Antonio Narro.