Ingeniería Investigación y Tecnología
volumen XXI (número 3), julio-septiembre 2020 1-12
ISSN 2594-0732 FI-UNAM artículo arbitrado
Información del artículo: Recibido: 24 de enero de 2020, aceptado: 29 de mayo de 2020

https://doi.org/10.22201/fi.25940732e.2020.21.3.023

Evaluación de una carta de control multivariada basada en profundidad de datos para observaciones no normales en presencia de autocorrelación

Evaluation of a multivariate control chart based on data depth for non-normal observations in the presence of autocorrelation

Olmos-Zepeda José Rodolfo
Colegio de Postgraduados, México
Departamento de Estadística
Correo: olmos.rodolfo@colpos.mx
https://orcid.org//0000-0002-3877-9756

Ramírez-Valverde Gustavo
Colegio de Postgraduados, México
Departamento de Estadística
Correo: gramirez@colpos.mx
https://orcid.org/0000-0003-3466-991X

Resumen

La carta de control T^2 de Hotelling es ampliamente utilizada para monitorear, simultáneamente, dos o más características de calidad continuas. Para que su uso sea válido se requiere que los datos a supervisar cumplan con dos suposiciones: normalidad multivariada e independencia entre observaciones, sin embargo, ambos supuestos son difíciles de sostener en un proceso real y, en consecuencia, el rendimiento de la gráfica de control se deteriora. Liu (1995) desarrolló una clase de cartas de control multivariadas no paramétricas o de distribución libre basadas en el concepto de profundidad de datos. La única suposición para usar estas gráficas es que las observaciones sean independientes. En este trabajo se estudia vía simulación, a través de la longitud promedio de corrida, el comportamiento de la carta de control basada en medidas profundidad cuando se relaja el supuesto de normalidad multivariada en presencia de autocorrelación, su comportamiento se compara con el de la gráfica T^2. Se encontró que la carta de control no paramétrica es menos afectada en la mayoría de los casos estudiados que el método paramétrico.
Descriptores: Carta de control, carta T^2 de Hotelling, profundidad de datos, profundidad de Mahalanobis, autocorrelación, carta de clasificación por rangos.

Abstract

The Hotelling T^2 control chart is widely used to simultaneously monitor two or more continuous quality features. For its use to be valid, the data to be monitored is required to met two assumptions: multivariate normality and independence between observations, however, both assumptions are difficult to sustain in a real process and, consequently, the performance of the control chart it deteriorates. Liu (1995) developed a class of multivariate nonparametric or free distribution control charts based on the concept of data depth. The only assumption to use these graphs is that the observations are independent. In this work, the behavior of the control chart based on depth measurements is studied via simulation, when the assumption of multivariate normality is relaxed in the presence of autocorrelation, its behavior is compared with that of the graph T^2. It was found that the non-parametric control chart is less affected in most of the cases studied than the parametric method.
Keywords: Control chart, Hotelling T^2 control chart, data depth, Mahalanobis depth, autocorrelation, *r*-chart.

Introducción

Cuando en un proceso de producción se monitorean dos o más características de calidad (o variables) continuas, posiblemente correlacionadas, el uso de múltiples procedimientos de control univariados tipo Shewhart es inadecuado, ya que, además de que no controlan la tasa general de falsas alarmas (Boone, 2010), tampoco toman en cuenta la correlación que con frecuencia presentan las variables en procesos multivariados (Mason y Young, 2002). Una recomendable alternativa al control univariado para el monitoreo de diferentes variables simultáneamente, son las cartas de control multivariadas. Entre las gráficas de control multivariadas más conocidas destacan la carta T^2 de Hotelling, la carta multivariada de promedios móviles exponencialmente ponderados (MEWMA) y la carta multivariada de sumas acumuladas (MCUSUM), de estas, la más popular es la gráfica T^2. El empleo de estos métodos de control se basa en dos suposiciones: que las observaciones sucesivas de las características de calidad a supervisar sean independientes (o no autocorrelacionadas) y que sigan una distribución normal multivariada. Sin embargo, a menudo en la práctica, los datos exhiben cierta dependencia en serie y un comportamiento distribucional que no puede suponerse normal. Particularmente el rendimiento de la carta T^2, cuando se ignora la presencia de autocorrelación o falta de normalidad, se afecta severamente (Montgomery, 2009). Chou *et al.* (2001) detectaron, por ejemplo, que la carta T^2 incrementa la proporción de falsas alarmas cuando el proceso está en control y las observaciones no provienen de una población normal multivariada. Vanhatalo y Kulahci (2015) observaron que la presencia de autocorrelación en datos normales multivariados disminuye la potencialidad de detección de cambios en la gráfica T^2 cuando el proceso está fuera de control. A fin de resarcir estos daños se han introducido en la literatura cartas de control alternativas y algunas modificaciones de la carta T^2. El método más conocido que se utiliza para atacar el problema de la autocorrelación consiste en ajustar primero un modelo de vectores autorregresivos de orden uno (VAR(1)) y después monitorear el proceso con una gráfica T^2 usando los residuales del modelo ajustado. Vanhatalo y Kulahci (2015) mostraron que, en general, el enfoque basado en residuales mejora el rendimiento de la carta T^2 que cuando se usan los datos originales. Sin embargo, este procedimiento es poco práctico y supone que el modelo VAR(1) es el adecuado para describir los datos. Respecto a la no normalidad, Chou *et al*. (2001) propusieron determinar el límite de control de la gráfica T^2 mediante la técnica de bootstrap. Esta modificación mostró ser muy eficiente comparada contra la auténtica carta T^2 para datos simétricos y asimétricos.

Un enfoque no paramétrico o de distribución libre de cartas de control multivariadas fue el desarrollado por Liu (1995), que se basa en el concepto de profundidad de datos. De acuerdo con Liu (1990), la profundidad de un dato se define como una medida de cuán profundo o central es un punto dado respecto a una distribución multivariada. La carta de control más notable dentro de este enfoque se conoce como gráfica de clasificación por rangos o carta *r*, la cual ha sido implementada bajo la profundidad de Mahalanobis. El supuesto básico que se requiere para usar esta carta de control es que las observaciones son independientes. Zertuche y Cantú (2008) compararon las cartas *r* y T^2 en datos distribuidos normal y no normal, y encontraron que cuando los datos no tienen distribución normal, el método no paramétrico es más eficiente que la gráfica T^2, pero cuando su distribución es aproximadamente normal, entonces su eficiencia es similar. Sin embargo, se desconoce el desempeño de la carta *r* cuando las mediciones están autocorrelacionadas.

El objetivo del presente artículo es evaluar, mediante simulación, el rendimiento de la gráfica de clasificación por rangos basada en la profundidad de Mahalanobis para datos que no cumplen con los supuestos de normalidad multivariada en presencia de autocorrelación y compararlo con el desempeño de la carta T^2 bajo las mismas condiciones. En la siguiente sección se proporcionan los procedimientos para construir las cartas de control T^2 y *r*, la metodología que se utiliza para simular datos autocorrelacionados y se describe el estudio de simulación para evaluar las gráficas de control. En la tercera sección se presentan y discuten los resultados y, finalmente, se dan las conclusiones de este trabajo.

Desarrollo

Para evaluar y comparar la eficiencia de las gráficas *r* y T^2 en presencia de autocorrelación, se realizó un estudio de simulación Monte Carlo en el que se consideraron los siguientes factores de estudio: estado del proceso (en control y fuera de control) grado de autocorrelación (con y sin independencia), distribución de los datos (datos normales y no normales) y relación entre las características de calidad (con y sin correlación en las características de interés).

Estudio de simulación

La comparación entre gráficas de control se realizó utilizando la longitud promedio de corrida (ARL). Cuando el proceso está en control, el ARL indica el número promedio de muestras que deberían tomarse para que aparezca una falsa alarma. Cuando no está en control el proceso, entonces el ARL es el número promedio de muestras que ocurriría hasta detectar la falta de control.

Sin pérdida de generalidad, las cartas de control se estudiaron simulando un proceso con dos características de calidad continuas a monitorear, la media de cada una de esas variables fue cero, la matriz de varianzas y covarianzas dependió de la situación simulada. Para cada una de las situaciones estudiadas, la fase I se construyó con una muestra de tamaño $n_1 = 1500$. Este punto es congruente con lo reportado por Vanhatalo y Kulahci (2015) respecto al número de observaciones que se requieren en la fase I para aproximarse al valor del ARL de 370. Posteriormente, en cada situación se simularon observaciones hasta detectar una señal fuera de control, registrando el número de realizaciones necesarias para lograrlo. Esto se repitió 1000 veces y el promedio es la estimación del ARL.

Los límites de las cartas de control se construyeron con los valores más usados en la práctica que corresponden a los límites que dejan una proporción de falsas alarmas de 0.0027 cuando se cumplen los supuestos de normalidad e independencia, lo que significa un ARL de 370 aproximadamente.

Métodos de control analizados

Carta de control T^2 de Hotelling

Esta carta de control se utiliza ampliamente para monitorear el vector de medias de un proceso. Sea $\boldsymbol{X}$ un vector aleatorio de orden p definido en $\mathbb{R}^p$ de la forma $\boldsymbol{X} = (X_1, X_2, \dots, X_p)^T$, donde cada X_i, para $i = 1, 2, \dots, p$, es una variable aleatoria unidimensional definida en $\mathbb{R}$. Suponga que $\boldsymbol{X}$ tiene distribución normal p-variada con vector de medias μ y matriz de varianzas y covarianzas Σ no singular, simétrica y positiva definida. Entonces, la estadística T^2 de Hotelling está dada por:

$$T^2 = (\boldsymbol{X} - \mu)^T \Sigma^{-1} (\boldsymbol{X} - \mu) \tag{1}$$

Cuando μ y Σ se conocen, la distribución de la estadística en la expresión (1) es una ji-cuadrada con ρ grados de libertad (Mason y Young, 2002). Sin embargo, en las aplicaciones estos parámetros se desconocen y se estiman por $\overline{X}$ y S, respectivamente, donde en una muestra aleatoria p-variada de tamaño n $X_1, X_2, \dots, X_n$, se tiene que:

$$\overline{\boldsymbol{X}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{X}_i \quad \text{y} \quad S = \frac{1}{n-1}\sum_{i=1}^{n} (\boldsymbol{X}_i - \overline{\boldsymbol{X}})(\boldsymbol{X}_i - \overline{\boldsymbol{X}})^T \tag{2}$$

Entonces, la estadística T^2 queda de la forma:

$$T^2 = (\boldsymbol{X} - \overline{\boldsymbol{X}})^T \boldsymbol{S}^{-1} (\boldsymbol{X} - \overline{\boldsymbol{X}}) \tag{3}$$

y en este caso su distribución es (Mason y Young, 2002):

$$\left\{ \frac{p(n+1)(n-1)}{n(n-p)} \right\} \times F_{(p,n-p)} \tag{4}$$

donde $F_{(p,\, n-p)}$ denota la distribución de F de Fisher-Snedecor con parámetros p y $n-p$ grados de libertad. El monitoreo de un proceso con la gráfica T^2 se realiza en dos fases. En una primera fase con el proceso en control (distribución de referencia) se define una base de datos histórica (BDH) de tamaño n, cuya distribución se usa para monitorear observaciones futuras (fase II). Con la información de la BDH se estiman μ y Σ y se obtiene el límite de control superior (LCS) dado por:

$$LCS = \left\{ \frac{p(n+1)(n-1)}{n(n-p)} \right\} \times F_{(\alpha,p,n-p)} \tag{5}$$

en el cual $F_{(\alpha;\, p.n-p)}$ es el α-ésimo cuantil superior de la distribución F de Fisher-Snedecor con parámetros p y $n - p$ grados de libertad. α es la proporción de falsas alarmas. Para cada observación en la fase II (fase de monitoreo), se calcula la estadística T^2 como se señala en la expresión (3), en tanto los puntos de la estadística T^2 no excedan el LCS, el proceso se declara bajo control. Si algún punto está por arriba del LCS, el proceso se declara fuera de control y es deseable localizar las causas responsables de la señal de fuera de control.

Carta de control basada en medidas de profundidad

Una medida de profundidad es una función que asigna un rango no negativo en el intervalo [0, 1] que provee un orden del centro hacia afuera (centro-exterior) de un punto dado $\boldsymbol{x} \in \mathbb{R}^p$ respecto a una nube de datos. Valores grandes de profundidad indican que $\boldsymbol{x}$ está más cerca del centro de la nube de puntos y valores pequeños que se encuentra ubicada en las regiones externas. Bajo este hecho, Liu y Singh (1993) propusieron una estadística definida por funciones de profundidad para comparar la periferia de una población multivariada en relación con otra, la cual trasladaron al esquema de cartas de control para verificar si el proceso que se está supervisando cumple con las características de la distribución de referencia. Existen varias funciones de profundidad, quizás la más popular de estas es la profundidad de Mahalanobis de un punto $x \in \mathbb{R}^p$ que se define como (Liu y Singh, 1993):

$$MD_H(x) = \frac{1}{\{1+(x-\mu_H)^T \Sigma_H^{-1}(\boldsymbol{x}-\mu_H)\}} \tag{6}$$

donde μ_H y Σ_H son el vector de medias y la matriz de varianzas y covarianzas, respectivamente, de la distribución de referencia H. Cuando estos parámetros son desconocidos, la versión muestral de la profundidad de Mahalanobis expresada en (7) es:

$$MD_{H_m}(x) = \frac{1}{\{1+(x-\bar{\boldsymbol{X}})^T \boldsymbol{S}_H^{-1}(\boldsymbol{x}-\bar{\boldsymbol{X}})\}} \tag{7}$$

donde $\bar{\boldsymbol{X}}$ y $\boldsymbol{S}$ son los estimadores muestrales expresados en (2) para una muestra aleatoria de la distribución de referencia de tamaño m denotada por H_m.

Aplicaciones de la carta de control que Liu (1995) propone, se registran en los trabajos de Hamurkaroglu *et al*. (2004); Messaoud *et al*. (2005); Zertuche y Cantú (2008). Para construir una carta de control basada en la profundidad de Mahalanobis se realiza el siguiente procedimiento (Zertuche y Cantú, 2008). Sea $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_m$ una muestra aleatoria p-variada de tamaño m de la distribución de referencia H (fase I). Sea $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots,$ (fase II) la muestra aleatoria de las p variables a monitorear cuya distribución es G. Si el proceso está en control se tendrá que las distribuciones H y G son iguales, pero si el proceso no está en control, la distribución de los puntos monitoreados en G será distinta de la distribución de referencia H. Entonces, la carta de control r se puede construir con los pasos siguientes:

1. (Fase I). Obtenga los estimadores $\bar{\boldsymbol{Y}}$ y $\mathbf{S}$ de la muestra de referencia y calcule posteriormente las profundidades de Mahalanobis de cada $\boldsymbol{Y}_j, j = 1, \ldots, m$ aplicando la expresión en (7).
2. Ordene ascendentemente las profundidades de cada $\boldsymbol{Y}_j$, $j = 1, \ldots, m$ y denótelas como $Y_{[1]}, Y_{[2]}, \ldots, Y_{[m]}$.
3. Use nuevamente la expresión en (7) para hallar las profundidades de Mahalanobis de cada nueva observación de $\boldsymbol{X}_i$, $i = 1, 2, \ldots$, a monitorear respecto a la distribución de referencia, es decir, empleando los estimadores $\bar{\boldsymbol{Y}}$ y $\boldsymbol{S}$ de la muestra de referencia.
4. Calcule la estadística de clasificación por rangos para cada nueva observación de $\boldsymbol{X}_i$, $i = 1, 2, \ldots$, como sigue:

$$r(\boldsymbol{X}_i) = \frac{\#\{\boldsymbol{Y}_j \mid MD_{H_m}(\boldsymbol{Y}_j) \le MD_{H_m}(\boldsymbol{X}_i), j = 1, \ldots, m\}}{m+1} \tag{8}$$

5. Grafique cada uno de los valores obtenidos en (8) contra el tiempo y trace conjuntamente una línea central igual a 0.5 y un límite de control inferior (LCI) igual a la proporción de falsas alarmas α.

Si el valor de la estadística en (8) está por debajo del LCI, entonces se declara al proceso fuera de control. Liu y Singh (1993) demostraron que la distribución de la estadística en (8) es uniforme en el intervalo [0,1]. Este hecho se usa para establecer el límite de control a la proporción de falsas alarmas α y la línea central de la gráfica igual a 0.5.

Factor grado de autocorrelación

En este trabajo se evalúa la gráfica paramétrica y no paramétrica en casos donde no hay normalidad y las observaciones no son independientes. Para generar situaciones donde los datos no son independientes se simularon observaciones con autocorrelación conocida, suponiendo que el mecanismo generador de los datos seguía un modelo de vectores autorregresivos de orden uno (VAR(1)). A fin de limitar la complejidad del modelo, se consideró que $p = 2$, entonces, matricialmente, el modelo bivariado VAR(1) está dado por (Vanhatalo y Kulahci, 2015):

$$\boldsymbol{X}_t = \boldsymbol{c} + \boldsymbol{\Phi} X_{t-1} + \boldsymbol{u}_t \tag{9}$$

para $t = 1, 2, \ldots, T$, donde:

$$\boldsymbol{X}_t = \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix}, \quad \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \quad \boldsymbol{\Phi} = \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{pmatrix} \quad \text{y} \quad \boldsymbol{u}_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{10}$$

En (10), $\boldsymbol{X}_t$ es un vector de orden 2 de procesos autorregresivos univariados en el tiempo t, $\boldsymbol{c}$ es un vector de constantes de orden 2, $\boldsymbol{\Phi}$ es una matriz de tamaño 2×2 integrada por coeficientes de autocorrelación $\phi \in (-1,1)$ que miden la dependencia entre el valor actual y el rezagado de $\boldsymbol{X}_t$, y $\boldsymbol{u}_t$ es un proceso de ruido blanco vectorial con vector de medias $\mathbf{0}_2$ y matriz de varianzas y covarianzas Σ_u invariantes en el tiempo. $\mathbf{0}_2$ es un vector de orden 2 cuyas entradas son todas iguales a cero. Por lo anterior, las tres componentes que se necesitan para generar observaciones multivariadas con autocorrelación conocida mediante el modelo VAR(1) son el vector de medias del proceso $\boldsymbol{\mu}$, la matriz de coeficientes autorregresivos $\boldsymbol{\Phi}$ y la matriz de varianzas y covarianzas de los residuos Σ_u. Para simplificar el número de casos posibles de autocorrelación, en este trabajo se considera el caso bivariado con $\boldsymbol{\Phi}$ de la forma:

$$\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{pmatrix} = \begin{pmatrix} \varphi_{11} & 0 \\ 0 & \varphi_{22} \end{pmatrix} \tag{11}$$

donde los casos estudiados resultan de combinar los valores de φ_{11} y φ_{22} para cada una de las siguientes situaciones:

a) Sin autocorrelación.
b) Autocorrelación baja (0.33).
c) Autocorrelación media (0.63).
d) Autocorrelación alta (0.93).
e) Autocorrelación negativa (–0.50).

Factor distribución de los datos

Para controlar la distribución de las observaciones bivariadas en presencia de autocorrelación se induce la distribución deseada en $\boldsymbol{u}_t$. En este trabajo es de interés generar mediciones con distribución simétrica, simétrica con colas pesadas y asimétrica, por lo que se consideran las distribuciones: normal bivariada, t de Student bivariada y gamma bivariada. Estas dos últimas se utilizan tal y como las definen Stoumbos y Sullivan (2002), para mayores detalles véase en Kotz *et al*. (2004). Las distribuciones que se consideran para $\boldsymbol{u}_t$ son:

1. Normal bivariada.
2. t de Student bivariada con grados de libertad $\upsilon = 3$ y 20.
3. Gamma bivariada con parámetro de forma $\nu = 1$ y 16.

Factor relación entre variables

Para controlar la relación entre las variables, se usaron las entradas que están fuera de la diagonal principal de Σ_u. Se consideran dos casos:

1. No relacionadas $\Sigma_u = \boldsymbol{I}_2$.
2. Relacionadas $\boldsymbol{\Sigma}_u = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$.

Estado del proceso

Para simular al proceso en control, se estableció el vector de medias del modelo VAR(1) en $\boldsymbol{\mu} = 0_2$. Para simular al proceso fuera de control, se ejecutaron tres diferentes escenarios de cambio en el vector de medias, cuyas intensidades se midieron mediante el parámetro de no centralidad λ expresado como sigue (Lowry *et al*., 1992):

$$\lambda = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^T \Sigma_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)} \tag{12}$$

en el cual, $\boldsymbol{\mu}_0$ y Σ_0 son el vector de medias y la matriz de varianzas y covarianzas del proceso en control y $\boldsymbol{\mu}_1$ es el vector de medias del parámetro desplazado. Los valores de λ que se consideran para analizar la eficiencia de las cartas de control a cambios en el vector de medias son pequeños (0.50), medios (2) y grandes (3). Los cambios se añaden a las variables con la misma magnitud.

Todo el procedimiento de simulación se ejecutó en el paquete estadístico R (R Core Team, 2019). Los códigos en R se encuentran disponibles bajo pedido a los autores de este trabajo.

Discusión y análisis de resultados

Caso distribución normal bivariada

El comportamiento en control de las dos gráficas de monitoreo para cuando los datos tienen distribución normal con diferentes escenarios de autocorrelación y relación entre variables se muestra en la Tabla 1. Note que cuando las observaciones son independientes el ARL_0 está muy cerca de 370 para las dos cartas de control. De la Tabla 1 puede notarse que el efecto de la autocorrelación, cuando el proceso está en control, produce valores más grandes de ARL_0 en ambas cartas de control a medida que el grado de autocorrelación, positivo o negativo, es más alto. El aumento del ARL_0 no es perjudicial en una carta de control, al contrario, es benéfico y deseable siempre que no lesione su capacidad de detectar cambios.

Este aumento en el ARL_0 de la gráfica T^2 para datos autocorrelacionados se corroboró con el trabajo de Vanhatalo y Kulahci (2015). De la Tabla 1 se puede apreciar también que cuando las variables están altamente relacionadas, el ARL_0 en presencia de autocorrelación es mayor que cuando las características de interés no están relacionadas. El efecto de la relación entre las variables en el caso fuera de control fue semejante, por lo que el análisis del ARL_1 se muestra solo considerando el caso en el que las variables están altamente relacionadas. Los valores de ARL_1 para las tres magnitudes diferentes de cambio se exhibe en la Tabla 2.

Cuando el proceso está fuera de control es deseable que el ARL sea pequeño, lo que indica que la carta detecta rápidamente que el proceso está fuera de control.

Tabla 1. Valores de ARL_0 para observaciones normales bivariadas con diferentes grados de autocorrelación

Φ		$\Sigma_u = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$		$\Sigma_u = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$	
Φ_{11}	Φ_{22}	T^2	r	T^2	r
	0.00	371.76	373.86	370.45	371.91
	0.33	392.64	389.38	421.27	410.50
0.00	0.63	410.17	401.72	499.16	432.77
	0.93	456.18	437.80	565.21	500.69
	–0.50	377.55	373.35	432.30	426.21
	0.00	390.98	385.86	418.41	406.32
	0.33	397.06	387.62	437.42	427.65
0.33	0.63	415.64	406.32	503.93	483.40
	0.93	495.97	483.81	574.77	522.12
	–0.50	387.61	382.20	441.65	437.77
	0.00	406.47	402.04	496.27	438.85
	0.33	413.79	404.04	507.18	488.97
0.63	0.63	439.08	425.16	569.88	547.31
	0.93	542.80	531.39	650.02	635.15
	–0.50	410.72	392.72	469.98	445.36
	0.00	457.34	440.20	568.38	509.30
	0.33	497.84	486.47	578.78	527.13
0.93	0.63	539.54	532.01	655.45	645.10
	0.93	656.01	646.66	703.57	698.40
	–0.50	494.12	477.61	561.07	543.20
	0.00	380.24	375.39	435.82	429.03
	0.33	385.02	383.76	443.31	436.55
–0.50	0.63	415.60	393.71	473.54	447.49
	0.93	495.58	475.69	563.92	554.76
	–0.50	430.00	399.20	422.25	402.45

En la Tabla 2 se observa que, para observaciones no autocorrelacionadas, la carta T^2 detecta más rápido que la carta no paramétrica las señales de fuera de control ante cambios pequeños y medios en el vector de medias. Para cambios grandes, la capacidad de detección de cambios en las dos gráficas es muy similar. Esta información es congruente con los resultados del estudio de Zertuche y Cantú (2008), quienes encontraron que el desempeño de la carta T^2 es más eficiente que el de la gráfica de clasificación por rangos, a menos que el cambio en la media sea grande, ya que en ese caso el desempeño de las dos cartas de control es similar. Se observó que el efecto de la autocorrelación en los datos provoca que incremente el ARL_1 de ambas cartas conforme crece la autocorrelación, esto genera una sustancial disminución en la rapidez con la que las cartas identifican señales de fuera de control. Esto también es congruente con el resultado de Vanhatalo y Kulahci (2015), quienes reportan que la autocorrelación empobrece la capacidad para detectar cambios en la carta T^2. No obstante, es importante hacer notar que cuando en una de las dos variables los datos presentan autocorrelación negativa y en la otra la autocorrelación es nula, baja y regular, el $ARL1$ es más pequeño, por lo que estos casos en particular no alteran el desempeño de ambas cartas de control. Además, en cambios grandes, se aprecia que la autocorrelación, a menos que sea muy alta, no daña severamente la capacidad de detectar puntos fuera de control en ambas cartas de control.

Tabla 2. Valores de ARL_1 con observaciones normales bivariadas, diferentes grados de autocorrelación y magnitudes de cambio considerando $Cov(u_{1t}, u_{2s}) = 0.9$

Φ		$\lambda = 0.50$		$\lambda = 2$		$\lambda = 3$	
ϕ_{11}	ϕ_{22}	T^2	r	T^2	r	T^2	r
0.00	0.00	215.04	219.69	10.01	11.47	3.18	3.35
	0.33	228.48	230.56	10.16	11.51	3.23	3.65
	0.63	240.02	245.26	10.68	11.85	3.39	3.98
	0.93	254.12	257.03	11.30	12.69	3.59	4.37
	–0.50	173.00	175.88	7.70	8.78	2.44	3.02
0.33	0.00	227.21	228.05	9.75	10.94	3.10	3.67
	0.33	236.65	238.58	10.53	11.91	3.34	4.10
	0.63	252.58	254.54	11.24	12.71	3.57	4.38
	0.93	263.05	266.71	11.70	13.32	3.72	4.59
	–0.50	181.91	184.66	8.09	8.92	2.57	3.07
0.63	0.00	236.63	238.48	10.53	11.91	3.34	4.01
	0.33	254.11	259.53	11.30	12.76	3.59	4.39
	0.63	257.73	264.29	11.46	13.20	3.64	4.55
	0.93	265.68	268.54	11.82	13.41	3.75	4.62
	–0.50	190.98	193.49	8.49	9.51	2.70	3.28
0.93	0.00	253.18	255.49	11.26	12.76	3.58	4.39
	0.33	262.63	267.16	11.68	13.34	3.71	4.59
	0.63	267.99	270.05	11.92	13.64	3.79	4.70
	0.93	334.34	337.14	19.87	16.09	10.72	7.54
	–0.50	235.80	227.51	10.49	11.36	3.33	3.91
–0.50	0.00	175.69	176.13	7.81	8.80	2.48	3.03
	0.33	183.36	185.13	8.16	9.24	2.59	3.18
	0.63	193.76	196.25	8.93	10.05	2.84	3.46
	0.93	245.12	250.57	10.90	12.51	3.46	4.31
	–0.50	352.20	356.00	15.67	17.43	4.98	6.00

Caso distribución *t* de Student bivariada

En la Tabla 3 se presentan las cantidades de ARL_0 de las dos cartas de control para datos bivariados distribuidos *t* con diferentes valores de grados de libertad, niveles de autocorrelación y relación entre variables. Note que, bajo independencia, el ARL_0 de la carta paramétrica para observaciones *t* ($\upsilon = 3$) disminuye inaceptablemente respecto al nominal bajo el supuesto de normalidad, a diferencia de la carta *r*, que exhibe un ARL_0 considerablemente cercano del nominal para $\upsilon = 3$ y $\upsilon = 20$ Así que, para cuando el proceso está en control, colas pesadas en datos independientes provoca una tasa desmesurada de falsas alarmas en la gráfica T^2, acorde con los resultados de Chou *et al.* (2001). La autocorrelación en este caso no produce ningún efecto significativo para la carta T^2, mientras que en la gráfica basada en la profundidad de Mahalanobis, la autocorrelación produce un desempeño bastante significativo, ya que su ARL_0 aumenta tanto como el grado de autocorrelación. Nuevamente, la relación alta entre variables no parece causar mayores incidencias en el comportamiento del ARL_0. Por otra parte, cuando el proceso está fuera de control, véase la Tabla 4, hay evidencia de que la carta *r* es superior que la T^2 para datos independientes en las tres magnitudes de cambio. Al inducir autocorrelación en las mediciones, se observa que la gráfica no paramétrica continúa siendo mejor que la T^2. Note que en cambios medios y grandes con combinaciones de autocorrelación baja y regular, incluyendo el caso de autocorrelación negativa, el desempeño de la carta *r* no se aleja demasiado de su desempeño nominal para datos

Tabla 3. Valores de ARL_0 para observaciones *t* bivariadas con diferentes grados de libertad y grados de autocorrelación

		$\upsilon = 3$				$\upsilon = 20$			
Φ		$\Sigma_u = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$		$\Sigma_u = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$		$\Sigma_u = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$		$\Sigma_u = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$	
ϕ_{11}	ϕ_{22}	T^2	r	T^2	r	T^2	r	T^2	r
	0.00	79.81	366.23	85.63	368.15	102.81	368.43	105.13	367.79
	0.33	84.29	381.43	97.38	406.34	108.58	383.73	119.55	405.95
0.00	0.63	88.05	393.52	115.38	428.39	113.43	395.88	141.66	427.97
	0.93	97.93	428.86	130.65	495.62	126.16	431.44	160.40	495.13
	−0.50	81.05	365.73	99.93	421.90	104.41	367.92	122.68	421.48
	0.00	83.94	377.99	96.72	402.21	108.13	380.26	118.74	401.81
	0.33	85.24	379.71	101.11	423.32	109.81	381.99	124.13	422.91
0.33	0.63	89.23	398.03	116.48	478.50	114.95	400.42	143.01	478.04
	0.93	106.48	473.93	132.86	516.84	137.16	476.78	163.11	516.33
	−0.50	83.21	374.40	102.09	433.34	107.19	376.65	125.33	432.92
	0.00	87.26	393.84	114.71	434.41	112.41	396.20	140.84	433.98
	0.33	88.83	395.79	117.23	484.02	114.43	398.17	143.93	483.55
0.63	0.63	94.26	416.48	131.73	541.77	121.43	418.98	161.73	541.24
	0.93	116.53	520.55	150.25	628.72	150.11	523.67	184.47	628.11
	−0.50	88.17	384.71	108.64	440.86	113.58	387.02	133.38	440.42
	0.00	98.18	431.21	131.38	504.14	126.48	433.80	161.30	503.65
	0.33	106.88	476.54	133.78	521.80	137.68	479.40	164.25	521.29
0.93	0.63	115.83	521.15	151.51	638.57	149.21	524.29	186.01	637.95
	0.93	140.83	633.46	162.63	691.33	181.42	637.27	199.67	690.66
	−0.50	106.08	467.86	129.69	537.70	136.65	470.67	159.23	537.17
	0.00	81.63	367.73	100.74	424.69	105.15	369.93	123.68	424.27
	0.33	82.66	375.93	102.47	432.14	106.48	378.19	125.81	431.71
−0.50	0.63	89.22	385.68	109.46	442.96	114.93	387.99	134.39	442.53
	0.93	106.39	465.98	130.35	549.14	137.05	468.77	160.04	548.61
	−0.50	92.31	391.06	97.60	398.38	118.92	393.40	119.83	397.99

no autocorrelacionados. Los valores de ARL_1 de la Tabla 4 se muestran solo para $\upsilon = 3$, ya que con $\upsilon = 20$, el comportamiento del ARL_1 es semejante al de $\upsilon = 3$.

Tabla 4. Valores de ARL_1 con observaciones *t* bivariadas, diferentes grados de autocorrelación y magnitudes de cambio considerando $Cov\ (u_{1t}, u_{2s})$ y 0.9 y $\upsilon = 3$

Φ		$\lambda = 0.50$		$\lambda = 2$		$\lambda = 3$	
ϕ_{11}	ϕ_{22}	T^2	r	T^2	r	T^2	r
	0.00	70.34	54.45	20.16	10.18	4.49	2.17
	0.33	79.99	60.10	22.93	11.24	5.11	2.40
0.00	0.63	94.78	63.36	27.16	11.85	6.05	2.53
	0.93	107.32	73.30	30.76	13.70	6.85	2.92
	–0.50	82.08	62.40	23.53	11.67	5.24	2.49
	0.00	79.45	59.49	22.77	11.12	5.07	2.37
	0.33	83.06	62.61	23.80	11.71	5.30	2.50
0.33	0.63	95.68	70.77	27.42	13.23	6.11	2.82
	0.93	109.14	76.44	31.28	14.29	6.97	3.05
	–0.50	83.86	64.09	24.03	11.98	5.35	2.55
	0.00	94.23	64.25	27.01	12.01	6.01	2.56
	0.33	96.30	71.59	27.60	13.38	6.15	2.85
0.63	0.63	108.21	80.13	31.01	14.98	6.91	3.19
	0.93	123.42	92.99	35.37	17.39	7.88	3.71
	–0.50	89.24	65.20	25.58	12.19	5.70	2.60
	0.00	107.92	74.56	30.93	13.94	6.89	2.97
	0.33	109.90	77.17	31.50	14.43	7.01	3.08
0.93	0.63	124.45	94.45	35.67	17.66	7.94	3.76
	0.93	133.59	102.25	38.29	19.12	8.53	4.07
	–0.50	106.53	79.53	30.53	14.87	6.80	3.17
	0.00	82.75	62.81	23.72	11.74	5.28	2.50
	0.33	84.18	63.91	24.13	11.95	5.37	2.55
–0.50	0.63	89.91	65.51	25.77	12.25	5.74	2.61
	0.93	107.08	81.22	30.69	15.18	6.83	3.24
	–0.50	80.18	58.92	22.98	11.02	5.12	2.35

Caso distribución gamma bivariada

La Tabla 5 exhibe el ARL_0 de las dos cartas de control para datos bivariados con distribución gamma, diferentes valores del parámetro de forma, grados de autocorrelación y relación entre variables, donde se observa, suponiendo independencia, que la carta paramétrica no compite para datos que se distribuyen poco o muy asimétricamente, ya que su ARL_0 para observaciones gamma ($\nu = 1$) y gamma ($\nu = 16$) baja excesivamente. Este resultado se puede constatar con los estudios de Stoumbos y Sullivan (2002) quienes observan que la gráfica T^2 exhibe una cantidad excesiva de falsas alarmas para distribuciones muy asimétricas. El ARL_0 de la carta *r* decae respecto al nominal, pero no radicalmente. Así como en el caso de la distribución *t*, se nota que la autocorrelación produce en la gráfica T^2 un ARL_0 significativamente pobre, pero altamente significativo en la gráfica no paramétrica. En cuanto al ARL_1, véase la Tabla 6, se aprecia nuevamente que la carta *r* tiene un mejor desempeño que la T^2 para datos independientes en las tres magnitudes de cambio. Por otro lado, la autocorrelación altera en mayor medida la capacidad de detección de cambios en la gráfica T^2 que en la gráfica *r*. Observe que el ARL_1 de la carta *r* no se aparta considerablemente de su desempeño nominal cuando los datos presentan autocorrelación baja y regular, excepto en cambios pequeños. Los valores de ARL_1 de la Tabla 6 se muestran solo para $\nu = 1$, ya que con $\nu = 16$ el comportamiento del ARL_1 es similar que al de $\nu = 1$.

https://doi.org/10.22201/fi.25940732e.2020.21.3.023

Tabla 5. Valores de ARL_0 para observaciones gama bivariadas con diferentes parámetros de forma y grados de autocorrelación

Φ		$\nu = 1$				$\nu = 16$			
		$\Sigma_u = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$		$\Sigma_u = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$		$\Sigma_u = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$		$\Sigma_u = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$	
ϕ_{11}	ϕ_{22}	T^2	r	T^2	r	T^2	r	T^2	r
	0.00	55.07	359.01	55.98	361.65	198.07	364.01	200.98	365.65
	0.33	58.16	373.91	63.66	399.17	209.19	379.12	228.55	403.58
0.00	0.63	60.76	385.76	75.43	420.83	218.53	391.13	270.81	425.48
	0.93	67.58	420.41	85.41	486.87	243.05	426.26	306.64	492.25
	–0.50	55.93	358.52	65.33	414.45	201.16	363.51	234.53	419.03
	0.00	57.92	370.54	63.23	395.11	208.31	375.70	227.00	399.48
	0.33	58.82	372.22	66.10	415.85	211.55	377.41	237.31	420.45
0.33	0.63	61.57	390.18	76.15	470.06	221.45	395.61	273.39	475.26
	0.93	73.47	464.59	86.86	507.71	264.25	471.06	311.83	513.33
	–0.50	57.42	367.02	66.74	425.69	206.51	372.13	239.61	430.40
	0.00	60.21	386.07	74.99	426.74	216.56	391.45	269.24	431.46
	0.33	61.30	387.99	76.64	475.48	220.46	393.40	275.16	480.73
0.63	0.63	65.04	408.27	86.12	532.20	233.94	413.96	309.18	538.09
	0.93	80.41	510.28	98.23	617.62	289.20	517.39	352.65	624.45
	–0.50	60.84	377.12	71.02	433.07	218.83	382.37	254.98	437.86
	0.00	67.75	422.71	85.89	495.24	243.67	428.60	308.36	500.72
	0.33	73.75	467.14	87.46	512.59	265.24	473.65	314.00	518.26
0.93	0.63	79.92	510.88	99.05	627.30	287.46	518.00	355.60	634.24
	0.93	97.18	620.97	106.32	679.13	349.52	629.62	381.71	686.64
	–0.50	73.20	458.64	84.79	528.20	263.26	465.03	304.40	534.05
	0.00	56.33	360.48	65.86	417.19	202.59	365.50	236.44	421.80
	0.33	57.03	368.52	66.99	424.51	205.13	373.65	240.51	429.20
–0.50	0.63	61.56	378.07	71.56	435.14	221.43	383.34	256.91	439.95
	0.93	73.41	456.79	85.22	539.45	264.04	463.15	305.94	545.41
	–0.50	63.70	383.35	63.81	391.34	229.10	388.69	229.08	395.67

Tabla 6. Valores de ARL_1 con observaciones gamma bivariadas, diferentes grados de autocorrelación y magnitudes de cambio considerando $Cov(u_{1t}, u_{2s}) = 0.9$ y $\nu = 1$

Φ		$\lambda = 0.50$		$\lambda = 2$		$\lambda = 3$	
ϕ_{11}	ϕ_{22}	T^2	r	T^2	r	T^2	r
	0.00	57.18	50.97	27.19	6.77	4.25	2.77
	0.33	65.02	56.26	30.92	7.47	4.83	3.06
0.00	0.63	77.05	59.31	36.64	7.88	5.73	3.22
	0.93	87.24	68.62	41.48	9.11	6.48	3.73
	–0.50	66.73	58.41	31.73	7.76	4.96	3.17
	0.00	64.58	55.69	30.71	7.40	4.80	3.03
	0.33	67.52	58.61	32.11	7.78	5.02	3.19
0.33	0.63	77.78	66.25	36.99	8.80	5.78	3.60
	0.93	88.72	71.56	42.19	9.50	6.59	3.89
	–0.50	68.17	60.00	32.42	7.97	5.07	3.26
	0.00	76.60	60.14	36.42	7.99	5.69	3.27
	0.33	78.28	67.01	37.23	8.90	5.82	3.64
0.63	0.63	87.96	75.01	41.83	9.96	6.54	4.08
	0.93	100.33	87.05	47.71	11.56	7.46	4.73
	–0.50	72.54	61.04	34.50	8.11	5.39	3.32
	0.00	87.73	69.80	41.72	9.27	6.52	3.79
	0.33	89.34	72.24	42.48	9.60	6.64	3.93
0.93	0.63	101.17	88.41	48.11	11.74	7.52	4.80
	0.93	108.60	95.71	51.64	12.71	8.07	5.20
	–0.50	86.60	74.44	41.18	9.89	6.44	4.05
	0.00	67.27	58.80	31.99	7.81	5.00	3.20
	0.33	68.43	59.83	32.54	7.95	5.09	3.25
–0.50	0.63	73.09	61.33	34.76	8.15	5.43	3.33
	0.93	87.04	76.03	41.39	10.10	6.47	4.13
	–0.50	65.18	55.16	30.99	7.33	4.84	3.00

Conclusiones

De acuerdo con los resultados obtenidos en este trabajo, se puede concluir que la carta de control T^2 de Hotelling es únicamente recomendable en el monitoreo de procesos donde se sostenga que las mediciones de las características de calidad tienen distribución normal multivariada y sean independientes. Cuando los datos son normales, pero están autocorrelacionados, la gráfica T^2 pierde severamente su capacidad deseable de detección de cambios en la media del proceso. Si las observaciones, además de presentar autocorrelación, se distribuyen asimétricamente o con colas muy pesadas, la carta T^2 incrementa la proporción de falsas alarmas en procesos que están en control estadístico, situación que resulta desfavorable. En contraste, la carta de control basada en la profundidad de Mahalanobis, no requiere normalidad en los datos a monitorear. Mostró ser eficiente en procesos distribuidos gamma y *t*-student bivariadas con observaciones independientes y robustas en la mayoría de los casos estudiados con autocorrelación baja y media. Una ventaja más que mostró la carta bajo el concepto de medidas de profundidad es que la autocorrelación baja, media y alta no afecta en ningún momento la proporción de falsas alarmas y tampoco agrava considerablemente su rapidez para detectar cambios grandes, sin embargo, en cambios pequeños, dicha rapidez disminuye. ¿Cómo mejorar esta situación? es una importante pregunta de investigación a seguir.

Referencias

Boone, J. (2010). *Contributions to multivariate control charting, studies of the Z chart and four nonparametrics chart. Tuscaloosa, Alabama*. (PhD Thesis. University of Alabama).

Chou, Y. M., Mason, R. L. & Young, J. C. (2001). The control chart for individual observations from a multivariate non-normal distribution. *Communications in Statistics: Theory and Methods,* 30(8), 291-303. https://doi.org/10.1081/STA-100105706

Hamurkaroglu, C., Mert, M. & Saykan, Y. (2004). Nonparametric control charts based on Mahalanobis depth. *Hacettepe Journal of Mathematics and Statistics,* 33, 57-67.

Kotz, S., Balakrishnan, N. & Johnson, N. L. (2004). *Continuous multivariate distributions, models and applications*. New Jersey: John Wiley & Sons.

Liu, R. (1995). Control charts for multivariate processes. *Journal of The American Statistical Association*, 90, 1380-1387.

Liu, R. (1990). On a notion of data depth based on random simplicities. *The Annals of Statistics*, 18, 405-414.

Liu, R. & Singh, K. (1993). A quality index based on data depth and multivariate Rank tests. *Journal of The American Statistical Association*, 88, 252-260. https://doi.org/10.1080/01621459.1993.10594317

Lowry, C. A., Woodal, W. H., Champ, C. W. & Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, 34(1), 46-53.

Mason, R. L. & Young, J. C. (2002). *Multivariate statistical process control with industrial applications*. Philadelphia: SIAM.

Messaoud, A., Theis, W., Weihs, C. & Hering, F. (2005). Application and use of multivariate control charts in a BTA deep hole drilling process. En *Classification the Ubiquitous Challenge*, 648-655. 10.1007/3-540-28084-7_77

Montgomery, D. C. (2009). *Introduction to statistical quality control* (Sixth edition). Arizona: John Wiley & Sons.

R Core Team (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Stoumbos, Z. G. & Sullivan, J. H. (2002). Robustness to non-normality of the multivariate EWMA control chart. *Journal of Quality Technology*, 34(3), 260-276. https://doi.org/10.1080/00224065.2002.11980157

Vanhatalo, E. & Kulahci, M. (2015). The effect of autocorrelation on the Hotelling T^2 control chart. *Quality and Reliability Engineering International,* 31(8), 1779-1796.

Zertuche-Luis, F. & Cantú-Sifuentes, M. (2008). Una comparación del desempeño de las cartas de control T^2 de Hotelling y de clasificación por rangos. *Ingeniería, Investigación y Tecnología*, 9(3), 185-195. https://doi.org/10.22201/fi.25940732e2008.09n3.16