

ANÁLISIS COMPARATIVO DE LOS GRÁFICOS DE CONTROL MULTIVARIADOS GENERADOS EN MINITAB Y R PARA EL MÉTODO T^2 DE HOTELLING: UNA EXPLORACIÓN DE LA ESTADÍSTICA MULTIVARIADA

A COMPARATIVE ANALYSIS OF MULTIVARIATE CONTROL CHARTS RUN IN MINITAB & R FOR THE T^2 HOTELLING TECHNIQUE: AN EXPLORATORY APPROACH TO MULTIVARIATE STATISTICS

José Othón Perez Villanueva

Tecnológico Nacional de México / IT de Celaya, México
m2203051@itcelaya.edu.mx

Manuel Darío Hernández Ripalda

Tecnológico Nacional de México / IT de Celaya, México
dario.hernandez@itcelaya.edu.mx

Recepción: 13/septiembre/2023

Aceptación: 17/junio/2024

Resumen

En este estudio, se compararon los gráficos de control multivariados generados en Minitab®, R y Excel, para el método T^2 de Hotelling para datos individuales y de subgrupos. Se evaluaron las capacidades de detección y precisión de cada herramienta, se observó que Minitab® y R produjeron resultados comparables en general.

Este estudio proporciona información relevante para seleccionar el método de gráfico de control multivariado más apropiado al monitorear un proceso industrial con variables altamente correlacionadas y elegir la herramienta de software adecuada para su análisis, especialmente en el caso de T^2 de Hotelling.

Los resultados indicaron que tanto Minitab® como R son opciones sólidas para generar gráficos de control multivariados, ya que mostraron una capacidad comparable para detectar desviaciones multivariadas. Sin embargo, es importante tener en cuenta la discrepancia encontrada para el gráfico T^2 de Hotelling entre R y Minitab® debido a las diferencias mostradas en los grados de libertad.

Palabras clave: Gráficos de control multivariados, T^2 de Hotelling, Minitab, R, Excel, Estadística multivariada.

Abstract

In this paper, multivariate control charts generated in Minitab, R and Excel were benchmarked for Hotelling's T^2 method for individual and subgroup data. The detection capabilities and accuracy of each tool were evaluated, and it was observed that Minitab and R produced generally comparable results.

This study provides relevant information for selecting the most appropriate multivariate control charting method when monitoring an industrial process with highly correlated variables and choosing the appropriate software tool for analysis, especially in the case of Hotelling's T^2 .

The results indicated that both Minitab and R, are robust options for generating multivariate control charts, as they showed comparable ability to detect multivariate deviations. However, it is important to note the discrepancy found for the Hotelling T^2 plot between R and Minitab due to the differences shown in the degrees of freedom.

Keywords: *Multivariate control charts, Hotelling's T^2 , Minitab, R, Excel, Multivariate statistics*

1. Introducción

Los gráficos de control multivariados se basan en el análisis de datos multivariantes, donde se recopilan mediciones simultáneas de múltiples variables en diferentes puntos en el tiempo o en diferentes unidades de producción. Estas variables pueden estar altamente correlacionadas, lo que significa que los cambios en una variable pueden estar asociados con cambios en otras variables. El análisis multivariante permite tener en cuenta estas relaciones y proporciona una visión más completa del proceso en comparación con los gráficos de control univariados.

Para [Campos Avendaño, 2020], en el ámbito del control estadístico de procesos, los gráficos de control son herramientas fundamentales para evaluar la estabilidad de un proceso productivo. Estas herramientas permiten supervisar el

comportamiento de las variables de calidad del sistema. Los gráficos de control multivariantes son especialmente ventajosos, ya que toman en cuenta la correlación entre las distintas variables de calidad.

Según [Barahona y González-Narváez, 2019], aunque los gráficos multivariantes son una opción superior para el control de procesos, los costos asociados al muestreo a menudo son elevados. Por esta razón, se busca desarrollar gráficos de control multivariantes que utilicen variables costosas solo cuando sea necesario, lo cual ayudaría a reducir los costos de muestreo.

Como bien apuntan los expertos [Ruelas, et al. 2020] la supervisión de la calidad en los sistemas de producción es una tarea muy importante. Para lograr este objetivo, es necesario supervisar la calidad de los productos desde un enfoque global que considere todas las características de calidad. Además, indican que una desventaja significativa del control estadístico multivariante es la falta de un análisis complementario que interprete la presencia de variaciones o señales fuera de control, lo que dificulta identificar la fuente de inestabilidad en el proceso de fabricación. Además, los expertos mencionan como existen numerosos gráficos de control y varios programas estadísticos que implementan funciones para crear gráficos de control, aunque estos son mayoritariamente univariantes. Desafortunadamente, no hay programas estadísticos que generen gráficos de control multivariantes, lo que facilitaría su aplicación y el control en los procesos.

Los procesos de fabricación actuales son cada vez más complejos y los productos tienen cada vez más funciones, por lo que es necesario controlar más de un atributo de calidad.

Uno de los gráficos de control multivariante más útiles para estos procesos es el gráfico de Hotelling T^2 , que se basa en la teoría de la distribución normal multivariante.

Como menciona [Navarro, 2003] en el sector industrial es necesario crear un gráfico de control multivariante que permita el análisis simultáneo de varias características de calidad, teniendo en cuenta el efecto de cada característica sobre la calidad final del producto refinado y el efecto de la interacción de estas características sobre la calidad final del producto.

Además [Montgomery, 2009], sugiere y evidencia como el control independiente de dos características de calidad para un mismo producto puede ser muy engañoso. “El uso de dos gráficos univariantes independientes ha distorsionado el control simultáneo de la respuesta. Provocando que caigamos dentro de un error Tipo I o Tipo II. Si las variables a estudiar son independientes entre sí, los gráficos univariantes son una posibilidad y los límites de control pueden ajustarse para tener en cuenta este hecho.” Esta distorsión en el procedimiento de control del proceso aumenta a medida que aumenta el número de características de calidad.

Así de esta manera, se genera un supuesto de independencia entre estas variables, si lo son, gráficos univariados son la solución, sin embargo [Montgomery, Jennings, & Kulahci, 2008] sugieren como los métodos tradicionales de SPC (Statistical Process Control) suponen que los datos son independientes en el tiempo. Sin embargo, debido a la dinámica del sistema las observaciones en serie de tiempo suelen estar siempre correlacionadas dichas variables.

Como bien apuntan los autores [Vanhatalo & Kulahci, 2015] sugieren que, en muchas ocasiones, el control estadístico del proceso mediante gráficos de control univariados no es posible, o más bien inadecuado, sería un engaño decir que el proceso está verdaderamente en control estadístico cuando no lo sabemos ya que las variables están correlacionadas.

Claramente, los sesgos al introducir más variables de control dentro de los gráficos de forma conjunta pueden ser graves incluso con valores de p (variables de control/atributos de calidad) modestos (rangos de 2-4), y si los atributos de calidad no son independientes y ni siquiera tenemos certeza del control estadístico del proceso.

El trabajo inicial sobre el control de calidad multivariado fue realizado por [Hotelling, 1947] quien aplicó su procedimiento a los datos de bombardeo durante la Segunda Guerra Mundial. Gracias a este trabajo, este tema es particularmente importante hoy en día, ya que los procedimientos de inspección automatizados hacen que sea relativamente fácil medir muchos parámetros o características de calidad para cada unidad de un producto fabricado en una serie de tiempo.

Base teórica del T^2 de Hotelling

Para el desarrollo del presente artículo, se revisarán fundamentos teóricos, así como supuestos o criterios a tomar en cuenta para la construcción y cálculo de un gráfico T^2 de Hotelling.

Distribución Normal Multivariada

La Distribución normal multivariada, es importante para la construcción de nuestro gráfico Hotelling. El estadístico T^2 de Hotelling es una variable aleatoria univariante, y como menciona [Navarro, 2003] se crea considerando información sobre la dispersión y la media.

En el control de calidad estadístico univariante, generalmente se utiliza la distribución normal univariada para describir el comportamiento de una característica de calidad en una serie de tiempo, por lo que podemos usar la distribución normal bivariada para monitorizar 2 características de calidad.

El método, creado por Harold Hotelling en su artículo "The Generalization of Student's Ratio" publicado en 1931 supone que las características de calidad de interés se distribuyen según una distribución normal multivariante. Además, para [Kariya, 1981] la distribución normal multivariante es una extensión de la distribución normal univariante a una situación con múltiples (p) variables o características de calidad. La ecuación 1 muestra la función de densidad normal multivariante. Un gráfico de control T^2 con dos características de calidad, $p = 2$, usa la distribución normal bivariada.

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)' \Sigma^{-1}(x-\mu)} \quad (1)$$

Donde:

$$-\infty < x_j < \infty \quad j = 1, 2, 3, \dots, p$$

Vector de medias y Matriz de Covarianza

Cuando obtenemos una muestra aleatoria de una distribución normal p -dimensional $N_p(\mu, \Sigma)$ tendremos un vector $x = [x_1, x_2, x_3, \dots, x_n]$, este es un vector

p-dimensional, donde su media es vector $p \times 1$. La ecuación 2 muestra la expresión para el vector de medias.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

Además, Σ es la matriz de varianza-covarianza poblacional $p \times p$, [Montgomery, 2006] expone la matriz de covarianza, de una muestra (S), Las ecuaciones 3 y 4 respectivamente reflejan estas expresiones matemáticas:

$$S = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})' \quad (3)$$

$$S = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22}^2 & & \vdots \\ \vdots & & & \\ \sigma_{1p} & \sigma_{1p} & \dots & \sigma_{pp}^2 \end{bmatrix} \quad (4)$$

Donde: σ_{ii}^2 es la varianza de la i –ésima variable y σ_{ij} es la covarianza entre la i –ésima y la j –ésima variable. Podemos decir entonces, la matriz de covarianza-varianza desempeña un papel importante. Para [Katebi & Moghadam, 2022] esta matriz captura las relaciones de covarianza y las varianzas entre las variables en un conjunto de datos multivariados. Proporciona información sobre cómo las variables están relacionadas entre sí y cómo se distribuyen en conjunto. La matriz de covarianza-varianza se representa como una matriz cuadrada simétrica.

Como se muestra en la ecuación 4, en su diagonal principal, se encuentran las varianzas de cada variable individual, mientras que los elementos fuera de la diagonal principal representan las covarianzas entre pares de variables.

Las ecuaciones 5 y 6, surgen de la obtención de la diagonal en la matriz covarianza, esto es las varianzas muestrales en la diagonal principal de la matriz S . y, las covarianzas.

$$S_{diagonal} = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \quad (5)$$

$$S_{jp} = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ip} - \bar{x}_p)' \quad (6)$$

Estadístico T^2 para datos Individuales

El gráfico T^2 de Hotelling se construye utilizando la matriz de covarianza-varianza para calcular un estadístico T^2 que permite comparar grupos o muestras en términos de sus medias multivariadas. Este estadístico se basa en la inversa de la matriz de covarianza-varianza y se utiliza para determinar si existen diferencias significativas entre las medias de los grupos en el espacio multivariado.

El gráfico de T^2 de Hotelling para datos individuales ocurre con frecuencia en la industria de procesos de manufactura, donde existen múltiples características de calidad que requieren supervisión.

Por lo tanto, resulta apropiado utilizar gráficos de control multivariante con tamaño de muestra $n = 1$. La ecuación 7 corresponde al estadístico T^2 .

$$T^2 = (x - \bar{x})' S^{-1} (x - \bar{x}) \quad (7)$$

Los límites de control estadístico del proceso multivariante están dados por las ecuaciones 8 y 9.

$$UCL = \frac{p(m+1)(m-1)}{m^2 - mp} F_{\alpha, p, m-p} \quad (8)$$

$$LCL = 0 \quad (9)$$

Estadístico T^2 para subgrupos y $p = 2$

Cuando tenemos dos características de calidad [Montgomery, 2006] explica que son por ejemplo A_1 y A_2 , y que se distribuyen conjuntamente utilizando una distribución normal bivariada. Los valores promedio de estas características son μ_1 y μ_2 , y las desviaciones estándar de A_1 y A_2 son σ_1 y σ_2 , respectivamente. La covarianza entre A_1 y A_2 se representa como σ_{12} .

Suponiendo que conocemos los valores de σ_1 , σ_2 y σ_{12} , obtenemos las medias muestrales y de estas dos características de calidad a partir de una muestra de tamaño n , el estadístico resultante sigue una distribución chi-cuadrada con 2 grados de libertad. Por lo tanto, esta ecuación puede servir como base para crear un gráfico de control del proceso, específicamente para monitorear las medias μ_1 y μ_2 del proceso. La ecuación 10 permite el cálculo del estadístico Chi-Cuadrada.

$$X_0^2 = \frac{n}{\sigma_1^2 \sigma_2^2 - (\sigma_{12})^2} [\sigma_2^2 (\bar{x}_1 - \mu_1)^2 + \sigma_1^2 (\bar{x}_2 - \mu_2)^2 - 2\sigma_{12} (\bar{x}_1 - \mu_1)(\bar{x}_2 - \mu_2)] \quad (10)$$

Con las ecuaciones 11 y 12 se calculan los límites de control estadístico del proceso multivariante.

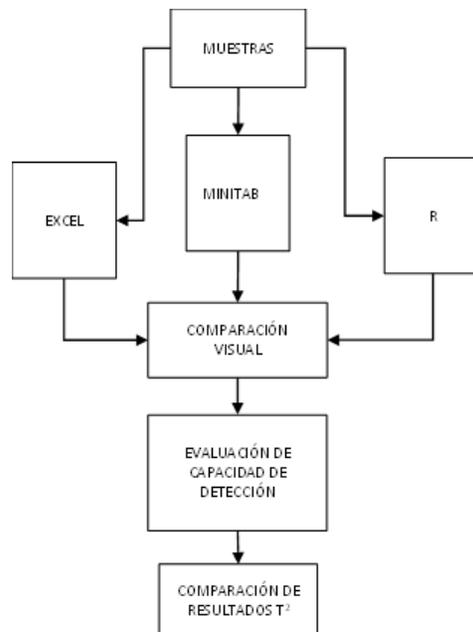
$$UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1} \quad (11)$$

$$LCL = 0 \quad (12)$$

2. Método

Se presenta el método de análisis superficial de los gráficos de control multivariados generados en Minitab®, R y Excel, así como el caso de estudio que se usara. La figura 1 contiene el método propuesto para el analisis comparativo entre los tres tipos de gráficos de control multivariado T^2 de Hotelling. El método consiste en los siguientes pasos:

- **Selección de datos:** Elegir un conjunto de datos multivariados con variables altamente correlacionadas que representen un proceso industrial o una situación relevante para el estudio para tomar una muestra.



Fuente: elaboración propia

Figura 1 Método de comparación entre los tres métodos de generación del gráfico.

- *Generación de gráficos de control multivariados en Minitab®*: Utilizar las funciones y herramientas disponibles en Minitab® para generar los gráficos de control multivariados, centrándose en el método T^2 de Hotelling para datos individuales y de subgrupos.
- *Generación de gráficos de control multivariados en Excel*: Utilizar las funciones y herramientas de Excel como hoja de cálculo, y el uso de las ecuaciones, para generar los gráficos de control multivariados siguiendo el método clásico establecido, incluyendo el método T^2 de Hotelling para datos individuales y de subgrupos.
- *Generación de gráficos de control multivariados en R*: Utilizar paquetes y funciones específicas de R para generar los gráficos de control multivariados, incluyendo el método T^2 de Hotelling para datos individuales y de subgrupos.
- *Comparación visual de los gráficos*: Observar y comparar visualmente los gráficos de control generados en cada software. Identificar cualquier diferencia notable en la presentación, diseño o representación de los datos.
- *Evaluación de la capacidad de detección*: Observar si los gráficos de control generados en cada software indican desviaciones multivariadas de manera clara y efectiva. Evaluar si los límites de control establecidos son consistentes y permiten identificar de manera adecuada patrones anormales en los datos.
- *Comparación de resultados*: Comparar los resultados obtenidos de los gráficos de control generados en cada software. Identificar cualquier discrepancia significativa en la detección de desviaciones multivariadas utilizando el valor de cada punto T^2 de Hotelling.

Caso de estudio: Manufactura de componente de automóviles

Se obtuvieron datos en serie de tiempo y altamente correlacionados de una línea de producción de una empresa, en la región laja-bajío, que fabrica componentes para a la industria automotriz, en este caso, mangueras hidráulicas para automóviles. Se obtuvieron 20 observaciones, con las características de calidad x_1 , x_2 , y x_3 . Siendo las características de calidad los diámetros en mm, d_{23} , d_{18} y d_{13} para x_1 , x_2 , y x_3 respectivamente. Los datos se muestran en tabla 1.

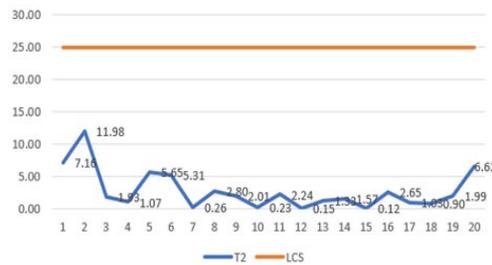
Tabla 1 Datos del proceso

Obs.	x_1	x_2	x_3
1	30.6775	51.6426	69.5642
2	31.9049	48.6413	68.7129
3	30.4495	50.7907	71.3386
4	28.8522	49.7166	71.0117
5	27.9394	48.4184	71.5400
6	28.5328	49.0509	69.7998
7	29.552	50.0653	71.0678
8	29.1609	48.8931	71.7061
9	30.4336	48.9001	70.7920
10	30.2861	49.9758	70.5586
11	30.6746	50.5962	71.4246
12	30.2067	49.8029	70.7522
13	29.4504	49.1902	70.2420
14	29.6636	50.8740	70.8518
15	29.8955	49.7391	70.9987
16	31.1606	50.5534	71.1258
17	30.6053	49.9328	71.0988
18	29.5305	49.1101	71.0864
19	29.9976	50.2463	71.7593
20	28.3486	51.0426	70.5166

Fuente: elaboración propia

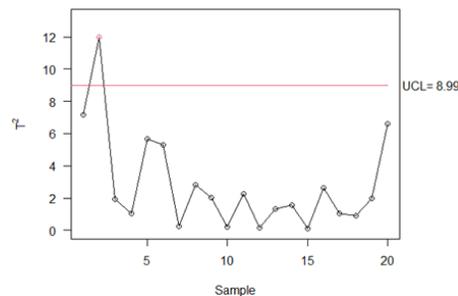
3. Resultados

Computamos los datos en Excel, haciendo uso de la ecuación 7, obteniendo el estadístico T^2 y construimos el gráfico para datos individuales. Computamos los datos en R y Minitab®, obteniendo los gráficos mostrados en figuras 2, 3 y 4.



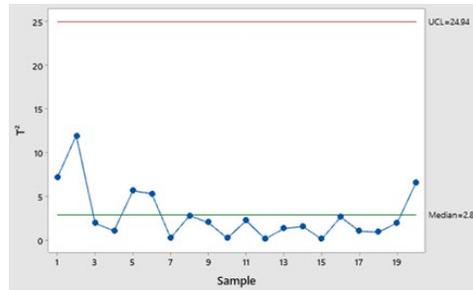
Fuente: elaboración propia

Figura 2 Gráfico de T^2 individuales en Excel.



Fuente: elaboración propia

Figura 3 Gráfico de T^2 individuales en R.

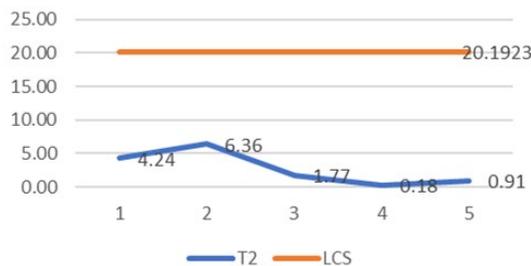


Fuente: elaboración propia

Figura 4 Gráfico de T^2 individuales en Minitab®.

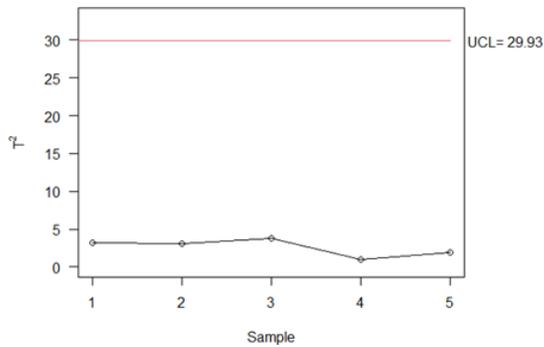
Podemos observar el mismo gráfico en las tres alternativas, todas las observaciones son semejantes, hasta el límite superior de control es idéntico, esto sugiere que los paquetes estadísticos siguen al pie de la letra la teoría multivariante y generan gráficos constantes con ella. Ahora procedemos a utilizar el mismo conjunto de datos, haciendo el análisis para subgrupos, donde tamaño de subgrupo es $n = 4$, y el número de muestras es de $m = 5$, analizando x_1 y x_3 únicamente.

Recordemos que hacer el gráfico en Excel equivale a hacer los cálculos de manera analítica en papel, esto es, los gráficos de control deberían ser iguales al fundamento teórico y por lo tanto a Excel. Computamos los datos en Excel, obteniendo el estadístico T^2 a través de la ecuación 10 y construimos el gráfico para datos individuales. En figuras 5, 6 7 notamos que las tres alternativas para el análisis multivariado, generan el mismo proceso en control estadístico, sin embargo, es evidente, que la forma en que fue calculado varía drásticamente, siendo este más holgado para los dos softwares, tanto Minitab® como R, generaron un límite superior de control más holgado.

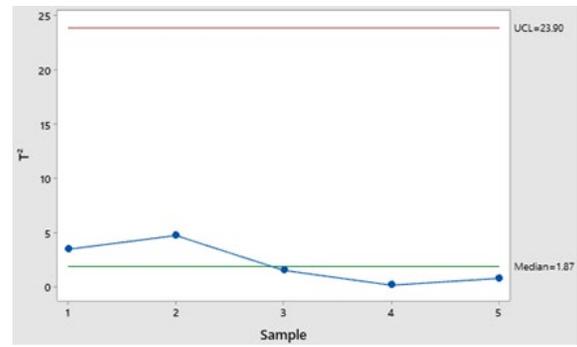


Fuente: elaboración propia

Figura 5 Gráfico de T^2 subgrupos en Excel.



Fuente: elaboración propia



Fuente: elaboración propia

Figura 6 Gráficos de T^2 subgrupos en R. Figura 7 Gráficos de T^2 subgrupos en Minitab.

La tabla 2 presenta una comparativa que resume las ventajas y desventajas de cada software utilizado, incluyendo una calificación ponderada del 1 al 5.

Tabla 2 Comparación de Software para Gráficos de Control Multivariado T^2 de Hotelling

Criterio	Minitab	R	Excel
Facilidad de uso	5 - Interfaz gráfica amigable y fácil de usar	3 - Requiere conocimientos de programación	4 - Interfaz conocida, pero requiere cálculos manuales
Presentación visual	5 - Atractiva y clara	3 - Menos llamativa, pero efectiva	2 - Básica, depende de las habilidades del usuario
Precisión en el cálculo	5 - Alta precisión	4 - Precisa, pero puede requerir paquetes adicionales	3 - Precisión depende de la implementación manual
Capacidad de detección	5 - Alta, detecta desviaciones claramente	5 - Alta, detecta desviaciones claramente	4 - Alta, si los cálculos se realizan correctamente
Flexibilidad	4 - Moderada, preconfigurada	5 - Alta, altamente customizable con paquetes adicionales	2 - Limitada, pero puede ser extendida mediante fórmulas
Requisitos de conocimiento	4 - Básico a intermedio	2 - Avanzado, requiere conocimientos de programación en R	3 - Básico, conocimientos de Excel y estadística
Consistencia con la teoría	5 - Consistente con la teoría multivariante	4 - Consistente con la teoría, pero necesita configuración	4 - Consistente, pero depende de la precisión del usuario
Tiempo de implementación	4 - Rápido	3 - Moderado a largo, dependiendo de la configuración	2 - Largo, requiere cálculos y configuraciones manuales
Coste	2 - Licencia comercial	5 - Gratuito (software libre)	2 - Licencia comercial
Soporte técnico y comunidad	5 - Amplio soporte y documentación	4 - Amplia comunidad de usuarios y paquetes de soporte	3 - Soporte limitado a la documentación de Microsoft
Total	44	38	29

Fuente: elaboración propia

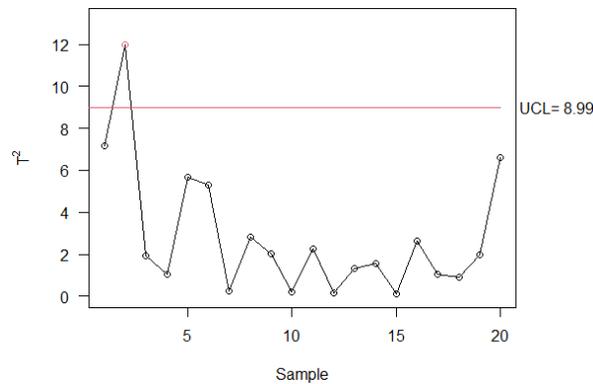
4. Discusión

Como se menciona previamente en la sección de resultados, los gráficos de control multivariado T^2 de Hotelling tanto para datos individuales, así como subgrupos son fácilmente reproducibles en los paquetes estadísticos comerciales,

esto es una noticia excelente para las organizaciones que pueden implementar este análisis para sus procesos y dejar de lado el análisis univariados que lleva a conclusiones erróneas y cometer errores tipo I y II.

Sin embargo, para la reproducción de los paquetes y la equivalencia con Excel que es una hoja de cálculo, usada para representar la teoría de forma más pura, se tiene que tomar en cuenta algunas consideraciones.

Para datos individuales, los 3 paquetes o software se comportaron de una manera idéntica, salvaguardando las distintas consolas e imágenes entregadas, numéricamente entregaron el mismo resultado, sin embargo, el paquete R entrego en primera instancia el gráfico mostrado en figura 8.



Fuente: elaboración propia

Figura 8 Gráfica errónea de control T^2 en R

Enseguida se puede apreciar la diferencia en el cálculo del límite superior de control, que debe ser calculado mediante la ecuación 8, esto no sugiere que el paquete estadístico R no sea capaz de calcular de forma adecuado dicho límite superior de control (UCL), pero necesita un paquete específico dentro de R. ya que el método en R de graficar no permite establecer 2 parámetros clave, el vector de medias y la matriz de covarianza.

Se necesita un vector medias, y la matriz de covarianza en el argumento del comando, de esta manera, investigando más a fondo el paquete (MSQC) para R, creado por Dr. Edgar Santos-Fernández, para producir el gráfico se ofrece la manera correcta de generar el gráfico es agregando los siguientes comandos al código:

```
"library(MSQC)
datos <- matrix(c(
  30.6775, 51.6426, 69.5642,
  31.9049, 48.6413, 68.7129,
  30.4495, 50.7907, 71.3386,
  28.8522, 49.7166, 71.0117,
  27.9394, 48.4184, 71.5400,
  28.5328, 49.0509, 69.7998,
  29.5520, 50.0653, 71.0678,
  29.1609, 48.8931, 71.7061,
  30.4336, 48.9001, 70.7920,
  30.2861, 49.9758, 70.5586,
  30.6746, 50.5962, 71.4246,
  30.2067, 49.8029, 70.7522,
  29.4504, 49.1902, 70.2420,
  29.6636, 50.8740, 70.8518,
  29.8955, 49.7391, 70.9987,
  31.1606, 50.5534, 71.1258,
  30.6053, 49.9328, 71.0988,
  29.5305, 49.1101, 71.0864,
  29.9976, 50.2463, 71.7593,
  28.3486, 51.0426, 70.5166),
  ncol = 3, byrow = TRUE)
Xmv <- mult.chart(datos, type = "t2") $Xmv
S <- mult.chart(datos, type = "t2") $covariance
colm <- nrow(datos)
mult.chart(datos, type = "t2", alpha = 0.0027, Xmv = Xmv, S = S, colm = colm)"
```

Este código genera de manera correcta el gráfico multivariante para datos individuales en R que mostramos en la parte de Resultados, y se comparte en esta sección para su análisis y reproducibilidad por parte de los interesados.

Por otro lado, observamos que, para subgrupos, el gráfico multivariante se comporta de manera similar con claras diferencias en el cálculo de UCL por parte de las tres alternativas, en futuras investigaciones se explorara en el cálculo específico de los paquetes contra la teoría multivariante.

Es claro que el valor de α empleado en el cálculo del UCL es usado de manera distinta por Minitab® y R, El nivel de $\alpha = 0.0027$, usado para el cálculo del UCL en la teoría multivariante, esto explicado por [Montgomery, 2009] como “Supervisar dos o más características de calidad de manera independiente puede llevar a

conclusiones engañosas. Sin embargo, al analizar simultáneamente dos o más variables, la probabilidad de que alguna de ellas supere los límites de control de tres sigmas es extremadamente baja, aproximadamente $\alpha = 0.0027$.”

A manera de discusión, y siguiendo el método propuesto se crea una matriz de comparación de las 3 alternativas en los 2 escenarios estudiados, una vez computados los resultados y teniendo los seis gráficos multivariantes (tres para datos individuales y tres para subgrupos).

Podemos crear esta matriz de relación entre las tres alternativas, evaluar la capacidad de detección es crítica ya que deben ser capaces de identificar desviaciones multivariadas numéricas de manera clara y efectiva además de evaluar si los límites de control establecidos son consistentes y permiten identificar de manera adecuada patrones anormales en los datos.

Por otro lado, es importante analizar cada punto T^2 Hotelling y así identificar cualquier discrepancia significativa en la detección de desviaciones multivariadas utilizando el valor de cada punto T^2 de Hotelling.

La tabla 2 proporciona una guía clara para ayudar a los profesionales a elegir entre Minitab, R y Excel, según sus necesidades específicas en términos de facilidad de uso, presentación visual, precisión y flexibilidad. La calificación total sugiere que Minitab es la opción más robusta, seguida por R y luego Excel.

5. Conclusiones

En este estudio, se realizó una comparación exhaustiva de tres herramientas de software para generar gráficos de control multivariados en el monitoreo de procesos industriales con variables altamente correlacionadas. Los resultados obtenidos son de gran relevancia para la selección adecuada del método de gráfico de control multivariado y la elección de la herramienta de software más apropiada.

En cuanto a la selección del software, se encontró que tanto Minitab® como R son opciones sólidas, ya que demostraron una capacidad comparable para detectar desviaciones multivariadas. Minitab® se destacó por su presentación visual atractiva y amigable, mostrando todos los elementos de manera clara. Por otro lado,

R proporcionó la información necesaria de forma concisa y efectiva, a pesar de tener una presentación visual menos llamativa.

En relación con la capacidad de detección, se observó que los tres softwares fueron capaces de identificar de manera clara los puntos fuera de control. Los límites de control establecidos fueron consistentes entre Excel, Minitab® y R para datos individuales. Sin embargo, se encontraron pequeñas diferencias en la detección de desviaciones multivariadas para datos agrupados.

Es importante destacar que se identificaron discrepancias significativas en la estimación de resultados utilizando el método T^2 de Hotelling, especialmente para datos agrupados. Se observaron discrepancias entre Excel y R, así como entre Minitab® y R en este aspecto. Estas diferencias deben tenerse en cuenta al seleccionar la herramienta de software más adecuada para el análisis de gráficos de control multivariados.

En conclusión, este estudio proporciona una guía útil para los profesionales que deseen utilizar gráficos de control multivariados en el análisis de procesos industriales. Tanto Minitab® como R son opciones recomendables, y la elección entre ellos dependerá de las preferencias específicas del usuario en términos de presentación visual y estimación de resultados. Se recomienda tener en cuenta las discrepancias encontradas en la implementación del método T^2 de Hotelling al seleccionar el software adecuado para las necesidades y requisitos individuales.

6. Bibliografía y Referencias

- [1] Barahona, J., & González-Narváez, M. (2019) Creación de un paquete en R para gráficos de control multivariante con dimensión variable, *Matemática* Vol. 17, No. 2.
- [2] Campos-Avenidaño, G., & Rueda-Molano, Á. (2020). Implementación de gráficos de control multivariantes para mediciones individuales en el sistema automatizado HAS-200.
- [3] Hotelling, H. (1947). Multivariate quality control. *Techniques of statistical analysis*.

- [4] Kariya, T. (1981). A robustness property of Hotelling's T2-test. *The Annals of Statistics*, 211-214.
- [5] Katebi, M., & Moghadam, M. B. (2022). A double sampling multivariate T2 control chart with variable sample size and variable sampling interval. *Communications in Statistics-Simulation and Computation*, 51(7), 3578-3595.
- [6] Montgomery , D. C., Jennings , C. L., & Kulahci , M. (2008). *Introduction to Time Series Analysis and Forecasting*. . NJ: Wiley: Hoboken.
- [7] Montgomery, D. C. (2009). *Introduction to Statistical Quality Control 6th Edition*. New Jersey: John Wiley & Sons, Inc.
- [8] Navarro, J. L. (2003). Una aproximación robusta de la T2 de Hotelling mediante truncamientos. *Asociación Española de Economía Aplicada, ASEPELT.*, p. 221.
- [9] Ruelas, E., López, J. A. V., Salgado, J. C., Márquez, J. A. S., Serrato, R. B., & García, J. A. J. (2020). Control estadístico de procesos multivariantes a través de la red neuronal artificial perceptron multicapa y el análisis del grafico mewma. *IEEE Latin America Transactions*, 18(6), 1041-1048.
- [10] Vanhatalo, E., & Kulahci, M. (2015). The effect of autocorrelation on the Hotelling T2 control chart. *Quality and Reliability Engineering International*, 31(8), 1779-1796.