

# **HERRAMIENTAS GRÁFICAS PARA VISUALIZAR FACTORES DE DISPERSIÓN**

## **GRAPHIC TOOLS TO VISUALIZE DISPERSION FACTORS**

***Cleidy García Rodríguez***

Tecnológico Nacional de México / IT de Celaya, México  
*m2203010@itcelaya.edu.mx*

***Manuel Darío Hernández Ripalda***

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

***Armando Javier Ríos Lira***

Tecnológico Nacional de México / IT de Celaya, México  
*armando.rios@itcelaya.edu.mx*

***Salvador Hernández González***

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

**Recepción:** 22/mayo/2023

**Aceptación:** 19/enero/2024

### **Resumen**

Taguchi plantea que el objetivo general de la Ingeniería de la Calidad es producir productos que sean robustos con respecto a todos los factores de ruido, esto es la variación. La robustez implica que la característica funcional del producto no es sensible a las variaciones en los factores de ruido, es decir que para obtener productos robustos es de suma importancia tener un control sobre la variabilidad. En el siguiente artículo se desarrolla un ejemplo que demuestra la eficacia de las gráficas Multi-vari como herramientas para visualizar la variabilidad y las fuentes que pueden causarlas, demostrándose que son un valioso instrumento para identificar patrones, tendencias y correlaciones que facilitan la comunicación efectiva de datos complejos y brindan información valiosa para la toma de decisiones.

**Palabras Clave:** dispersión, gráficas, diseño de experimento, simulación Montecarlo.

## **Abstract**

*Taguchi states that the general objective of Quality Engineering is to produce products that are robust with respect to all noise factors, that is, variation. Robustness implies that the functional characteristic of the product is not sensitive to variations in noise factors, that is, to obtain robust products it is extremely important to have control over variability. In the following article, an example is developed that demonstrates the effectiveness of Multi-vari graphs as tools to visualize variability and the sources that can cause them, demonstrating that they are a valuable instrument to identify patterns, trends, and correlations that facilitate effective communication of information complex data and provide valuable information for decision making.*

**Keywords:** *dispersion, graphs, experiment design, Montecarlo simulation.*

## **1. Introducción**

Un gráfico Multi-vari es una representación gráfica de las relaciones entre los factores y una variable de respuesta. Se utiliza para presentar el análisis de los datos de varianza en forma gráfica, especialmente en las etapas preliminares del análisis de datos para ver los datos, las posibles relaciones y las causas fundamentales de la variación. Los gráficos Multi-vari son especialmente útiles para comprender las interacciones [Minitab, 2017]. La inspiración para estos gráficos provino de los gráficos utilizados para representar la variación en los precios de las acciones. Dorian Shainin, este adoptó los gráficos Multi-vari como una forma de converger rápidamente en la causa raíz de un problema al que se refirió como la X Roja [Torres & Hernandez-Ripalda, 2019].

La X roja se refiere a la causa raíz de un problema que debe identificarse y eliminarse. La Metodología Shainin Red X (o simplemente Metodología Red X) es una metodología de resolución de problemas desarrollada por Dorian Shainin que se enfoca en encontrar y eliminar la causa raíz de los problemas [Cox, Garside, Kotsialos, & Vitanov, 2013]. La metodología involucra varias herramientas como análisis Multi-vari, gráfico de concentración, búsqueda de componentes, etc [Benchmark Six sigma , 2023], [Nair & Pregibon, 1988].

El gráfico Multi-vari es una de las herramientas utilizadas en la Metodología Red X. Muestra una representación gráfica de la relación entre factores o variables de entrada del proceso y una variable de respuesta o salida del proceso. La característica de calidad de interés se traza a través de múltiples variables o factores en un gráfico. Al analizar el gráfico, se vuelve más fácil identificar qué factor(es) están causando la variación en la variable de salida. Esto ayuda a identificar qué factores necesitan mejorar o eliminar [Minitab, 2017], [Steiner, MacKay, & Ramberg, 2007]. En resumen, un gráfico Multi-vari es una herramienta importante utilizada en la Metodología Red X de Shainin para identificar y eliminar las causas raíz de los problemas. Muestra las relaciones entre los factores o las variables de entrada del proceso y las variables de respuesta o de salida del proceso. Al analizar este gráfico, se vuelve más fácil identificar qué factor(es) están causando la variación en la variable de salida. Es una técnica gráfica que se centra en la causa más probable de un problema al eliminar las causas de variación que no contribuyen. Este es el principio del estudio Multi-vari. Es una técnica de filtrado a través de la cual se elimina un gran número de causas de variación que no contribuyen hasta que, sistemáticamente, se puede destilar la X Roja. En la mayoría de las aplicaciones, el Multi-vari actúa como el primer filtro y generalmente es seguido por:

- Otras técnicas de generación de pistas, si aún quedan demasiadas variables relevantes.
- Juicio de ingeniería: ahora que el confuso desorden de conjeturas se ha reducido sustancialmente.
- B versus C (prueba de dos muestras Tukey-Duckworth) si la variable relevante es solo una y la mejora tiene que ser confirmado como permanente [Tukey, 1959]
- Búsqueda de variables, si el resto de las variables relevantes están entre cinco y veinte, o el factorial completo, si son de dos a cuatro [Bhote & Bhote, 1999].

Un Ejemplo claro de porque el uso de los gráficos de Multi-vari se puede ver en el Ejercicio de Multi-vari en página 131 - 132 del libro de Bhote y Bhote [Bhote & Bhote, 1999]. Dicho ejercicio fue analizado en el Minitab de la siguiente forma:

En la tabla 1, 2, 3 y figura 1 se muestran los datos que se introduce al Minitab la variable respuesta “Y” y los diversos factores que intervienen. (Se re-etiquetan las variables para poderlas trabajar en el Minitab).

Tabla 1 Datos a introducir en el Minitab para el tiempo 1.

Tile	Y	tira
1	66	1
2	56	1
3	58	1
4	65	1
5	67	1
1	59	2
2	58	2
3	56	2
4	48	2
5	63	2
1	54	3
2	32	3
3	59	3
4	48	3
5	72	3

*Fuente: elaboración propia.*

Tabla 2 Datos a introducir en el Minitab para el tiempo 2.

Tile	Y	tira
1	60	1
2	53	1
3	44	1
4	50	1
5	58	1
1	57	2
2	37	2
3	46	2
4	44	2
5	52	2
1	47	3
2	45	3
3	48	3
4	49	3
5	56	3

*Fuente: elaboración propia.*

Tabla 3 Datos a introducir en el Minitab para el tiempo 3.

Tile	Y	tira
1	38	1
2	9	1
3	54	1
4	57	1
5	60	1
1	14	2
2	43	2
3	8	2
4	38	2
5	60	2
1	56	3
2	39	3
3	60	3
4	58	3
5	60	3

Fuente: elaboración propia.

The image shows the 'Enter Data' dialog box in Minitab. On the left, a list of variables includes 'Tile', 'Y', 'tira', and 'tiempo'. The 'Response' field is set to 'Y'. The 'Factor 1' field is 'Tile', 'Factor 2' is 'tira', and 'Factor 3' is 'tiempo'. The 'Factor 4' field is empty. There are buttons for 'Options...', 'Select', 'Help', 'OK', and 'Cancel'.

Fuente: elaboración propia.

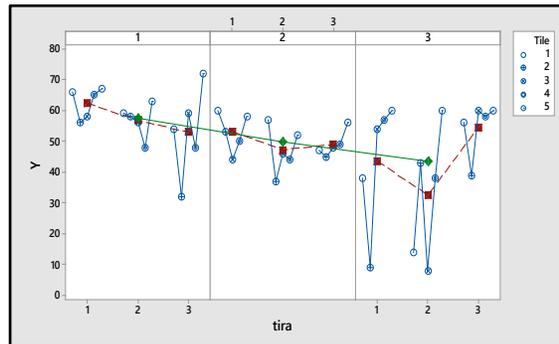
Figura 1 Ventana para datos introducidos al Minitab.

En la figura 2 y 3 se muestra el gráfico Multi-vari para Y por tile –tiempo. En ellos se pueden observar que en la tercera hora hay un cambio notable de variación. Las piezas número 5 aparecen siempre en lugar alto, mientras que las piezas 2 tienden a ser de menor valor, se puede probar una diferencia significativa entre las piezas 5 y 2.

A partir de los resultados obtenidos en el grafico Multi-vari, se realizan un grupo de análisis estadísticos, con ayuda del Minitab 17, que demuestran y explican los resultados sugeridos de la figura 2. En la figura 3 se muestra el grafico para el

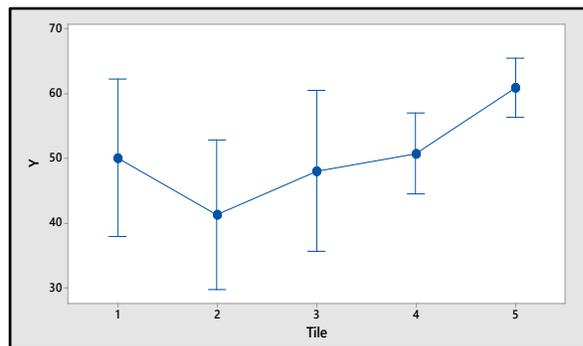
intervalo de confianza de la media con un  $\alpha=0.05$ , para ello se usaron desviaciones estándar individuales para calcular los intervalos.

La media de la pieza 5 es significativamente mayor que la de la pieza 2, las varianzas de las piezas también son diferentes, confirmando la lectura inicial en Multi-Vari. En la tabla 4 también se puede observar las diferencias que existen entre el promedio, la desviación y sus respectivos logaritmos para las piezas 2 y 5.



Fuente: elaboración propia.

Figura 2 Gráfico Multi-vari para Y por tile – tiempo.



Fuente: elaboración propia.

Figura 3 Intervalo de confianza del 95% para la media.

Tabla 4 Análisis de logaritmo natural.

Pieza	Promedio	Desv-est (s)	LN (promedio)	LN (s)
1	50.11111	15.75947	3.91424	2.75744
2	41.33333	15.00833	3.72166	2.70860
3	48.11111	16.14345	3.87351	2.78151
4	50.77778	8.07430	3.92745	2.08868
5	60.88889	5.90433	4.10905	1.77568

Fuente: elaboración propia.

En la tabla 5 se muestran los resultados de la prueba F de Fisher para la varianza de las piezas 2 y 5 y la prueba T de Student para la media de dichas piezas con la ayuda de las funciones de Excel:

- PRUEBA.F(T83: T91,T110:T118). Hipótesis para prueba F:  
H<sub>0</sub>: Las varianzas para la pieza 2 y 5 son iguales.  
H<sub>1</sub>: Las varianzas para la pieza 2 y 5 son diferentes.
- PRUEBA.T(T83: T91,T110:T118,1,3). Hipótesis para prueba T:  
H<sub>0</sub>: Las medias para la pieza 2 y 5 son iguales.  
H<sub>1</sub>: Las medias para la pieza 2 y 5 son diferentes.

Tabla 5 Cálculo de valor P para las pruebas F y T.

Pieza 2 y 5	Valor P
Prueba F	0.01610
Prueba T	0.00213

Fuente: elaboración propia.

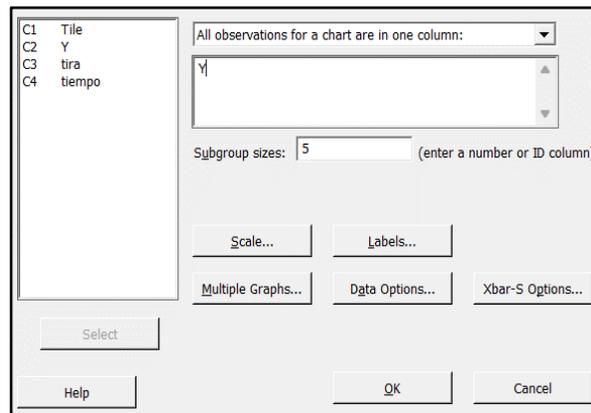
- Región Crítica para prueba F:
  - ✓ Si Valor P <  $\alpha$ , rechazo H<sub>0</sub>: 0.016 < 0.05, Se cumple, rechazo H<sub>0</sub> lo que significa que las varianzas para la pieza 2 y 5 son diferentes.
- Región Crítica para prueba T:
  - ✓ Si Valor P <  $\alpha$ , rechazo H<sub>0</sub>: 0.002 < 0.05, Se cumple, rechazo H<sub>0</sub> lo que significa que las medias para la pieza 2 y 5 son diferentes.

La media de 5 es significativamente mayor que la de 2 (prueba T), las varianzas de las piezas 5 y 2 también son diferentes (prueba F).

Existen varias alternativas de gráficas de variabilidad una de ellas son los Gráficos de Control, Media - Desviación estándar como se muestra en la figura 5.

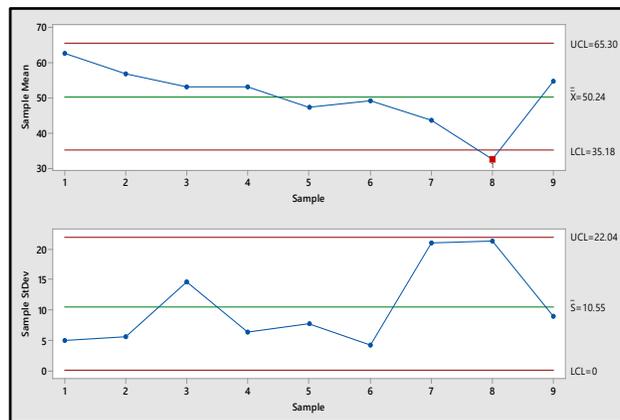
Los valores correspondientes a 1, 2 y 3 corresponden a la hora 1; los valores 4, 5, y 6 a la hora 2 y los valores 7, 8 y 9 a la hora 3. Por lo cual se puede observar que los mismos datos, en gráfica de control, evidencian que en la tercera hora hubo un cambio de varianza, que provoca una señal fuera de control en la segunda tira de la tercera hora. Esto corrobora lo señalado por Multi-Vari.

En comparación con otros tipos de gráficos, un gráfico Multi-vari está diseñado específicamente para identificar la causa raíz de un problema mediante el análisis de múltiples variables simultáneamente. Puede ser particularmente útil cuando se trata de problemas complejos que involucran múltiples factores. Otros tipos de gráficos pueden ser más apropiados para mostrar varios tipos de datos o relaciones entre variables [Six sigma study guide, 2020] [Minitab, 2017].



Fuente: elaboración propia.

Figura 4 Introducción de datos para el Gráfico de control en Minitab 17.



Fuente: elaboración propia.

Figura 5 Gráfico de control Media - Desviación estándar en Minitab 17.

## 2. Método

La simulación Monte Carlo se puede utilizar para realizar pruebas estadísticas en Excel. Estos son los pasos generales para crear una simulación de Montecarlo en Excel:

1. Defina las variables de entrada: identifique las variables de entrada que afectan el resultado de su prueba estadística y defina sus distribuciones de probabilidad.
2. Genere valores aleatorios: use las funciones integradas de Excel como ALEATORIO() o NORM.INV(ALEATORIO(), media, desviación estándar) para generar valores aleatorios para cada variable de entrada.
3. Calcule la salida: use una fórmula que calcule la salida en función de las variables de entrada.
4. Repita los pasos 2 y 3: Repita los pasos 2 y 3 muchas veces (por ejemplo, 10 000) para generar una gran cantidad de salidas.
5. Analice los resultados: Analice los resultados de su simulación de Montecarlo utilizando funciones estadísticas como PROMEDIO (), DESVEST() o PERCENTIL().

### **3. Resultados**

Se desarrolla un ejemplo de aplicación de Grafico Multi-vari apoyándonos del Simulación Monte Carlos para generar y tratar los datos.

- Se generan cinco variables ( $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$ ,  $x_5$ ) con la función de Excel DISTR.NORM.INV(probabilidad; media; desv-estándar) donde la probabilidad es un número aleatorio, la media es 0 y la desviación estándar es 10 para  $x_1$ ,  $x_3$ ,  $x_4$ ,  $x_5$  y 20 para  $x_2$ ., esto con el fin de comprobar si las gráficas Multi-vari son capaz de determinar que  $X_2$  es la X roja. Se generan 1000 datos para cada variable.
- Se calcula la variable respuesta (Y) como la suma de  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$  y  $x_5$  como se muestra en la tabla 6 para cada renglón.
- Luego se identifican los cuartiles a los que pertenece cada una de las variables según la función:

$SI(B2>CUARTIL(B2:B1001;3);4;SI(B2>CUARTIL(B2:B1001;2);3;SI(B2>CUARTIL(B2:B1001;1);2;1)))$

Donde la columna B corresponde a la variable  $x_1$ , se realiza el mismo procedimiento para  $x_2$ ,  $x_3$ ,  $x_4$ ,  $x_5$  como se muestra en tabla 7.

Tabla 6 Variables aleatorias y variable respuesta.

	x1	x2	x3	x4	x5	Y
1	-19.71139	11.65332	0.45639	8.58400	10.69718	11.67950
2	3.11484	3.71776	10.38441	7.77700	-1.59009	23.40393
3	-3.75047	6.11878	11.94220	10.15738	-5.58571	18.88218
4	1.66159	16.63290	-7.78229	-16.10332	0.17107	-5.42006
5	-0.05543	45.44226	1.57799	4.34956	-10.51664	40.79775
6	-2.96969	-8.56672	-8.30515	-0.76414	-8.22417	-28.82987
7	4.24274	-37.84733	-11.96550	19.13949	-6.89842	-33.32902
8	-3.82181	-10.62465	10.12568	-6.82625	-6.41861	-17.56564
9	-1.49597	46.75175	10.42572	1.11284	1.68414	58.47848
10	10.23681	16.56675	10.52999	1.85035	4.48077	43.66467
...	...	...	...	...	...	...
990	-0.42142	13.71679	-11.31336	-5.12007	8.89606	5.75800
991	-19.95618	34.95333	2.38791	-1.55239	7.48021	23.31290
992	17.88985	-5.89419	9.30697	-29.63406	-2.43976	-10.77119
993	22.05481	1.54356	15.92210	-1.94189	14.51251	52.09108
994	-23.86424	-22.23934	-3.78447	11.44679	0.11848	-38.32279
995	9.84881	40.35502	-6.53652	-7.66711	-1.52783	34.47237
996	16.68328	0.92439	-14.54256	4.31361	-7.50799	-0.12927
997	-10.70690	21.65581	-8.78442	1.61415	5.01476	8.79340
998	10.18040	-0.29517	1.71760	7.34420	-3.27623	15.67080
999	2.33716	1.38739	8.65778	-17.07729	-19.48893	-24.18389
1000	10.16019	-24.39294	-15.51362	6.43049	28.60596	5.29007

Fuente: elaboración propia.

Tabla 7 Cuartiles para cada variable.

	q1	q2	q3	q4	q5
1	1	3	3	4	4
2	3	3	4	4	2
3	2	3	4	4	2
4	3	4	1	1	2
5	2	4	3	3	1
6	2	2	1	2	1
7	3	1	1	4	1
8	2	2	4	1	1
9	2	4	4	3	3
10	4	4	4	3	3
...	...	...	...	...	...
990	2	3	1	2	4
991	1	4	3	2	4
992	4	2	4	1	2
993	4	3	4	2	4
994	1	1	2	4	2
995	4	4	2	1	2
996	4	2	1	3	1
997	1	4	1	3	3
998	4	2	3	4	2
999	3	3	4	1	1
1000	4	1	1	3	4

Fuente: elaboración propia.

- Para explicar que variables tienen mayor variación sobre la respuesta se ordenan las Y de menor a mayor en la tabla 8.

Tabla 8 Orden de la Y de menor a mayor.

	q1	q2	q3	q4	q5	Y
1	1	1	2	2	1	-95.67501
2	1	1	2	1	1	-92.09218
3	2	1	1	1	1	-87.61733
4	1	1	1	2	1	-83.42284
5	2	1	1	1	2	-83.42259
6	1	1	1	3	1	-77.37091
7	1	1	2	3	1	-76.64018
8	1	1	2	2	1	-76.24521
9	1	1	3	1	1	-76.17585
10	1	1	2	2	3	-76.01783
...						
990	4	4	4	2	4	69.33260
991	4	4	1	3	4	70.08900
992	4	4	4	1	3	72.37434
993	4	4	1	3	4	74.35628
994	1	4	4	4	4	74.62650
995	2	4	4	3	2	75.35832
996	4	4	4	4	2	76.10341
997	4	4	4	4	3	76.18377
998	3	4	4	4	2	79.62197
999	4	4	3	4	4	92.42220
1000	4	4	4	4	3	94.10792

Fuente: elaboración propia.

Se propone una prueba simple (Tukey-Duckworth) para comparar dos muestras. Si una muestra contiene el valor más alto y la otra el valor más bajo de las dos muestras combinadas, el total de dos recuentos se toma como estadística. Estos recuentos son de (i) el número de individuos en una muestra por encima de todos los individuos en el otro, y (ii) el número de individuos en el otro por debajo de todos los de uno. Los valores críticos de 5%, 1% y 0,1% son aproximadamente 7, 10 y 13 respectivamente. El resultado es que para  $x_2$  (identificada como q2) los cuartiles extremos tienen una cuenta superior a 7

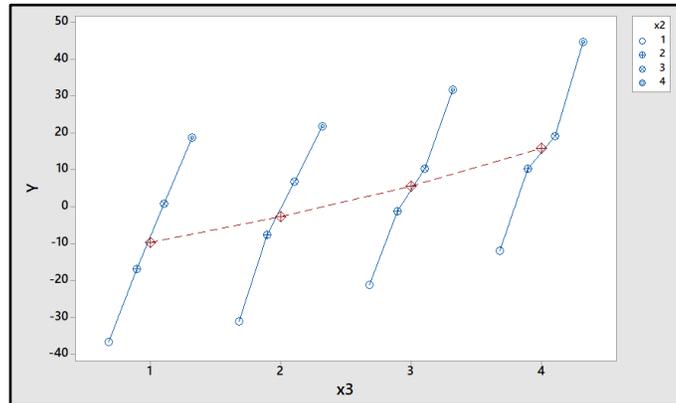
En la figura 7 se encuentra un análisis de varianza realizado en el Minitab en el cual se observa que la variable con mayor efecto sobre la variable respuesta es  $x_2$  pues es la variable con mayor suma de cuadrado (379,512).

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
x1	3	83885	27962	248.82	0.000
x2	3	379512	126504	1125.70	0.000
x3	3	82467	27489	244.61	0.000
x4	3	73185	24395	217.08	0.000
x5	3	98735	32912	292.86	0.000
Error	984	110580	112		
Falta de ajuste	630	70669	112	0.99	0.525
Error puro	354	39911	113		
Total	999	854717			

Fuente: elaboración propia.

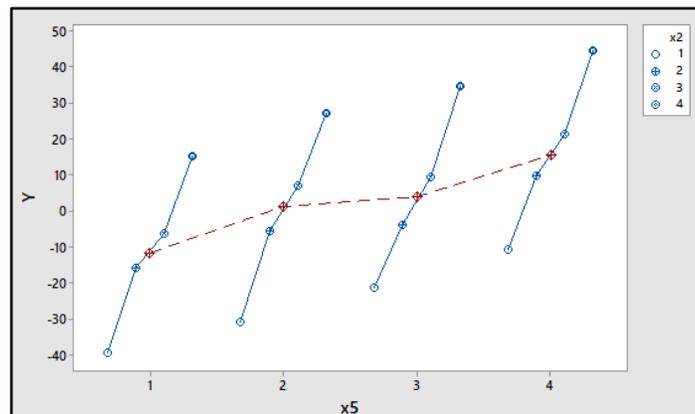
Figura 7 Análisis de varianza.

En las figuras 8, 9 y 10 se presentan varias Gráficas Multi-vari que muestran la interacción entre algunas variables, demostrando nuevamente que la variable con mayor efecto sobre la variabilidad de la respuesta es  $x_2$ .



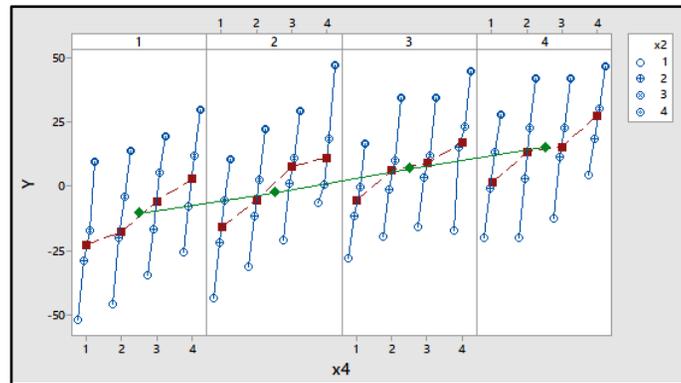
Fuente: elaboración propia.

Figura 8 Gráfica Multi-vari para las variables  $x_2$  y  $x_3$ .



Fuente: elaboración propia.

Figura 9 Gráfica Multi-vari para las variables  $x_2$  y  $x_5$ .



Fuente: elaboración propia.

Figura 10 Gráfica Multi-vari para las variables  $x_1$ ,  $x_2$  y  $x_4$ .

## 4. Discusión

Las gráficas Multi-vari son representaciones visuales que permiten explorar y analizar la relación entre múltiples variables en un solo gráfico. Estas representaciones son útiles cuando se busca comprender la interacción y la dependencia entre diferentes variables en un conjunto de datos.

La grafica Multi-vari es una herramienta estadística y de calidad capaz de resaltar patrones de variación dentro de subgrupos y entre subgrupos, de acuerdo a familias de variación, como tiempo-tiempo, máquina-máquina, localización-localización. Es muy efectiva en la búsqueda de la variable o las variables que tienen una significancia mayor en la variabilidad del proceso.

La adecuada utilización de estas graficas permite la reducción de una lista de varias posibles causas de variabilidad, ofreciendo un mayor enfoque en las causas más significativas. Brindan además un panorama rápido del proceso sin la necesidad de tener mucha información. Son un valioso instrumento para visualizar y analizar la relación entre múltiples variables. Permiten identificar patrones, tendencias y correlaciones, facilitan la comunicación efectiva de datos complejos y brindan información valiosa para la toma de decisiones.

## 5. Conclusiones

- Un estudio Multi-vari permite reducir un gran número de causas desconocidas e inmanejables de variación a una familia mucho más pequeña de variables relacionadas que contienen el rojo. X, es decir, la causa dominante.

- Los resultados obtenidos por una gráfica Multi-vari pueden ser corroborados con Pruebas F, Pruebas T, Intervalos de confianza para la media y Gráficos de control para la Media contra desviación estándar.
- La X roja arrojada por el Multi-vari puede ser corroborada con un análisis de varianza, analizando cual es la variable con mayor peso en la suma de cuadrado.
- Las gráficas de Multi-vari permiten identificar qué variables están causando la variación en la variable de salida; o cual de las variables es la que más influye en la variación. Esto ayuda a identificar qué factores se necesitan mejorar o eliminar.

## **6. Bibliografía y Referencias**

- [1] Benchmark Six sigma . (31/enero/2023). <https://www.benchmarksixsigma.com/forum/topic/39415-red-x-methodology/>.
- [2] Bhote, K. R., & Bhote, A. K. (1999). World class quality: using design of experiments to make it happen. AMACOM Div American Mgmt Assn.
- [3] Cox, S., Garside, J., Kotsialos, A., & Vitanov, V. (2013). Concise process improvement definition with case studies. *International Journal of Quality & Reliability Management*, 9(30), 970-990.
- [4] Feldman, K. (07 de nov de 2018). <https://www.isixsigma.com/dictionary/multi-vari-chart/>. Obtenido de <https://www.isixsigma.com/dictionary/multi-vari-chart/>.
- [5] Goethals, P. L., & Cho, B. R. (2011). The development of a robust design methodology for time-oriented dynamic quality characteristics with a target profile. *Quality and Reliability Engineering International*, 4(27), 403-414.
- [6] Joseph, C. J., & Wu, C. J. (2002). Robust parameter design of multiple-target systems. *Technometrics*, 44(4), 338-346. 44, 4, 338-348.
- [7] Lunani, M., Nair, V. N., & Wasserman, G. S. (1997). Graphical methods for robust design with dynamic characteristics. *Journal of Quality Technology*, 3(29), 327-338.
- [8] Nair, V. N., & Pregibon, D. (1988). Signal-to-Noise Ratios, Performance Criteria, and Transformations: Discussion. *Technometrics*, 1(30), 24-29.

- [9] Minitab. (2017). <https://support.minitab.com/en-us/minitab/21/help-and-how-to/quality-and-process-improvement/quality-tools/supporting-topics/multi-vari-chart-basics>.
- [10] Six sigma study guide. (feb de 2020). <https://sixsigmastudyguide.com/multi-vari-study-multi-vari-charts/>.
- [11] Steiner, S. H., MacKay, R. J., & Ramberg, J. S. (2007). An overview of the Shainin System™ for quality improvement. *Quality Engineering*, 1(20), 6-19.
- [12] Torres, R. L., & Hernandez-Ripalda, M. D. (2019). Gráfico Multi-Vari Vs Gráfico de Variabilidad: Análisis Comparativo. *Pistas Educativas*, 133(41).
- [13] Tukey, J. W. (1959). A quick compact two sample test to Duckworth's specifications. *Technometrics*, 1(1), 31-48.