

UN ÍNDICE DE CAPACIDAD DE PROCESOS PARA DISTRIBUCIONES MULTIVARIADAS NORMALES Y NO NORMALES, DE VARIABLES CORRELACIONADAS Y NO CORRELACIONADAS

Erasmus Rafael Salazar*

UNIVERSIDAD DE ORIENTE, NÚCLEO DE SUCRE, VENEZUELA

Recibido: 31 de enero del 2020 / Aprobado: 7 de mayo del 2020

doi: 10.26439/ing.ind2020.n038.4814

RESUMEN: En el análisis de capacidad de procesos multivariados existen muchos índices que solo se aplican cuando los datos son *normales* y otros cuando los datos son *no normales*; lo mismo ocurre cuando las variables de calidad están *correlacionadas* y *no correlacionadas*. En este trabajo se propone un índice de capacidad multivariado *CPME*, desarrollado bajo el uso inicial de un índice univariado, según sea el caso normal o no normal, y cualquier correlación entre variables, para luego, a través de una función característica, extenderlo para el caso multivariado. Este índice puede ser aplicado para todos los casos anteriores. Como utilidad, presentamos ejemplos de aplicación de esta alternativa sobre un conjunto de datos reales y simulados, donde se encontró un amplio desempeño del índice propuesto frente a otros de capacidad similares.

Palabras clave: índice de capacidad de proceso multivariado / índice de capacidad univariado / distribución no normal / distribución normal / métodos estadísticos

A PROCESS CAPABILITY INDEX FOR MULTIVARIATE NORMAL AND NON-NORMAL DISTRIBUTIONS OF CORRELATED AND UNCORRELATED VARIABLES

ABSTRACT: The multivariate process capability analysis includes many indices that are only used when the data is normal and others, when the data is not normal. The same occurs with correlated and uncorrelated quality variables. In this research work, a CPME multivariate capability index was developed by initially using a univariate

* erasmo10464407@gmail.com

index—depending on whether the data was or not normal—and any correlation between variables, and then through a characteristic function for the multivariate data. This index may be used in all the aforementioned cases. Some examples of this alternative are presented on a set of real and simulated data, where a broad performance of our proposed index was found against other similar capability indices.

Keywords: multivariate process capability index / univariate capability index / non-normal distribution / normal distribution / statistical methods

1. INTRODUCCIÓN

Los estudios de los índices de capacidad de procesos (*ICP*) se han desarrollado en el campo univariado y multivariado (*ICPM*) y son estimaciones numéricas de la capacidad del proceso, es decir, nos dan una idea de a qué nivel el proceso cumple con las especificaciones establecidas. Los *ICP* son estadísticos útiles, ya que aparte de ser sencillos de calcular, no tienen unidades de medida; por lo que permiten comparar distintos procesos. Básicamente, son el cociente entre la amplitud tolerable del proceso (la distancia entre los límites de tolerancia o límites de especificación), y la amplitud real o natural del proceso (habitualmente, la distancia entre los límites de control es de 6 sigma). Algunos de estos estadísticos se definen a partir de la media del proceso o del objetivo. Los índices de capacidad asociados con la variación a corto plazo son C_p , C_{pk} , CPU y CPL ; y por otro lado, los asociados con la variación a largo plazo son P_p , P_{pk} , PPU , PPL y C_{pm} . En la práctica, se suele considerar que 1,33 es el valor mínimo aceptable para un índice de capacidad; es decir, cualquier valor por debajo de esta cifra indicaría que, aunque esté bajo control estadístico, el proceso no cumple con las especificaciones deseadas. Estos índices se usan bajo el supuesto de que el valor de las mediciones de las variables de calidad representa valores de variables aleatorias independientes con distribución normal y que el proceso se encuentra en control estadístico, como lo establece Montgomery (2009); pero si la distribución fundamental es no normal, entonces se pueden generar resultados erróneos. En la industria, existen muchos procesos de manufactura donde la variabilidad del proceso se puede describir por una distribución normal y otros por una distribución no normal. En este último caso de poca literatura, no se deberían usar las fórmulas de la distribución normal, pues Somerville y Montgomery (1996) reportaron que el uso de índices de capacidad de procesos no normales, en vez de normales, generan resultados no confiables para este tipo de procesos. Sin embargo, podemos usar el método de percentil de Clements para estimar los índices de capacidad para estos procesos. Este método se utiliza para calcular los índices C_p y C_{pk} mediante una familia de curvas de Pearson y consiste en estimarlos mediante una simple modificación de los *ICP* normales, utilizando los valores de los puntos percentiles de distribuciones no normales, usando la mediana como medida central y no la media. Salazar y Fermín (2016 y 2017) definen un índice multivariado para distribuciones normales y luego para distribuciones no normales, respectivamente, aplicando estadísticos univariados, como los de Clements, para casos no normales, y otros comunes, para los casos normales; donde se trabaja en ambos casos con variables correlacionadas y no correlacionadas. Cabe destacar, muchos autores de índices no toman en consideración la correlación de las variables, hecho que es muy importante. En los estudios de índices de capacidad multivariados, según Foster, Barton y Gautam (2005), aún no hay una metodología que sea consistente para calcular índices de capacidad multivariados (*ICPM*) y, sobre todo, para el caso no normal; por lo tanto, hasta la fecha no existe consenso sobre el uso de un índice en

particular. Cuamea y Rodríguez (2014) señala también que los índices de capacidad más usados suponen que las variables de calidad sean no correlacionadas. Por lo tanto, no existe un *ICPM* que sea consistente y suficiente para el caso normal y no normal, para variables correlacionadas y no correlacionadas. De aquí que el diseño del índice *CPME* viene a constituir un índice que puede ser aplicado para cualquier tipo de dato multivariado y cualquier tipo de correlaciones, a diferencia de los otros índices multivariados que solo se aplican a un caso particular; lo que constituye una herramienta de amplio uso en el campo de la capacidad de procesos.

El método de percentiles de Clements, propuesto para determinar los *ICP*, reemplaza en la ecuación del C_p el valor 6σ por la longitud del intervalo entre el límite superior e inferior de los puntos percentiles 0,135 y 99,865 de una distribución X , esto es respectivamente $U_p - L_p$.

$$CC_p = \frac{U - L}{U_p - L_p} \quad (1)$$

donde U_p es el percentil 99,865 y L_p es el percentil 0,135.

Para el C_{pk} , la media del proceso μ es estimado por la mediana M_d y los valores de 3σ son estimados por $M_d - L_p$ y $U_p - M_d$, respectivamente. Así se obtienen las siguientes expresiones:

$$CC_{pl} = \frac{M_d - L}{M_d - L_p} \quad (2)$$

$$CC_{pu} = \frac{U - M_d}{U_p - M_d} \quad (3)$$

$$CC_{pk} = \min \{CC_{pk}, CC_{pu}\} \quad (4)$$

Los *ICPM*, en general, pueden ser obtenidos de tres maneras: a) desde la proporción de una región de tolerancia a una región de proceso, b) calculando la probabilidad de productos que no cumplen con las reglas y c) por los enfoques en los que las funciones de pérdida son empleadas. Además, según Cuamea y Anaya (2009):

En el estudio multivariado, las variables críticas deben estar dentro de sus respectivas especificaciones para que el producto se considere aceptable. Si solo una de ellas no cumple con las especificaciones esto será suficiente para que el cliente considere inaceptable el producto.

En base a esto se continúa el índice para el caso multivariado, en donde se hace el estudio univariado a cada variable de datos normales y no normales, para luego concluir respecto al caso multivariado.

2. ÍNDICES DE CAPACIDAD DE PROCESOS MULTIVARIADOS

Citando algunos autores, como Hubele, Shahiari y Cheng (1991), se tiene que, usando la distribución multivariada normal, definen un *ICPM* como la razón de una región de tolerancia rectangular y una región de tolerancia modificada que es el rectángulo más pequeño alrededor de la elipse con un error especificado de tipo I ($\alpha = 0,0027$). El número de las variables de calidad en el proceso es tenido en cuenta tomando la v -ésima raíz de la proporción, donde v representa el número de las variables de calidad.

$$C_{PM} = \sqrt[v]{\frac{\text{Volumen de la región de tolerancia de ingeniería}}{\text{Volumen de la región de procesos de ingeniería}}}$$

Taam, Subbaiah y Liddy (1993) definen un índice de capacidad multivariado (MC_{PM}) como la razón del volumen de la región de tolerancia modificada y el volumen de la escala 99,73 % de la región de proceso. La región de tolerancia modificada es definida como el elipsoide más grande centrado en el blanco totalmente ubicado dentro de la región de tolerancia original. Luego, para ajustar el cálculo estimado del índice de capacidad normal multivariado, usaron:

$$\widehat{MC}_{PM} = \frac{C_{PM}}{\widehat{D}}$$

donde:

$$\widehat{D} = \sqrt{1 + \frac{n}{n-1} (\bar{X} - \mu_0) S^{-1} (\bar{X} - \mu_0)^T}$$

\bar{X} es el vector media muestral

S es la matriz de varianza-covarianza muestral

μ_0 es la media objetivo del proceso

Wang, Hubele, Lawrence, Miskulin y Shahriari (2000) compararon los *ICPM* mencionados, anteriormente, y presentaron algunos ejemplos gráficos para ilustrarlos. Sin embargo, la correlación entre las variables no fue tenida en cuenta. Castagliola y Castellanos (2005) prolongaron el método univariado de Castagliola a la distribución multivariada, reemplazando la función de densidad de probabilidad univariada, $f(x)$, con su versión multivariada. Mientras que Wang y Du (2000) proponen un método que utiliza el análisis de componentes principales para describir el comportamiento de un

proceso para datos multivariados. Cuamea y Anaya (2009) proponen los índices C_{PM} y C_{PKM} para una distribución normal multivariada para variables múltiples con correlación, sin embargo, C_{PM} puede usarse para variables sin correlación.

3. MATERIALES Y MÉTODOS

Como se mencionó, los $ICPM$ pueden ser obtenidos desde la proporción de una región de tolerancia a una región de proceso; y éste es nuestro asunto. En el caso univariado, se tiene que los límites de especificación inferior y superior, para el control de la variable de calidad X , el objetivo que se persigue es que: $(LEI \leq X \leq LES)$, donde LEI y LES son los límites de especificación inferior y superior establecidos para la variable. Ahora, para el caso multivariado, lo anterior alcanzaría a extenderse y el intervalo sería reemplazado por una región de especificación; dicha región estaría construida por los intervalos de especificación para cada una de las variables de calidad. La región de especificación sería representada mediante las desigualdades:

$$I_{i=1}^v (LEI_i \leq X_i \leq LES_i)$$

Como se puede observar, en todos trabajos de $ICPM$, los índices expuestos se basan en un solo caso de distribución: normal o no normal. Además, como lo señala Cuamea y Rodríguez (2014), ninguno considera la correlación que pudiera existir entre las variables. En este trabajo consideramos X_1, X_2, \dots, X_v v variables aleatorias obtenidas de un proceso de calidad, modeladas por una función de distribución multivariada normal o no normal, correlacionadas o no correlacionadas.

La metodología consiste en que, dado el caso de las variables consideradas, según la distribución de los datos y definida la región de tolerancia del proceso, se obtienen las estimaciones de los parámetros para el cálculo de cada ICP para cada $X_i, i = 1, v, \dots, v$. Luego, se aplica una función característica a cada ICP para cada $X_i, i = 1, v, \dots, v$, para luego determinar el índice multivariado CPME (capacidad de proceso multivariado estadístico), el cual será comparado con otros índices multivariados.

Se usaron varias muestras de diferentes tamaños donde las variables estuvieron correlacionadas y no correlacionadas.

3.1 Proposición de índice de capacidad proceso (CPME)

Sea v el número de variables que se consigue del estudio de un proceso de calidad, donde se obtiene de un muestreo un conjunto de datos modelados por una distribución, normal o no normal, de la cual se consiguen las estimaciones de los parámetros para determinar el $ICPM$ que va fijar la capacidad del proceso de producción. En función a esto, se

evaluó el desempeño de los métodos de estimación de capacidad existentes, para datos con variables de calidad multivariados normal y no normal; y variables correlacionadas y no correlacionadas.

Sea $\mathbf{X} = (X_1, X_2, \dots, X_v)'$ el vector de las v variables de calidad del proceso

Sea $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_v)'$ el vector de medias y Σ su matriz de covarianzas

Sea $S = \{\mathbf{X} \in \mathbb{R}^v: LEI_i \leq X_i \leq LES_i\} i = 1, \dots, v$, la región de tolerancias del proceso

Se obtienen las estimaciones de los parámetros para el cálculo de cada ICP para cada $X_i, i = 1, \dots, v$ y se propone el $ICPM$ definido de la siguiente manera:

Sea

$$\chi_{A_i}(Icp_i) = \begin{cases} 1 & \text{si } Icp_i \in A_i \\ 0 & \text{si } Icp_i \notin A_i \end{cases}$$

donde

$A_i = \text{Rango}(Icp_i) = \{Icp_i/Icp_i \geq 1,33\}$ para cada $X_i, i = 1, \dots, v$

Icp_i es un índice de capacidad univariado seleccionado

$Icp = CC_p$ es adecuado para el caso no normal; caso normal $Icp = C_p$ o PP

Se define el índice propuesto, $CPME$ como:

$$CPME = \frac{\sum_{i=1}^v \chi_{A_i}(Icp_i)}{v}$$

Ahora bien, por definición de $CPME$ se tiene que:

Si $CPME = 1$, entonces el proceso de producción es estable, ya que todas las variables están en el rango de especificación dado.

Si $CPME \neq 1$, entonces el proceso de producción no es estable, ya que existe al menos una $X_i, i = 1, \dots, v$ que está fuera del rango de especificación dado.

4. RESULTADOS

En las siguientes secciones están los ejemplos de la obtención del índice de capacidad $CPME$ propuesto en este trabajo y su comparación con otros índices; tomando en consideración los casos con variables no normales y normales, con variables correlacionadas y no correlacionadas.

4.1 Cálculo de CPME caso no normal y normal. Comparación con otros índices propuestos

4.1.1 Caso no normal

En este proceso se usaron varios casos de muestras diferentes de datos no normales simulados, obtenidos a través del método de Monte Carlo, y se trabajó con las variables correlacionadas y no correlacionadas. En cada uno de los casos se calcularon índices univariados; luego, el índice multivariado propuesto en este trabajo, y se compararon con otros índices obtenidos. Además, para que el estudio fuera eficiente y suficiente, se generaron aleatoriamente las variables C_1 y C_2 y no normales, para los 6 tamaños diferentes de muestras $n = 10, 20, 50, 100, 1000$ y $10\ 000$.

Este caso se analizó de manera univariante y bivariante, usando los paquetes estadísticos de Minitab 17, para el caso univariante, y Statgraphics Centurion, para el caso multivariante. Sin embargo, ya que las variables son no normales, se escogió la técnica aplicada de Box y Cox (1964), para el cálculo de los índices multivariados de Cuamea y Rodríguez (2014), Capacidad del Proceso Multivariante (MCP) y de Taam *et. al.* (1993). Así, se obtienen las siguientes tablas de resultados.

4.1.1.1 Variables correlacionadas

Tabla 1

Resumen descriptivo para las variables no normales C_1 y C_2 para diferentes tamaños de muestras

Muestra		n = 10			
Variable	LEI	LES	Objetivo	Media	Desv.Est.(General)
C_1	0,2	3,2	1,2	1,17	1,1
C_2	0,05	10,3	5,0	3,6	3,4
Muestra		n = 20			
Variable	LEI	LES	Objetivo	Media	Desv.Est.(General)
C_1	0,01	3,6	1,2	0,9	0,97
C_2	0,2	20,0	5,0	5,9	5,6
Muestra		n = 50			
Variable	LEI	LES	Objetivo	Media	Desv.Est.(General)
C_1	0,01	2,9	1,2	0,99	0,79
C_2	0,17	21,9	5,0	4,8	4,81
Muestra		n = 100			
Variable	LEI	LES	Objetivo	Media	Desv.Est.(General)
C_1	0,05	7,2	1,2	1,1	1,18
C_2	0,17	29,0	5,0	6,6	6,7
Muestra		n = 1000			

(continúa)

(continuación)

Variable	LEI	LES	Objetivo	Media	Desv. Est. (General)
C_1	0,00008	7,0	1,2	1,07	1,07
C_2	0,0017	38,0	5,0	5,2	5,25
Muestra			n = 10 000		
Variable	LEI	LES	Objetivo	Media	Desv. Est. (General)
C_1	0,00001	9,7	1,2	1,01	1,01
C_2	0,0002	44,0	5,0	5,02	5,01

Elaboración propia

Tabla 2
ICP univariado de las variables no normal C_1 y C_2 para diferentes tamaños de muestras

Variables	Tamaño de muestra											
	n = 10		n = 20		n = 50		n = 100		n = 1000		n = 10 000	
	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
Índice de capacidad												
PP	0,49	0,29	0,57	0,51	1,39	0,93	1,08	0,69	0,96	1,06	1,44	1,32
PPI	0,79	0,98	0,99	0,95	0,94	0,95	0,94	0,97	1,00	1,00	1,00	1,00
PPu	0,44	0,25	0,52	0,46	1,47	0,71	1,10	0,65	0,96	1,06	1,49	1,25
P_{pk}	0,44	0,25	0,52	0,46	0,94	0,71	0,94	0,65	0,96	1,00	1,00	1,00
CC_p	1,00	0,99	0,98	1,00	0,98	1,00	1,00	1,00	1,00	1,00	1,44	1,28
CC_{pk}	0,98	0,99	0,98	0,99	0,98	1,00	1,00	1,00	1,00	0,98	1,00	1,00
Correlación de Pearson	0,904		0,982		0,953		0,951		0,997		1,00	

Elaboración propia

Ahora, para el caso multivariado, los *ICPM* fueron los siguientes:

Tabla 3
Cálculo de *ICPM* para diferentes tamaños de muestra respecto a la variable $C=(C_1, C_2)$ no normal

Índice de capacidad multivariado		n = 10	n = 20	n = 50	n = 100	n = 1000	n = 10 000
Cuamea y Rodríguez (2014)	C_{pM}	0,44	0,44	0,57	0,6	21,5	82,06
Statgraphics (2015)	MCP	0,35	0,42	0,52	0,60	0,78	0,86
Taam <i>et al.</i> (1993)	MC_p	3,8	5,7	6,8	4,15	12,3	∞
Índice propuesto en esta investigación	$CPME$ (CC_p)	$\neq 1$					

Elaboración propia

Los gráficos de tolerancia para algunos casos se observan en la figura 1.

Gráfico de tolerancia para C_1, C_2 . Variables correlacionadas. n=10

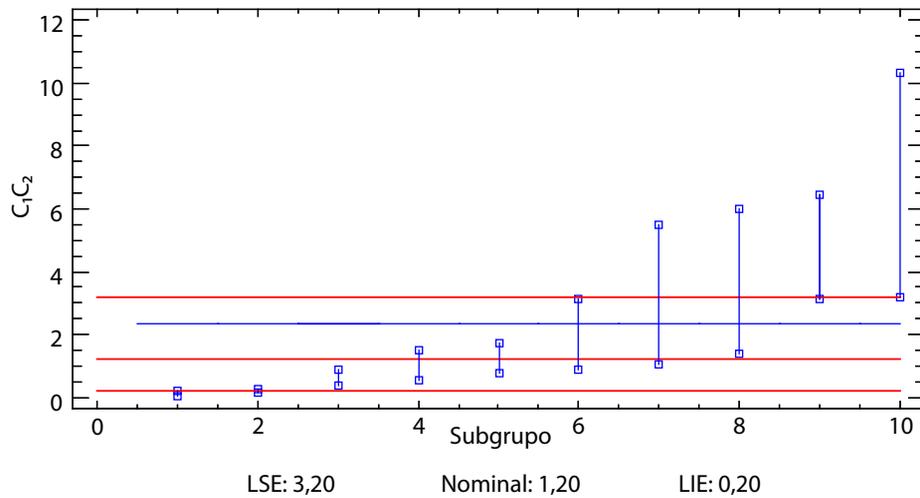


Gráfico de tolerancia para C_1, C_2 . Variables correlacionadas. $n=20$

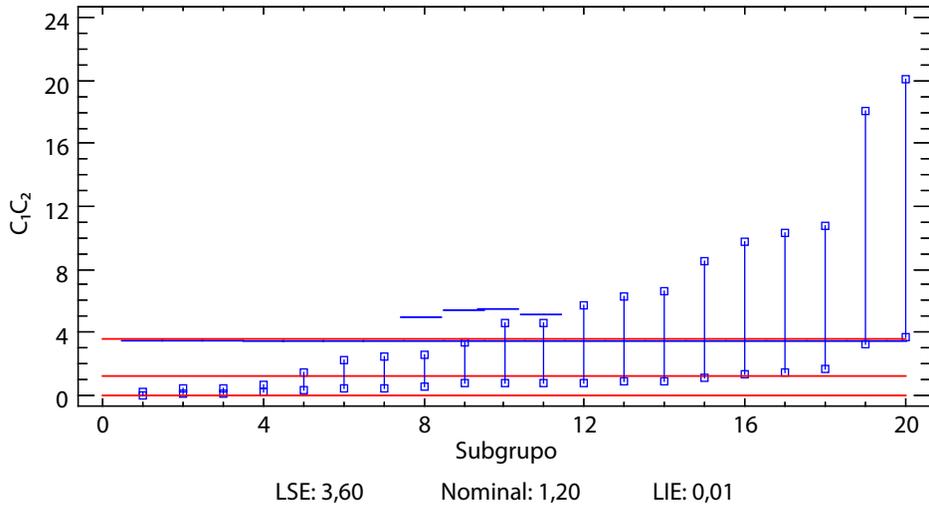


Gráfico de tolerancia para C_1, C_2 . Variables correlacionadas. $n=50$

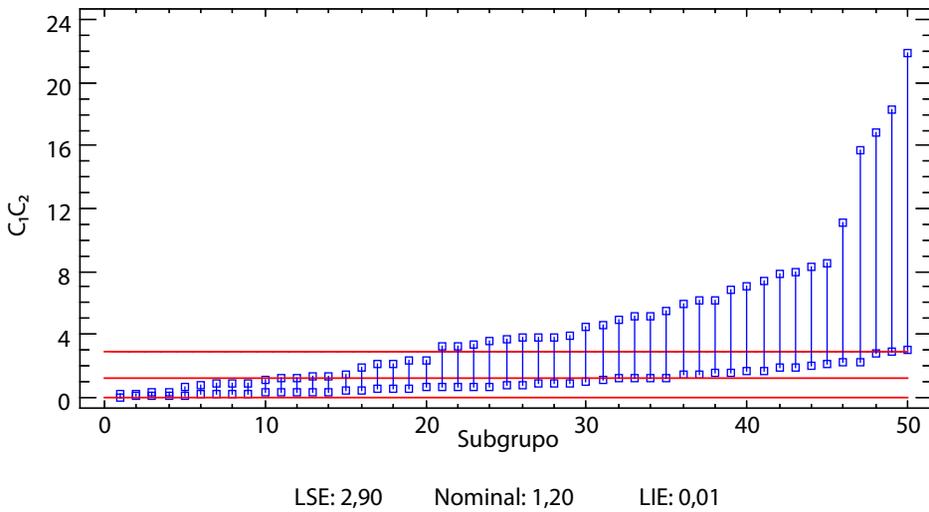


Gráfico de tolerancia para C_1, C_2 . Variables correlacionadas. $n=100$

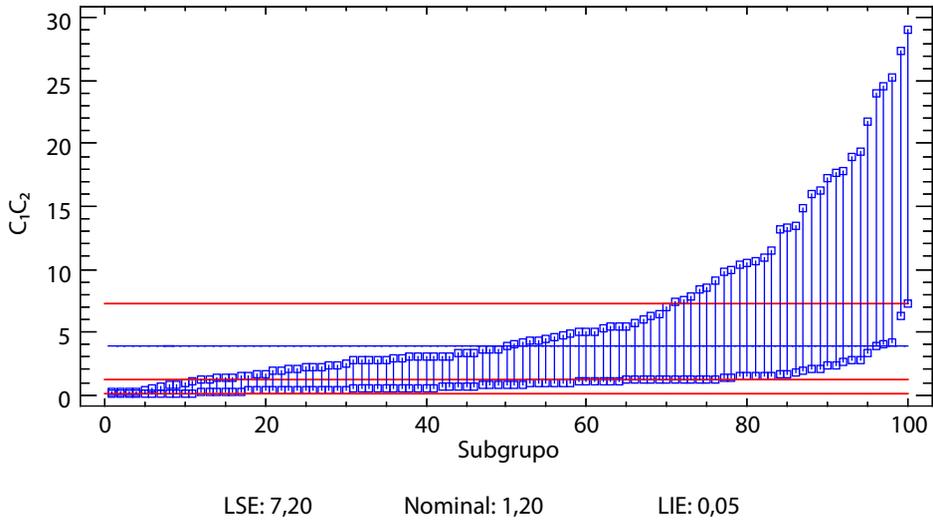


Gráfico de tolerancia para C_1, C_2 . Variables correlacionadas. $n=1000$

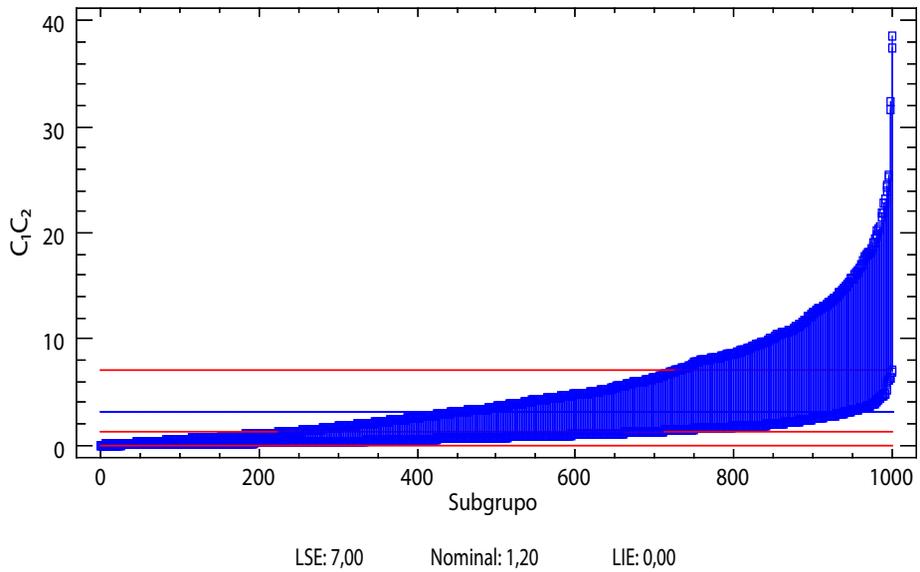


Gráfico de tolerancia para C_1, C_2 . Variables correlacionadas. $n=10\ 000$

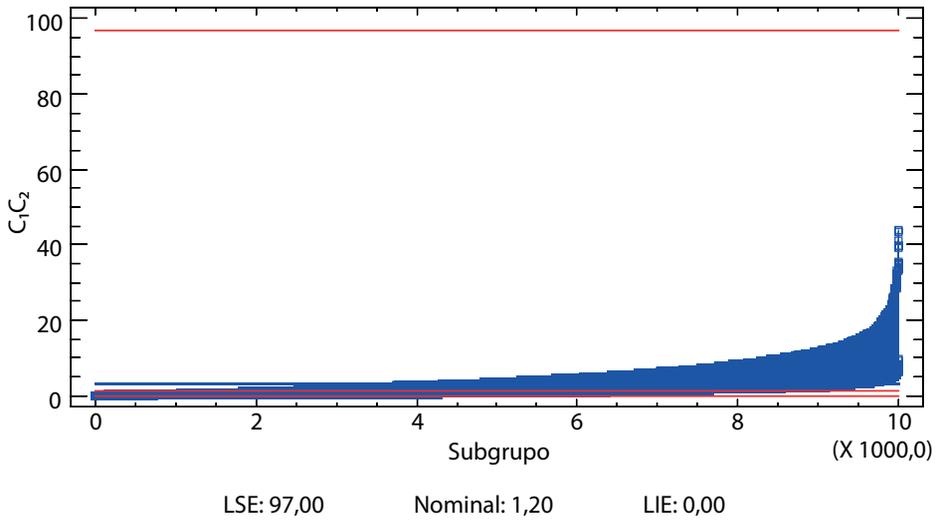


Figura 1. Gráficos de tolerancia para las variables C_1, C_2 . Caso variables correlacionadas. 10, 20, 50, 100, 1000 y 10 000

Elaboración propia

4.1.1.2 Caso de variables no correlacionadas

Para el estudio del caso de variables no correlacionadas, los resultados fueron los siguientes:

Tabla 4

Resumen descriptivo de las variables no normales C_1 y C_2 para diferentes tamaños de muestras

Muestra		n = 10			
Variable	LEI	LES	Objetivo	Media	Desv.Est.(General)
C_1	0,04	1,2	0,4	0,51	0,408
C_2	0,05	13,0	6,0	4,63	4,235
Muestra		n = 20			
Variable	LEI	LES	Objetivo	Media	Desv.Est.(General)
C_1	0,02	3,7	0,4	0,954	0,999
C_2	0,25	21,1	6,0	6,116	6,186

(continúa)

(continuación)

Muestra n = 50					
Variable	LEI	LES	Objetivo	Media	Desv.Est.(General)
C ₁	0,08	3,6	0,4	1,09	0,963
C ₂	0,02	19,9	6,0	4,71	4,431
Muestra n = 100					
Variable	LEI	LES	Objetivo	Media	Desv. Est. (General)
C ₁	0,01	7,2	0,4	1,04	1,164
C ₂	0,004	22,8	6,0	5,02	4,857
Muestra n = 1000					
Variable	LEI	LES	Objetivo	Media	Desv. Est. (General)
C ₁	0,001	7,0	0,4	1,032	1,041
C ₂	0,008	22,8	6,0	4,956	4,910
Muestra n = 10 000					
Variable	LEI	LES	Objetivo	Media	Desv. Est. (General)
C ₁	0,04	4,0	0,4	1,3333	0,5095
C ₂	0,16	3,2	1,7	0,5695	0,4781

Elaboración propia

Tabla 5
ICP univariado de las variables no normales C₁ y C₂ para diferentes tamaños de muestras

	Tamaño de muestra											
	n = 10		n = 20		n = 50		n = 100		n = 1000		n = 10 000	
Variables	C ₁	C ₂	C ₁	C ₂	C ₁	C ₂	C ₁	C ₂	C ₁	C ₂	C ₁	C ₂
Índice de capacidad												
C _p	0,42	0,38	0,58	0,52	0,62	0,70	0,88	0,66	1,03	0,7	1,28	1,15
P _{pl}	0,90	0,98	0,97	0,94	0,91	1,00	0,99	1,00	1,00	1,00	1,06	1,12
P _{pu}	0,34	0,32	0,49	0,47	0,57	0,66	0,88	0,63	1,03	0,67	1,43	1,18
P _{pk}	0,34	0,32	0,49	0,47	0,57	0,66	0,88	0,63	1,00	0,67	1,06	1,12
CC _p	1,00	0,99	0,98	1,00	0,98	1,00	2,00	1,14	1,00	0,7	1,40	1,38
CC _{pk}	0,77	0,99	0,97	1,00	0,98	1,00	1,09	1,00	1,00	0,66	1,04	1,12
Correlación de Pearson	- 0,22		- 0,174		- 0,229		- 0,055		- 0,025		0,003	

Elaboración propia

Ahora, para el caso multivariado, los *ICPM* fueron los siguientes:

Tabla 6
Cálculo de *ICPM* para diferentes tamaños de muestra respecto a la variable $C = C_1 C_2$ no normal

Índice de capacidad multivariado		n = 10	n = 20	n = 50	n = 100	n = 1000	n = 10 000
Cuamea y Rodríguez (2014)	C_{pM}	0,40	0,67	0,71	0,67	0,87	0,96
Statgraphics (2015)	MCP	0,23	0,35	0,44	0,60	0,79	1,00
Taam <i>et al.</i> (1993)	MC_p	1,1	1,00	1,03	1,00	1,00	1,00
Índice de esta investigación	$CPME$ (CC_p)	≠ 1	≠ 1	≠ 1	≠ 1	≠ 1	1

Elaboración propia

Los gráficos de tolerancia para algunos casos se observan en la figura 2.

Gráfico de tolerancia para C_1, C_2 . Variables no correlacionadas. n=10

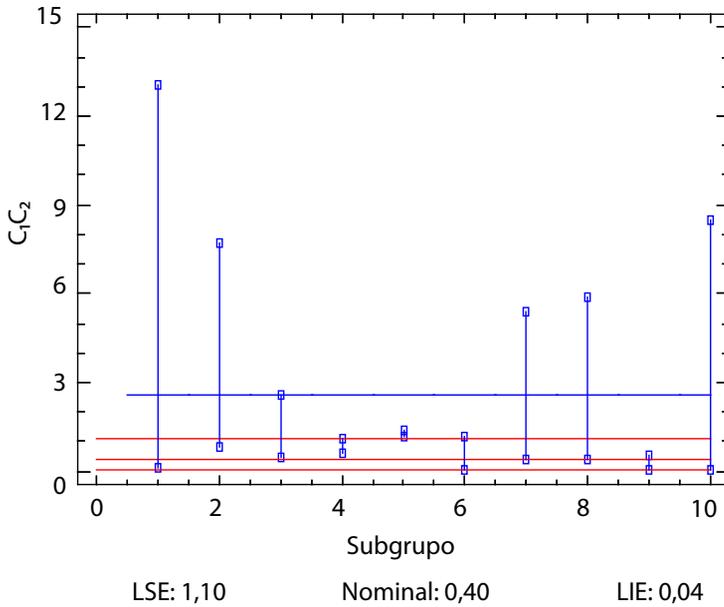


Gráfico de tolerancia para C_1, C_2 . Variables no correlacionadas. $n=20$

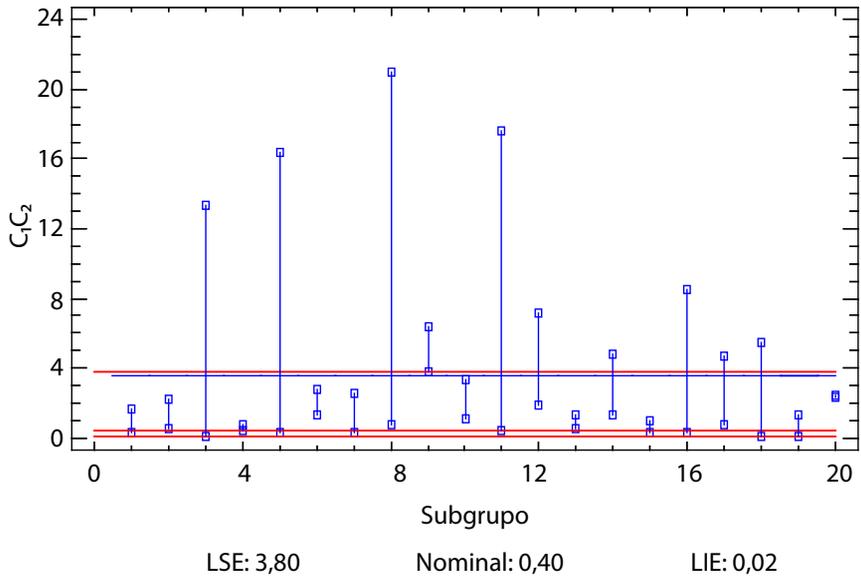


Gráfico de tolerancia para C_1, C_2 . Variables no correlacionadas. $n=50$

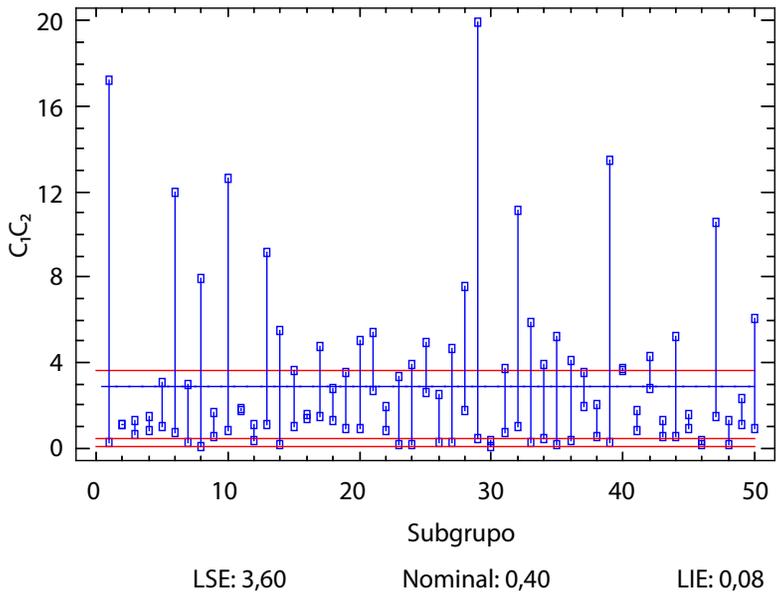


Gráfico de tolerancia para C_1, C_2 . Variables no correlacionadas. $n=100$

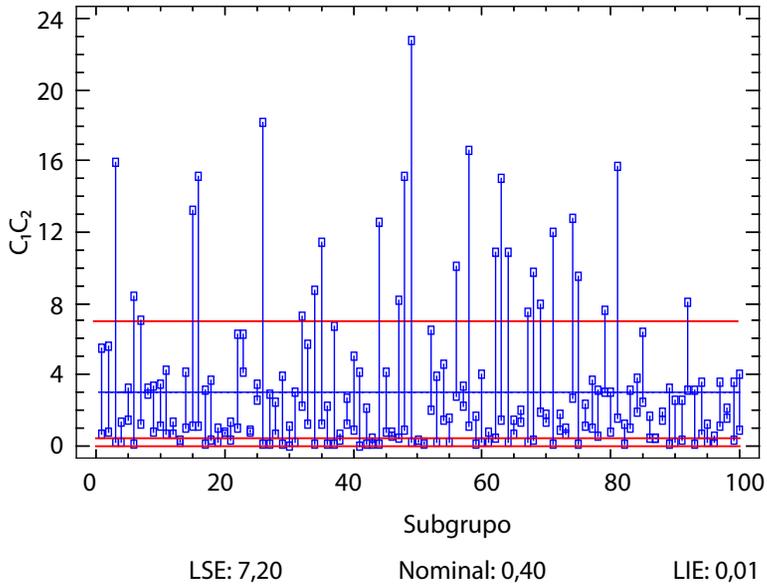


Gráfico de tolerancia para C_1, C_2 . Variables no correlacionadas. $n=1000$

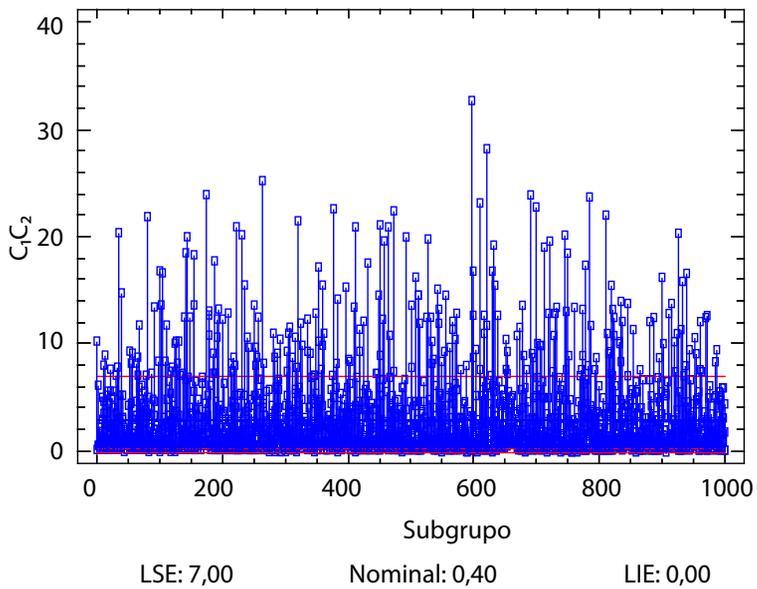


Gráfico de tolerancia para C_1, C_2 . Variables no correlacionadas. $n=10\ 000$

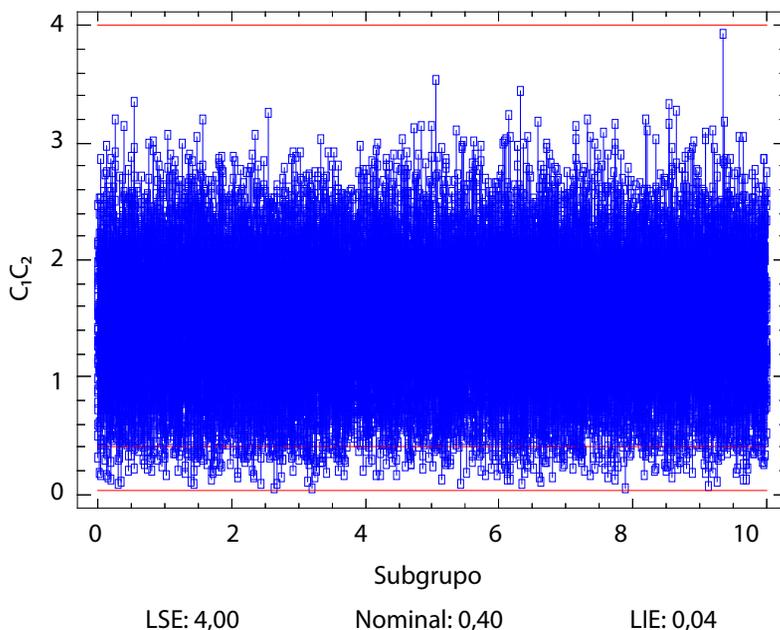


Figura 2. Gráficos de tolerancia para las variables C_1, C_2 . Caso variables no correlacionadas. 10, 20, 50, 100, 1000 y 10 000

Elaboración propia

El paquete estadístico Statgraphics emplea el $ICPM$, $MCP = \frac{Z}{K/2}$, donde Z es el valor de la variable aleatoria normal estándar correspondiente a los DPM calculados. El DPM es el índice de capacidad defecto por millón, que representa el número estimado de las muestras por millón que estarían fuera de la especificación en una o más variables. K es el múltiplo del sigma; en general, $K = 6$. Además, con el paquete estadístico Minitab 17, se calcularon los índices de capacidad de procesos para el caso no normal PP, PPI, PPu y Ppk . Otros índices calculados se determinaron usando el software Matlab V. 7.6.0, Excel y el paquete estadístico SPSS V. 15.0 para Windows.

Siendo así las cosas, la aplicación del índice $CPME$, para variables no normales, donde se presentaron casos para variables correlacionadas y no correlacionadas, resultó útil y reflejó la capacidad del proceso, tal como lo presentan los gráficos, ya que pocos índices logran evidenciar el proceso y algunos se distorsionan al aumentar el tamaño de la muestra. Hubo poca coincidencia del reflejo de la capacidad del proceso con otros índices, como lo fue con el índice MCP y el índice Cpm , en algunos casos de muestras pequeñas.

4.1.2 Caso normal

En Chen (1994) se discuten dos ejemplos numéricos, donde se aplicaron tres de los índices de capacidad que aparecen en la literatura. En este trabajo se usaron los datos de Chen (1994) y se trabajó con los casos de variables correlacionadas y no correlacionadas. Utilizando los datos del primer ejemplo, se calculó el índice de capacidad propuesto y se comparó con los índices obtenidos por las tres propuestas; el resumen se presenta en la tabla 7. El ejemplo trata de una distribución normal bivariada, en la cual la dureza Brinell (H) y la resistencia a la tensión (S) son dos variables de calidad de un producto industrial. Las tolerancias de ingeniería, para ambas variables de calidad, vienen dadas por [112,3; 241,3] y [32,7; 73,3], respectivamente; y el vector de valores nominales para H y S , por [177, 53].

4.1.2.1 Caso variables con correlación

Después de realizar la colección de 25 mediciones de ambas variables de calidad, se obtuvieron los valores de la tabla 7. Al realizar la prueba de Shapiro-Wilk, se encontró que las 25 mediciones seguían una distribución normal multivariada con el vector de medias [177,52; 52,188] y la siguiente matriz de varianzas covarianzas:

	H	S
H	329,093 (25)	91,8898 (25)
S	91,8898 (25)	38,8653 (25)

Tabla 7
Dureza Brinell (H) y resistencia a la tensión (S) para un producto industrial

H	S	H	S	H	S
143	31,0	141	47,3	178	50,9
200	57,0	175	57,3	196	57,9
168	47,5	187	58,5	160	45,5
181	53,4	187	58,2	183	53,9
148	47,8	186	57,0	179	51,2
178	51,5	172	49,4	194	57,5
162	45,9	182	57,2	181	55,6
215	59,1	177	50,6		
161	48,4	204	55,1		

Fuente: Cuamea y Rodríguez (2014)

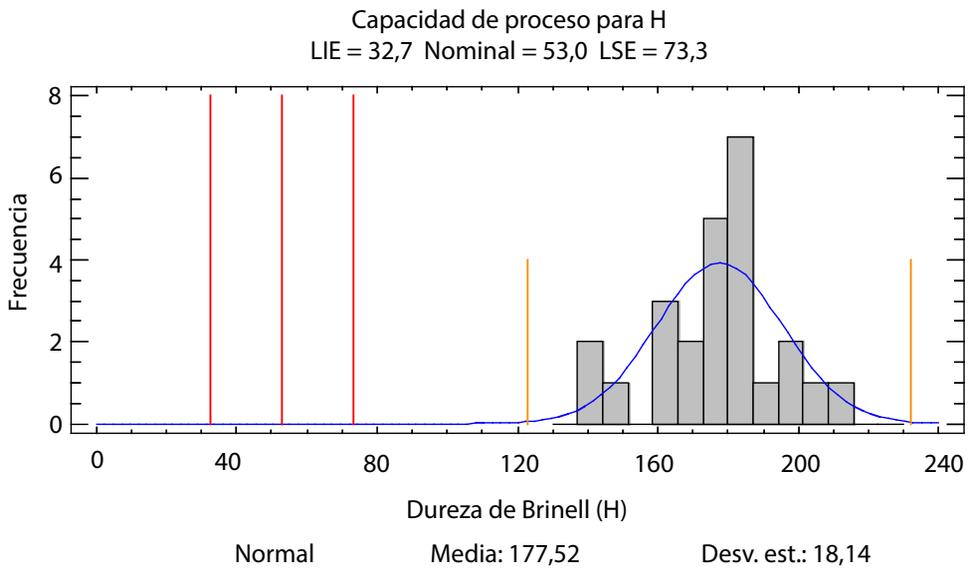
Con un $p > 0,150$, se concluye, según la prueba Kolmogorov-Smirnov, que H y S siguieron una distribución normal.

KS = 0,129 y KS = 0,140, para H y S , respectivamente.

Tabla 8
Índices de capacidad univariado para las variables H y S

Variable	Índice de capacidad	Capacidad (dentro de)	Capacidad general
H	C_p/P_p	1,06602	1,18517
	C_{pk}/P_{pk}	1,05412	1,17194
	C_{pm}		1,18
S	C_p/P_p	1,1211	1,08541
	C_{pk}/P_{pk}	1,07625	1,04199
	C_{pm}		1,16
Correlación de Pearson		0,813	

Elaboración propia



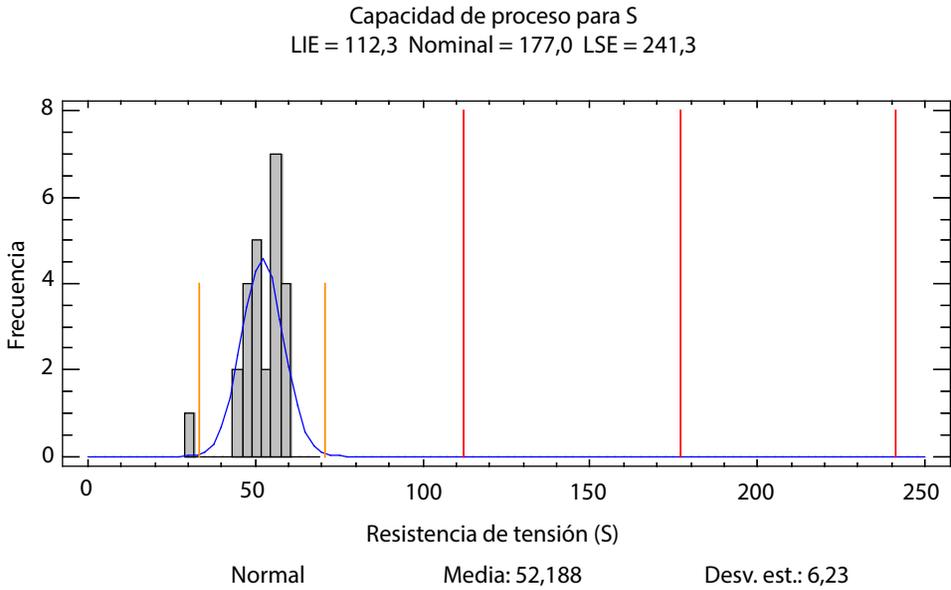


Figura 3. Gráficos de capacidad del proceso para la dureza (H) y la resistencia a la tensión (S). Caso variables correlacionadas

Elaboración propia

Para realizar el cálculo del índice de capacidad potencial, que se propone en este trabajo, se ubicó el promedio del proceso en el centro de las especificaciones.

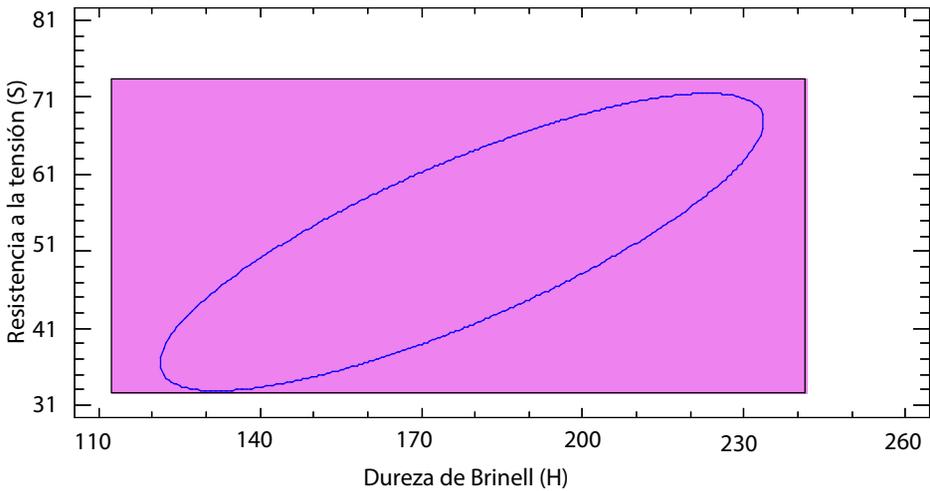


Figura 4. Elipse de capacidad del proceso para la dureza (H) (a) y la resistencia a la tensión (S). (b) Caso variables correlacionadas

Elaboración propia

Tomando $I_{cp_i} = C_p$, entonces:

$$CPME = \frac{\sum_{i=1}^p X_{A_i}(C_p)}{v} = \frac{0+0}{2} = 0, \text{ así } CPME(C_p) < 1.$$

De forma análoga, se obtuvieron los resultados que se muestran en la tabla 9.

En la figura 4 se observa que el proceso se salió ligeramente de la región de especificación, lo que estará reflejado por el cálculo del CPME, el cual para este proceso deberá tener un valor de este índice menor que 1.

Tabla 9
Comparación de varios índices de capacidad para 2 variables de calidad con correlación

Taam et al. (1993)		Índice de Cuamea y Rodríguez (2014)	Índice de Statgraphics (2016)	Índice CPME
\widehat{MC}_p	\widehat{MC}_{pM}	C_{pM}	MCP	$CPME(C_p)$ $CPME(P_p)$ $CPME(C_{pm})$
1,71	1,65	0,95	1,05	< 1

Elaboración propia

4.1.2.2 Caso variables no correlacionadas

Después de efectuar 25 mediciones de ambas variables de calidad, se obtuvieron los valores de la tabla 10. Al realizar la prueba de Shapiro-Wilk, se encontró que las 25 mediciones seguían una distribución normal multivariada con el vector de medias [177,106; 53,157] y la siguiente matriz de varianzas covarianzas:

	H	S
H	0,765983 (25)	- 0,0296604 (25)
S	- 0,0296604 (25)	1,09009 (25)

Tabla 10

Dureza Brinell (H) y resistencia a la tensión (S) para un producto industrial

H	S	H	S	H	S
176,437	53,5317	175,588	52,3967	176,818	54,1309
176,596	54,7648	176,807	52,1516	176,800	51,9326
176,092	51,7810	177,441	52,7447	176,051	54,0050
176,566	53,5707	176,926	52,0609	178,095	51,8588
176,940	53,1057	177,540	54,6906	178,384	53,9194
178,321	53,8484	177,634	53,5954	179,515	52,6633
177,161	52,4804	176,789	52,2339	177,645	52,7505
177,128	53,0898	175,812	54,4680		
177,209	55,3306	177,357	51,8202		

Elaboración propia

Con un $p > 0,150$, se concluyó, según la prueba Kolmogorov-Smirnov, que para H y S no había evidencia suficiente que sugiriera que los datos no siguen una distribución normal $KS = 0,109$ y $KS = 0,131$ para H y S , respectivamente.

Tabla 11

Índices de capacidad univariado para las variables H y S

Variable	Índice de capacidad	Capacidad (dentro de)	Capacidad general
H	C_p/P_p	24,33	24,37
	C_{pk}/P_{pk}	24,22	24,45
	C_{pm}		24,30
S	C_p/P_p	6,34	6,48
	C_{pk}/P_{pk}	6,29	6,43
	C_{pm}		6,41
Correlación de Pearson		- 0,033	

Elaboración propia

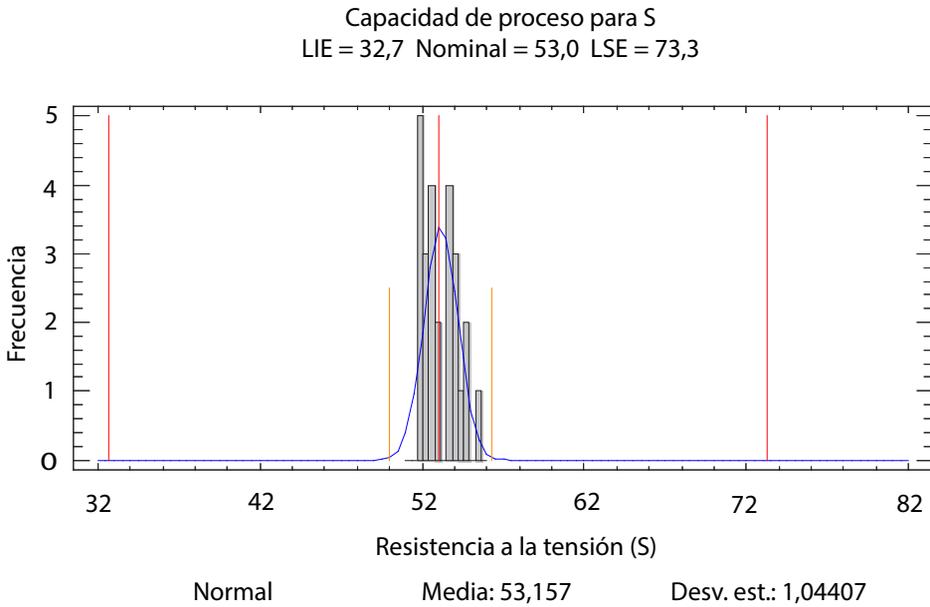
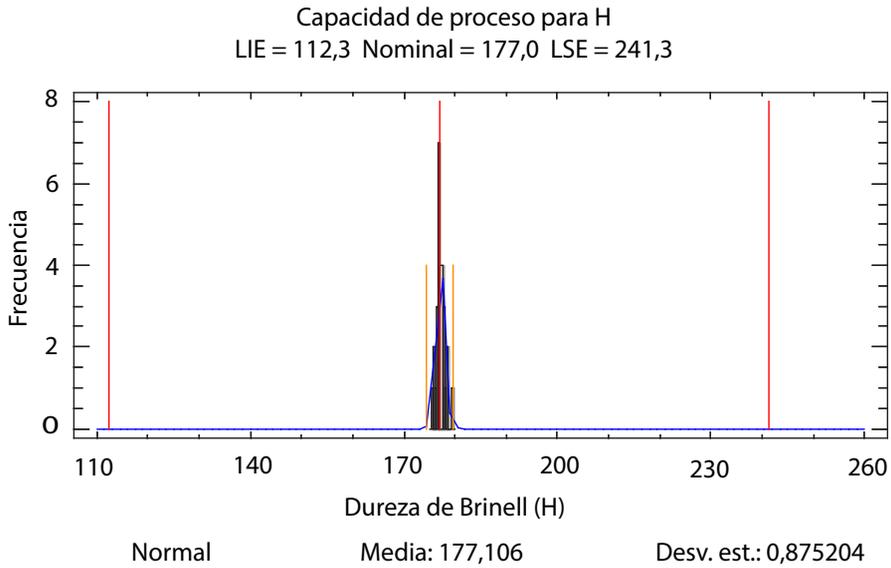


Figura 5. Gráficos de capacidad del proceso para la dureza (H) (c) y la resistencia a la tensión (S).
(d) Caso sin correlación

Elaboración propia

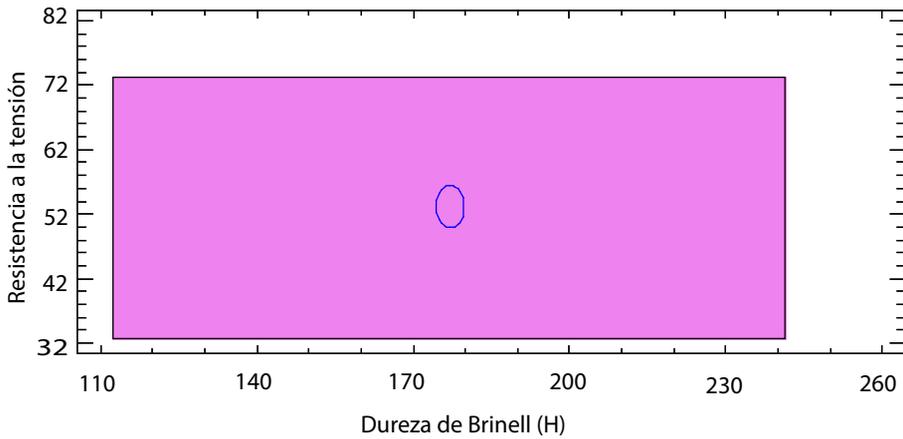


Figura 6. Elipse de capacidad del proceso para la dureza (H) y la resistencia a la tensión (S)
Elaboración propia

En la figura 6 se observa que el proceso está totalmente contenido en la región de especificación, lo que estará reflejado por el cálculo del $CPME$, el cual tendrá un valor de este índice igual que 1, para este proceso.

Tabla 12

Comparación de varios índices de capacidad para dos variables de calidad sin correlación

Taam <i>et al.</i> (1993)		Índice de Cuamea y Rodríguez (2014)	Índice de Statgraphics (2016)	Índice
\widehat{MC}_P	\widehat{MC}_{PM}	C_{pM}	MCP	$CPME (C_p)$ $CPME (P_p)$ $CPME (C_{pm})$
1,0	0,81	1,433	> 1,33	1

Elaboración propia

Así las cosas, la aplicación del índice $CPME$ para variables normal, donde se presentan casos para variables correlacionadas y no correlacionadas, resultó útil y evidenció la capacidad del proceso tal como lo presentan los gráficos, dado que pocos índices reflejan el proceso y algunos se distorsionan al aumentar el tamaño de la muestra. Hubo coincidencia del reflejo de la capacidad del proceso con otros índices, como lo fue con el índice MCP y el índice C_{pM} .

5. DISCUSIÓN

Para el estudio de variables no normal, en los casos estudiados de variables con correlación y sin correlación, estas variables pueden modelarse a través de una distribución de Weibull multivariada.

En el caso para variables correlacionadas, para los tamaños muestrales $n = 10, 20, 50$ y 100 , nuestro índice *CPME* coincidió con el índice de Cuamea y Rodríguez (2014), en cuanto a la indicación de que el proceso no era capaz; pero no hay coincidencia para los casos muestrales $n = 1000$ y $n = 10\ 000$. Sin embargo, el índice de Taam *et al.* (1993) indicó que el proceso era capaz para todos los casos muestrales y el índice de Statgraphics coincidió con nuestro índice, en cuanto a que el proceso no era capaz para todos los casos muestrales. En la figura 1 se ve que el proceso no era capaz para los diferentes casos muestrales, corroborando la información indicada por nuestro índice *CPME* y el índice obtenido por Statgraphics.

Para el caso de variables no correlacionadas, nuestro índice *CPME* mostró que, para los tamaños muestrales $n = 10, 20, 50, 100$ y 1000 , el proceso no era capaz, hecho que coincidió con los índices de Cuamea y Rodríguez (2014), Taam *et al.* (1993) y Statgraphics; sin embargo, para el caso de $n = 10\ 000$, solo nuestro índice *CPME* indicó que el proceso era capaz, hecho que no lo muestran los índices de Cuamea y Rodríguez (2014), Taam *et al.* (1993) y Statgraphics. Además, nuestro índice es respaldado por los gráficos de tolerancia de la figura 2, lo que indica su fortaleza.

Para el caso normal, también se estudiaron variables correlacionadas y no correlacionadas, las variables fueron modeladas a través de una distribución normal multivariada, y en ambos casos nuestro índice de capacidad *CPME* resultó con fortaleza. Su interpretación iba acorde con la situación presentada por los datos, además, fue aplicable para cualquier dimensión.

En el primer caso, la figura 4 muestra que el proceso se salió ligeramente de la región de especificación, reflejado por el índice *CPME*, el cual indicó que, en ese caso, la capacidad de proceso bivariable (H, S) no era estable; cuestión que no es reflejada por los índices de Taam *et al.* (1993). Sin embargo, para el siguiente caso, la figura 6 muestra que el proceso está totalmente contenido en la región de especificación, lo cual es reflejado por el cálculo del *CPME* el que indica que la capacidad de proceso bivariable (H, S) es estable, cuestión que es reflejada por los demás índices.

Como se observa, el uso de gráficos y regiones de especificación para procesos multivariados resultan de gran ayuda, ya que facilitan la comprensión de los índices de capacidad de procesos y, en particular, el índice *CPME* expresa mejor la situación del proceso que los demás índices.

La metodología del cálculo del índice *CPME* es fácil y tiene alta efectividad, pues puede ser aplicado para los casos donde la distribución de datos es normal y no normal; y para variables con correlación (caso más complicado, ya que los otros *ICPM* fueron distorsionados por este caso) y sin correlación, obteniendo resultados semejantes a otros índices. Vale la pena anotar, que para los casos en que hay discrepancia, el gráfico de la capacidad de proceso respalda nuestro índice, mostrando su fortaleza y efectividad. Así, el índice *CPME* viene a constituir un nuevo índice multivariado de aplicación más amplia que los demás índices multivariados, pues estos son aplicados a casos particulares.

REFERENCIAS

- Box, G., y Cox, D. R. (1964). An analysis of transformation. *Journal of the Royal Statistical Society: Series B*, 26(2), 211-243. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
- Castagliola, P., y Castellanos, J. V. G. (2005). Capability indices dedicated to the two quality characteristics case. *Quality Technology and Quantitative Management*, 2(2), 201-220. <https://doi.org/10.1080/16843703.2005.11673094>
- Castagliola, P., y Castellanos, J. V. G. (2008). Process capability indices dedicated to bivariate non-normal distributions. *Journal of Quality in Maintenance Engineering*, 14(1), 87-101. <https://doi.org/10.1108/13552510810861969>
- Chen, H. (1994). A multivariate process capability index over a rectangular solid tolerance zone. *Statistica Sinica*, 4(2), 749-758.
- Chou, Y., Polansky, A., y Mason, R. (1998). Transforming non-normal data to normality in statistical process control. *Journal of Quality Technology*, 30(2), 133-141. <https://doi.org/10.1080/00224065.1998.11979832>
- Cuamea, G., y Anaya, C. (2009). Evaluación de la calidad en productos o procesos con múltiples características de calidad correlacionadas. *Epistemos*, 6, 12-16.
- Cuamea, G., y Rodríguez, M. (2014). Propuesta para evaluar la capacidad de procesos de manufactura multivariados. *Revista Ingeniería Industrial*, 13(2), 35-47.
- Foster, E. J., Barton, R. R., y Gautam, N. (2005). The process-oriented multivariate capability index. *International Journal of Production Research*, 43(10), 2135-2148. <https://doi.org/10.1080/00207540412331530158>
- Hubele, N., Shahiari, H., y Cheng, C. (1991). A bivariate capability vector. En J. B. Keats y D. C. Montgomery (Eds.), *Statistics and design in process control: Keeping pace with automated manufacturing* (pp. 229-310). Marcel Dekker.
- Johnson, N. L. (1949). System of frequency curves generated by methods of translation. *Biometrika*, 36(2), 149-176. doi:10.2307/2332539

- Liu, P. H. y Chen, F. L. (2006). Process capability analysis of non-normal process data using the Burr XII distribution. *International Journal of Advanced Manufacturing Technology*, 27(9), 975-984. <https://doi.org/10.1007/s00170-004-2263-8>
- Montgomery, D. (2009). *Statistical quality control. A modern introduction* (6.ª ed.). John Wiley and Sons, Inc.
- Polansky, A. M., Chou, Y. M., y Mason, R. L. (1998). Estimating process capability indices for a truncated distribution. *Quality Engineering*, 11(2), 257-265. <https://doi.org/10.1080/08982119808919237>
- Salazar, E., y Fermín, J. (2016). Un índice de capacidad de procesos para distribuciones multivariadas normales de variables correlacionadas y no correlacionadas. *Ingeniería Industrial*, (34), 57-73. doi:10.26439/ing.ind2016.n034.1337
- Salazar, E., y Fermín, J. (2017). Un índice de capacidad de procesos para distribuciones multivariadas no normales de variables correlacionadas y no correlacionadas. *Ingeniería Industrial*, (35), 55-75. doi:10.26439/ing.ind2017.n035.1790
- Shahriari, H., y Abdollahzadeh, M. (2009). A new multivariate process capability vector. *Quality Engineering*, 21(3), 290-299. <https://doi.org/10.1080/08982110902873605>
- Shinde, R. L., y Khadse, K. G. (2009). Multivariate process capability using principal component analysis. *Quality and Reliability Engineering International*, 25(1), 69-77. <https://doi.org/10.1002/qre.954>
- Somerville, S., y Montgomery, D. (1996). Process capability indices and non-normal distributions. *Quality Engineering*, 9(2), 305-316. <https://doi.org/10.1080/08982119608919047>
- Statgraphics (16.1.15) [Software] (2015). Recuperado de <https://statgraphics.net/>
- Taam, W., Subbaiah, P., y Liddy, J. W. (1993). A note of multivariate capability indices. *Journal of Applied Statistics*, 20(3), 339-351. <https://doi.org/10.1080/02664769300000035>
- Tang, L. C., y Than, S. E. (1999). Computing process capability indices for non-normal data: a review and comparative study. *Quality and Reliability Engineering International*, 15(5), 339-353. [https://doi.org/10.1002/\(SICI\)1099-1638\(199909/10\)15:5<339::AID-QRE259>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1099-1638(199909/10)15:5<339::AID-QRE259>3.0.CO;2-A)
- Wang, F. K. (2006). Quality evaluations of a manufactured Product with multiple characteristics. *Quality and Reliability Engineering International*, 22(2), 225-236. doi:<https://doi.org/10.1002/qre.712>
- Wang, F. K., y Du, T. C. (2000). Using principal component analysis in process performance for multivariate data. *Omega*, 28(2), 185-194. [https://doi.org/10.1016/S0305-0483\(99\)00036-5](https://doi.org/10.1016/S0305-0483(99)00036-5)

- Wang, F. K., Hubele, N. F., Lawrence, F. P., Miskulin, J. D., y Shahriari, H. (2000). Comparison for three multivariate process capability indices. *Journal of Quality Technology*, 32(3), 263-275. <https://doi.org/10.1080/00224065.2000.11980002>
- Yeh, A., y Chen, H. (2001). A nonparametric multivariate process capability index. *International Journal of Modelling and Simulation*, 21(3), 218-223. <https://doi.org/10.1080/02286203.2001.11442205>

