Ride Revista Iberoamericana para la Investigación y el Desarrollo Educativo
ISSN 2007 - 7467

https://doi.org/10.23913/ride.v13i25.1322

Artículos científicos

Construcción de índices de capacidad para el análisis y evaluación de procesos con múltiples respuestas

Construction of Capacity Indices for the Analysis and Evaluation of Processes with Multiple Responses

Construção de índices de capacidade para análise e avaliação de processos com múltiplas respostas

Guillermo Cuamea Cruz
Universidad de Sonora, México
guillermo.cuamea@unison.mx
https://orcid.org/0000-0001-8884-6825

Manuel Arnoldo Rodríguez Medina
Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Juárez, México
manuel_rodriguez_itcj@yahoo.com
https://orcid.org/0000-0003-1676-0664

Eduardo Rafael Poblano Ojinaga
Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Juárez, México
pooe_65@hotmail.com
https://orcid.org/0000-0003-3482-7252

Rafael García Martínez
Tecnológico Nacional de México, Instituto Tecnológico de Hermosillo, México
gamrmx@yahoo.com
https://orcid.org/0000-0001-7175-5361

Diego Adiel Sandoval Chávez
Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Juárez, México
dsandoval@itcj.edu.mx
https://orcid.org/0000-0002-2536-1844

Resumen

Los productos fabricados actualmente tienen varias características de calidad, todas tienen importancia para el cliente, y controlarlas y evaluarlas se ha convertido en una actividad de primer interés. La industria automotriz establece índices de capacidad para evaluar la capacidad de los procesos donde se tiene únicamente una variable de respuesta, y cuando estos son estables se recomienda utilizar el C_p y el C_{pk} como medibles de la capacidad del proceso de manufactura para manufacturar productos que cumplan con las especificaciones y sean catalogados como productos de calidad. Para evaluar la calidad completa de un producto, la cual depende de que se cumpla con varias características de calidad simultáneamente, existen propuestas en la literatura de cómo medir la capacidad de procesos multivariados; la mayoría de estas coinciden en que se debe establecer claramente una región de especificación que represente lo que establece el cliente y otra región que muestre la variación como medida de desempeño del proceso. En la definición de ambas regiones existe una gran controversia entre los autores, lo que lleva a presentar esta nueva propuesta, mediante la cual se definen de manera confiable estas dos regiones mencionadas, y al hacer una comparación entre ellas se pueden obtener los índices de capacidad multivariados C_{pm} y el C_{pkm} y el como una extensión de los índices univariados C_p y el C_{pk}. El documento incluye el análisis de datos de un proceso donde el producto, para ser de buena calidad, debe cumplir con dos características de calidad simultáneamente. Las mediciones obtenidas del proceso se pueden representar por una distribución normal multivariada, lo que permite medir la capacidad del proceso utilizando los índices propuestos y realizar una interpretación de estos en relación con el desempeño del proceso.

Palabras clave: capacidad del proceso, índices de capacidad, procesos multivariados, regiones de especificación.

Abstract

Currently manufactured products have several quality characteristics, all of which are important to the customer, and controlling and evaluating them has become an activity of prime interest. The automotive industry establishes capacity indices to evaluate the capacity of processes where there is only one response variable, and when these are stable, it recommends using *Cp* and *Cpk* as measures of the capacity of the manufacturing process to manufacture products that comply with specifications and are cataloged as quality products. To evaluate the complete quality of a product, which depends on meeting several quality characteristics simultaneously, there are proposals in the literature on how to measure the capability of multivariate processes; most of these agree that a specification region should be clearly established to represent the customer's requirements and another region to show the variation as a measure of process performance. In the definition of both regions there is a great controversy among the authors, which leads to present this new proposal, by which these two regions are defined in a reliable way, and by making a comparison between them, the multivariate capacity indexes C_{pm} and C_{pkm} can be obtained as an extension of the univariate indexes C_p and C_{pk}. The document includes the data analysis of a process where the product to be of good quality must meet two quality characteristics simultaneously. The measurements obtained from the process can be represented by a multivariate normal distribution, which allows measuring the process capability using the proposed indexes and interpreting them in relation to the process performance.

Keywords: process capability, capability indices, multivariate processes, specification regions.

Resumo

Os produtos fabricados atualmente possuem diversas características de qualidade, todas importantes para o cliente, e controlá-los e avaliá-los tornou-se uma atividade de grande interesse. A indústria automotiva estabelece índices de capacidade para avaliar a capacidade de processos onde há apenas uma variável de resposta, e quando estes estão estáveis, recomenda o uso de *Cp* e *Cpk* como medidas da capacidade do processo de fabricação em fabricar produtos que atendam às especificações e sejam catalogados como produtos de qualidade. Para avaliar a qualidade completa de um produto, que depende do cumprimento de várias características de qualidade simultaneamente, existem propostas na literatura sobre como medir a capacidade de processos multivariados; a maioria concorda que uma definição clara deve ser estabelecida. que representa o que o cliente estabelece e outra região que mostra a variação como medida de desempenho do processo. Na definição de ambas as regiões há grande controvérsia entre os autores, o que leva a apresentar esta nova proposta, pela qual estas duas regiões são definidas de forma fiável, e fazendo uma comparação entre elas é possível obter os índices de capacidade multivariados C_{pm} e C_{pkm} como uma extensão dos índices univariados C_p e C_{pk}. O documento inclui a análise de dados de um processo onde o produto para ser de boa qualidade deve atender a duas características de qualidade simultaneamente. As medições obtidas a partir do processo podem ser representadas por uma distribuição normal multivariada, permitindo que a capacidade do processo seja medida utilizando os índices propostos e uma interpretação destes índices em relação ao desempenho do processo.

Palavras-chave: capacidade de processo, índices de capacidade, processos multivariados, regiões de especificação.

Fecha Recepción: Enero 2022 **Fecha Aceptación:** Octubre 2022

Introduction

International globalization or internationalization, the increased competitiveness of companies, productivity growth can be achieved through the efficient application and innovation of existing technologies (Prokopenko, 1998). At the same time, Prokopenko himself (1998) mentions that a trained workforce with a positive attitude towards change and learning new concepts increases the productivity and competitiveness of companies and the country.

The increase in modern technology has caused that the demand on products and processes has also increased, and that the controls and evaluations on these are stricter and of a different nature, since products with characteristics that satisfy the needs are needed. and customer expectations (Poblano, Sanchez, Rodriguez, Valles, and Gonzalez, 2020; Rodriguez, Poblano, Rodriguez, and Alvarado, 2021). Thus, the controls on the process and the preventive controls must be carried out with greater knowledge of the occurrence of failures due to the existence of the causes that provoke them. Design failure modes must be more widely understood to prevent process failure modes from altering the quality and reliability of finished products. Therefore, it must be clear whether the current manufacturing process has the ability or capacity to meet the product design goals, while meeting quality and reliability goals.

The initial process studies, considered from the normative point of view, are carried out to verify if the level of initial capacity or performance of the process is acceptable. In this regard, quality or performance indices should be considered, depending on the behavior of the process, this only in the case of processes with a univariate response variable. In addition, it is considered that, if the process is stable, the indices should be the C_p and C_{pk}, whose names are potential capacity indices and actual capacity index, respectively; here $\hat{\sigma} = \frac{\bar{R}}{d_2}$ o $\hat{\sigma} = \frac{\bar{S}}{C_4}$. In the case of unstable processes, the indices are called performance indices, citing as P_p y P_{pk}, where the estimation of the variation must be made with all the measured data (at least 100 from the normative point of view and under statistical rigor). Juran (1974) was one of the first to present a formal way to measure quality using the C_p, which is a ratio of variations, the variation allowed by the client in the allowed tolerance and the variation of the process considered the difference between the upper and lower natural tolerance limits; that is:

$$C_p = \frac{LSE - LIE}{LTNS - LTNI} = \frac{Tolerancia}{6\sigma}$$

It should be noted that the Cp does not consider in its calculation where the process is located in relation to the specifications, so it restricts the analysis to verify if the variation of the process is adequate to the variation required by the client.

Kotz and Johnson (2002) simplify the notation by representing LSE as U and LEL as L. Furthermore, for univariate measurements, the variable will be represented by X and its expected value and variance will be given by E(X) and Var (X), respectively. So:

$$Cp = \frac{U - L}{6\sigma} = \frac{d}{3\sigma}$$

It should be noted that in the above formula $d = \frac{U-L}{2}$. Now, considering $M = \frac{U+L}{2}$, and being a specification-centric process, the distances from any of the specifications to the nominal value take the same value, that is, $M - LIE = LSE - M$. If we consider that the process can be modeled as a normal distribution in which the assumption is made that the mean of the process is the nominal value of the specifications, and without altering the original variation of the process, the algebraic operations can be carried out following, as suggested by (Cuamea & Rodriguez, 2014), which does not alter the value that would be obtained with the original formula:

$$C_p = \frac{LSE - LIE}{LTNS - LTNI} = \frac{Tolerancia}{6\sigma}$$

Well then:

$$C_p = \frac{(LSE - LIE)/2}{3\sigma}$$

which can be written as:

$$C_p = \frac{M - LIE}{3\sigma} = \frac{LSE - M}{3\sigma}$$

and it can be seen that:

$$C_p = \frac{M - LIE}{3\sigma} = \frac{Z_c}{3} = \sqrt{\frac{Z_c^2}{9}} = \sqrt{\frac{\chi_{c,1}^2}{9}}$$

$$C_p = \frac{LSE - M}{3\sigma} = \frac{Z_c}{3} = \sqrt{\frac{Z_c^2}{9}} = \sqrt{\frac{\chi_{c,1}^2}{9}}$$

The proposed equations provide the same result of the capability of a process, which has only one critical or important quality characteristic in the sense that for any standard normal distribution one can associate a χ^2 with degrees of freedom equal to 1; On the other hand, the value of 9 corresponds to a χ^2 with a degree of freedom that covers a density of 0.9730, as described (Cuamea & Rodriguez, 2014). In this way, a more convenient way of calculating the *Cp* in this research, which is going to be taken as a basis to be able to extend it to multivariate processes using the Mahalanobis distance (it is a measure of distance that takes into account the correlation that could exist between two variables and is used to determine the similarity between two variables), as specified (Peña, 2002). The probability distribution function associated with the Mahalanobis distance corresponds to a chi-square with degrees of freedom equal to the number of variables or critical quality characteristics of the process (Peña, 2002).

Kane (1986) proposed an index to determine if a given process has the capacity to produce good quality products and a characteristic of this proposal is that the location of the process is taken into account in its calculation. This index is given by:

$$C_{pk} = min\left\{\frac{LSE - \mu}{3\sigma}, \frac{\mu - LIE}{3\sigma}\right\}$$

The C_{pk} analyzes the location of the process, but without taking into account its objective value (T). Using the notation given by Kotz and Johnson (2002), the C_{pk} will be:

$$C_{pk} = \frac{d - |U - M|}{3\sigma} = \frac{min\{U - \mu, \mu - L\}}{3\sigma}$$

The index C_{pm} was then created to correct that problem, and is calculated by:

$$C_{pm} = \frac{LSE - LIE}{6\sqrt{\sigma^2 + (\mu - T)^2}}$$

The modification of C_{pm} it will be as follows:

$$C_{pm} = \frac{d}{3\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{d}{3\sqrt{E[(X - T)^2]}}$$

The C_{pm} It is called the Taguchi index by its creator, Genichi Taguchi (1985).

Pearn, Kotz and Johnson (1992) developed the index C_{pmk} , which can be used for those situations where the target value (T) is not within the specifications. This is a combination of C_{pk} y el C_{pm}, and the way to get it is as follows:

$$C_{pmk} = min\left[\frac{LSE - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LIE}{3\sqrt{\sigma^2 + (\mu - T)^2}}\right]$$

The modification of C_{pmk} shown below as described by Wu, Pearn and Kotz (2009), which is given as:

$$C_{pmk} = \frac{d - |\mu - M|}{3\sqrt{E[(X - T)^2]}}$$

Kotz and Johnson (2002) show the relationship between the different indices. It can clearly be established that:

$$C_p \geq C_{pk} \geq C_{pmk} \text{ y } C_p \geq C_{pm} \geq C_{pmk}.$$

And the relationship between C_{pk} y C_{pm} shows up right away:

$$C_{pk} = C_p - \frac{1}{3}\left|\frac{\mu - M}{\sigma}\right|$$

$$C_{pm} = \frac{C_p}{\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}}$$

By the way, yes $T = M$, the relationship between C_p, C_{pk}, C_{pm} Will be:

$$C_{pk} = C_p - \sqrt{\left(\frac{C_p}{C_{pm}}\right)^2 - 1}$$

Problem Statement

For those processes where the quality of their products depends on the simultaneous fulfillment of several quality characteristics, some measures of process capacity have been proposed that can be classified as extensions of their univariate counterparts. However, as established by Foster, Barton, Gautam, Truss and Tew (2005), there is still no consistent methodology to calculate multivariate capacity indices, so to date there is no consensus on the use of a particular index. The proposal of this work is to modify the equations to obtain the capacity indices proposed by Cuamea and Rodriguez (2014), considering that the population parameters of the multivariate normal distribution, such as the mean vector and the variance-covariance matrix, in practical applications will not be known, since there is only one sample of observations for two or more quality characteristics of the process, in such a way that with this sample the quality of a process can be evaluated by relating the engineering specification region with a region of natural variation of a multivariate process. Said evaluation will result in the calculation of capacity indices that represent the performance of the process in a consistent manner, taking into account the information contained in each of the quality variables or characteristics (Cuamea & Rodriguez, 2014). From the above it can be seen that it is important to consider whether or not the variables are correlated before measuring the process capability, but this can be complex when making the calculations and interpreting the results, since the specification limits for each characteristic of quality cannot be represented simply as two vertical lines that resemble a goal; in this case, on the one hand, they must be represented by hyperrectangles, and on the other, the region of natural variation of the process will have to be represented by ellipsoids.

Research work focused on defining capability indices for multivariate processes began in the 1990s and to date different authors have released several proposals. These proposals can be divided into four different groups, as suggested first (Khadse & Shinde, 2009) and later (de-Felipe & Benedito, 2017). These groups are:

- Group 1: capacity indices that are obtained from the ratio between the tolerance region and the variation region of the process, one of these proposals is the one presented by (Taam et al., 1993) (Pana & Lee, 2010).

- Group 2: these indices are based on the probability of non-conforming product, such as the index proposed by Wierda (1994), (Bothe, 1999), (Castagliola & Castellanos, 2008) and (de-Felipe & Benedito, 2017).
- Group 3: those indices based on principal component analysis are grouped here. One of the most cited works in this group is the proposal they made (Wang & Chen, 1998).
- Group 4, called others, in this group is the proposal of Shahriari, Hubele and Lawrence (1995) and Barreto and Herrera (2021).

A summary of the proposals presented by the different authors is found in table 1 and it is easy to see, looking at column three of the table, that most of the proposed indices require a multivariate normal distribution, and in column four of The table shows that most of them belong to the first group.

Table 1. Capacity indices found in the literature in the period 1993-2021

Autor	Índice	DNM	Grupo
Taam *et al*. (1993)	MC_{pm}	Sí	1
Chen (1994)	MC_p	Sí	1
Shahriari *et al*. (1995)	*MPCV*	Sí	4
Wang y Chen (1998)	MC_p, MC_{pk}, MC_{pm} y MC_{pmk}	Sí	1 y 3
Wang y Du (2000)	MC_p y MC_{pc}	Sí	1 y 3
Yeh y Chen (2001)	MC_f	No	2
Castagliola *et al*. (2005)	BC_p y MC_{pk}	Sí	2
Wang (2005)	MC_p y MC_{pk}	Sí	1 y 3
Wang (2006)	MC_{pc}	No	4
Pearn *et al*. (2007)	MC_p	Sí	1
Castagliola *et al*. (2008)	BC_p y BC_{pk}	No	2
Shahriari *et al*. (2009)	*NMPCV*	Sí	1
Ahmad *et al*. (2009)	PNC_{Total}	No	2
González y Sánchez (2009)	C_n^s	Sí	2 y 3
Shinde y Khadse (2009)	$Mp1$ y $Mp2$	Sí	2 y 3
Pan y Lee (2010)	NMC_{pm}	Sí	1
Goethals y Cho (2010)	MC_{pmc}	Sí	4
Dharmasenaa *et al*. (2016)	$T\ Spk,PC;\beta$	Sí	2 y 3
Cuamea y Rodriguez (2014)	C_{pM} , C_{pkM}	Si	1
Zainab *et al*. (2019)	$MC''\ p(u, v)$	Sí	1
Barreto y Herrera (2021)	*MCpCR*	Sí	3 y 4

Source: (Shinde & Khadse, 2009) and (Barreto & Acosta Roberto, 2021)

Méthod

The method used for the development of multivariate capacity indices mainly uses the theoretical bases of the multivariate normal distribution, as well as the properties associated with said distribution.

Basis for obtaining the index C_{pM}

Peña (2002) describes the scalar normal distribution as a function of density:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Thus, described as $x \sim N(\mu, \sigma^2)$, states that x has a normal distribution with mean and variance $\mu\ y\ \sigma^2$, respectively.

Generalizing the previous function, it can be said that a vector x follows a p-dimensional normal distribution if its density function is:

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{|\Sigma|}} (2\pi)^{-p/2} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})`\Sigma^{-1}(\boldsymbol{x}-\mu)}$$

In this case, μ is the mean of the normal random vector and Σ is the matrix of variances and covariances.

A vector x with p-dimensional normal distribution with matrix Σ nonsingular can be converted to a p-dimensional normal z vector with mean 0 and variance-covariance matrix equal to the identity matrix I. The density function of $\boldsymbol{z}$ will be:

$$f(\boldsymbol{z}) = \frac{1}{(2\pi)^{p/2}} e^{-\frac{1}{2}\boldsymbol{z}`\boldsymbol{z}} = \prod_{i=1}^{p} \frac{1}{(2\pi)^{1/2}} e^{-\frac{1}{2}\boldsymbol{z}_i^2}$$

The multivariate normal distribution has a very important basic property whereby every level curve in the distribution is an ellipsoid with a confidence value $(1-\alpha)$, la cual tiene una distribución χ^2 with degrees of freedom equal to the number of variables (Peña, 2002). Thus, for any value of one of the random variables, an ellipsoid can be associated with confidence $(1-\beta_i)$. The calculation of C_{pM} proposed by (Cuamea & Rodriguez, 2014) is given by the following equation:

$$C_{pM} = min\left\{\sqrt{\frac{\chi^2_{\beta_i}}{\chi^2_{0.0027,m}}}\ |i = 1,2, \dots, m\right\}$$

For each of the specifications, two values of $\chi^2_{\beta_i}$, where each value is obtained by relating each specification limit to its nominal value, as shown below for the first variable:

$$\boldsymbol{\chi^2_{\beta_1,inf}} = \frac{(LEI_1 - \mu_1)^2}{|\Sigma||\Sigma_1^{-1}|}, \boldsymbol{\chi^2_{\beta_{1,sup}}} = \frac{(LES_1 - \mu_1)^2}{|\Sigma||\Sigma_1^{-1}|}$$

Over there μ_1 is the nominal value of the specifications.

If the standard deviations of all the variables considered in the process are modified by $\breve{\sigma}_i = C_{pM}\sigma_i$, the process will potentially be able to meet all specifications simultaneously. If the transformed values are substituted into the variance-covariance matrix Σ, values on the main diagonal will take the form:

$$C_{pM}^2\sigma_i^2$$

And the values outside the main diagonal, that is, the covariances in row i and column j will be:

$$C_{pM}^2\sigma_i\sigma_j$$

And the new matrix of variances and covariances will be $\breve{\Sigma} = C_{pM}^2\Sigma$, fulfilling that $|\breve{\Sigma}| = C_{pM}^{2m}|\Sigma|$. The inverse of the new matrix of variances and covariances is calculated by:

$$|\breve{\Sigma}^{-1}| = \frac{1}{|\breve{\Sigma}|} = \frac{1}{C_{pM}^{2m}|\Sigma|}$$

The upper and lower natural tolerance limits of quality characteristic i are:

$$LTN_i = \mu_{ci} \pm \sqrt{|\breve{\Sigma}||\breve{\Sigma}_i^{-1}|\chi_{0.0027,m}^2}$$

Substituting the values, we have:

$$LTN_i = \mu_{c_i} \pm \sqrt{C_{pM}^2|\Sigma||\Sigma_1^{-1}|\chi_{0.0027,m}^2} = \mu_{c_i} \pm \sqrt{\frac{\chi_{\beta_i}^2|\Sigma||\breve{\Sigma}_i^{-1}|\chi_{0.0027,m}^2}{\chi_{0.0027,m}^2}}$$

And simplifying:

$$LTN_i = \mu_{c_i} \pm \sqrt{|\Sigma||\Sigma_1^{-1}|\chi_{\beta_i}^2}$$

Obtaining the index C_{pkM}

The C_{pkM} proposed in the work of (Cuamea & Rodriguez, 2014) is an index that must take into account the current location of the process, similar to the univariate case. The equation for the calculation will be:

$$\boldsymbol{C_{pkM} = min\left\{\sqrt{\frac{\chi_{\beta_i}^2}{\chi_{0.0027,m}^2}}\right\} para\ i = 1, 2, \dots, m}$$

As in the case of C_{pM}, for each variable we must obtain two values of $\chi_{\beta_i}^2$, associated with the lower and upper specifications of each of the variables, for example, for variable 1, we will use the following notation:

$$\chi^2_{\beta_{1,inf}} = \frac{(LEI_1 - \mu_1)^2}{|\Sigma||\Sigma_1^{-1}|}, \chi^2_{\beta_{1,sup}} = \frac{(LES_1 - \mu_1)^2}{|\Sigma||\Sigma_1^{-1}|}$$

Where the values of μ_i are the real averages of the respective variables. For each variable with a lower and a higher specification, two possible values must be obtained for the C_{pkM} and it is common that the processes are not centered, and from here $C_{pkM} < C_{pM}$, and as $\mu \longrightarrow \mu_c$, as μ_c is the nominal value of the specifications, then $C_{pkM} \longrightarrow C_{pM}$.

Decisions about the process according to the indices

Remembering that the C_{pkM} is an index that considers the location of the process and the variation contributed by the variables involved, having a value of $C_{pkM} > 1$ indicates that the process is capable of meeting specifications; otherwise, the process requires adjustments. If $C_{pkM} < C_{pM}$, the required process adjustments should be to move the mean toward the center of the specifications. If the $C_{pM} < 1$, Mainly the adjustments in the process should focus on reducing its variability.

The previous development is the one proposed by (Cuamea & Rodriguez, 2014), which requires that the vector of means of the multivariate normal distribution be known, as well as the variance-covariance matrix of said distribution. The proposal presented in this work is based on the work of (Cuamea & Rodriguez, 2014), yes, but it takes into consideration that in practice the real values of the variances are not known and, therefore, neither is the real correlation between variables, so sample values should be used. The same happens with the vector of means. The maximum likelihood method applied to the multivariate normal leads to the best estimators being the vector of sample means $\bar{X}$ and the sample variance-covariance matrix S. A result in multivariate statistics states that the expression $(x - \mu)^T S^{-1} (x - \mu)$ follow a distribution T^2 of Hotelling (Peña, 2002), which is related to the distribution *F* Fisher's using the relation $F_{m,n-m} = \frac{(n-m)T^2_{m,n-m}}{m(n-1)}$. Therefore, in the previous developments it is only required to make the change of $\chi^2_{(1-\alpha),m}$ for variable $F_{m,n-m}$ that satisfies the value $1 - \alpha$ required. The process variation regions will still be ellipses (in two dimensions): $(x - \mu)^T S^{-1} (x - \mu) = T^2_{m,n-m} = \frac{m\,(n-1)F_{m,n-m}}{(n-m)}$.

If the equation is taken as a basis $LTN_i = \mu_{c_i} \pm \sqrt{|\Sigma||\Sigma_1^{-1}|\chi^2_{\beta_i}}$ To obtain the natural tolerance limits of the process, which assumes knowledge of the process parameters, it must be modified using the estimators for the vector of means and for the variance-covariance matrix as follows: $LTN_i = \bar{X}_{c_i} \pm \sqrt{|S||S_i^{-1}|T^2_{m,n-m}}$. If the corresponding value of the $T^2_{m,n-m}$, it is obtained that:

$$T^2_{m,n-m} = \frac{\left(LTN_i - \bar{X}_{c_i}\right)^2}{|S||S_i^{-1}|}, i = 1,2, \dots, m$$

Since the natural tolerance limits are equally spaced from the vector of means, they will result in the same value of $T^2_{m,n-m}$ for the same variable or quality characteristic.

Results

Chen (1994) considers two numerical examples where it applies capacity indices cited in the literature. In this work, the data from the first example shown in Table 2 were used to calculate the capacity indices with the proposed modification.

Table 2. Variable data, brinell hardness (H) and tensile strength (S)

H	S	H	S	H	S
143	34.2	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		

Source: Chen (1994)

In this example, a bivariate normal distribution is used. The variables cited are Brinell hardness (H) and tensile strength (S). Engineering tolerances for the listed quality characteristics are:

$$Dureza\ LEI = 112.3\ LES = 241.7$$

$$Resistencia\ LEI = 32.7\ LES = 73.3$$

$$Vector\ de\ valores\ nominales: [177 \quad 53]$$

When the 25 measurements obtained for the quality characteristics were analyzed, it was found that the averages for both variables, respectively, are: [(177.2&52.316)] and that the variance and covariance matrix is given by:

$$S = \begin{bmatrix} 338 & 88.8925 \\ 88.8925 & 33.6247 \end{bmatrix}$$

Also, its determinant is 3463.3, and its matrix inverse is:

$$S^{-1} = \begin{bmatrix} 0.0097 & -0.0257 \\ -0.0257 & 0.0976 \end{bmatrix}$$

Figure 1 shows both the specification and process regions. Taking into account the data provided and for the calculation of the C_{pM}, We will assume that the process average is centered relative to the specifications, which means that the $\bar{x}_i$ represent the nominal values of the specifications for each of the variables or quality characteristics.

Obtaining the values of $T^2_{m,n-m}$ with respect to the nominal values for the calculation of the C_{pM}

The $T^2_{m,n-m}$ associated with the specifications of the hardness variable are the following:

$$T^2_{2,23} = \frac{\left(X_{1,inf} - \bar{x}_1\right)^2}{|S||S_1^{-1}|} = \frac{(LEI_1 - \bar{x}_1)^2}{|S||S_1^{-1}|} = \frac{(112.3 - 177)^2}{(3463.3)(0.0976)} = 12.3842$$

$$T^2_{2,23} = \frac{\left(X_{1,sup} - \bar{x}_1\right)^2}{|S||S_1^{-1}|} = \frac{(LES_1 - \bar{x}_1)^2}{|S||S_1^{-1}|} = \frac{(241.7 - 177)^2}{(3463.3)(0.0976)} = 12.3842$$

The value of $T^2_{m,n-m}$ associated with the resistance variable will be:

$$T^2_{2,23} = \frac{\left(X_{2,inf} - \bar{x}_2\right)^2}{|S||S_2^{-1}|} = \frac{(LEI_2 - \bar{x}_2)^2}{|S||S_2^{-1}|} \frac{(32.7 - 53)^2}{(3463.3)(0.00976)} = 12.1914$$

$$T^2_{2,23} = \frac{\left(X_{2,sup} - \bar{x}_2\right)^2}{|S||S_2^{-1}|} = \frac{(LES_2 - \bar{x}_2)^2}{|S||S_2^{-1}|} \frac{(73.3 - 53)^2}{(3463.3)(0.00976)} = 12.1914$$

Given that there are two variables, 25 observations and using a confidence of 99.73%, we have that the $F_{2,23}$ associated with the process takes the value of 7.73346, which implies a value of:

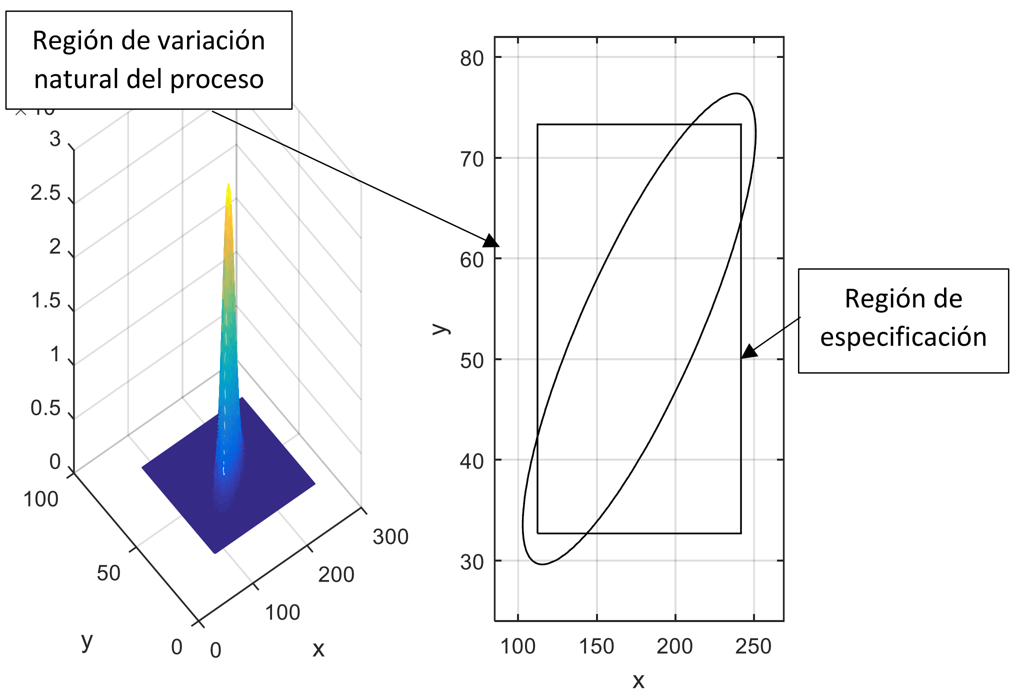
$$T^2 = \frac{m\,(n-1)F_{m,n-m}}{(n-m)} = \frac{2(25-1)7.73346}{(25-2)} = 16.1394$$

So the value of C_{pM} Is given by:

$$C_{pM} = min\sqrt{\frac{T^2_{m,n-m}}{T^2_{(1-\alpha),2,23}}} = min\left[\sqrt{\frac{12.3842}{16.1394}}, \sqrt{\frac{12.1914}{16.1394}}\right] = min[0.8759, 0.8691]$$
$$= 0.8691$$

As can be seen in figure 1, the region of natural variation of the process is larger than the region of the specifications, so it is easy to deduce that the process is not potentially capable of meeting the specifications, which is also reflected in the value of C_{pM}. The assumption that the process is centered facilitates the calculation of the index.

Figure 1. Specification region and process region for H and S, assuming that the process is centered



Source: self made

Obtaining the values of $T^2_{m,n-m}$ with respect to the vector of means for the calculation of the C_{pkM}

The current average for H is 177.2 and for S it is 52.316. With these values, the value of the C_{pkM}. Figure 2 shows the current location of the process. First, we will obtain the values of $T^2_{m,n-m}$ for the upper and lower specifications of H and S.

$$T_{2,23}^2 = \frac{\left(X_{1,inf} - \bar{x}_1\right)^2}{|\mathrm{S}||\mathrm{S}_1^{-1}|} = \frac{(112.3 - 177.2)^2}{(3463.3)(0.0976)} = 12.4609$$

$$T_{2,23}^2 = \frac{\left(X_{1,sup} - \bar{x}_1\right)^2}{|\mathrm{S}||\mathrm{S}_1^{-1}|} = \frac{(241.7 - 177.2)^2}{(3463.3)(0.0976)} = 12.3078$$

$$T_{2,23}^2 = \frac{\left(X_{2,inf} - \bar{x}_2\right)^2}{|\mathrm{S}||\mathrm{S}_2^{-1}|} = \frac{(32.7 - 52.316)^2}{(3463.3)(0.00976)} = 11.3836$$

$$T_{2,23}^2 = \frac{\left(X_{2,sup} - \bar{x}_2\right)^2}{|\mathrm{S}||\mathrm{S}_2^{-1}|} = \frac{(73.3 - 52.316)^2}{(3463.3)(0.00976)} = 13.0268$$

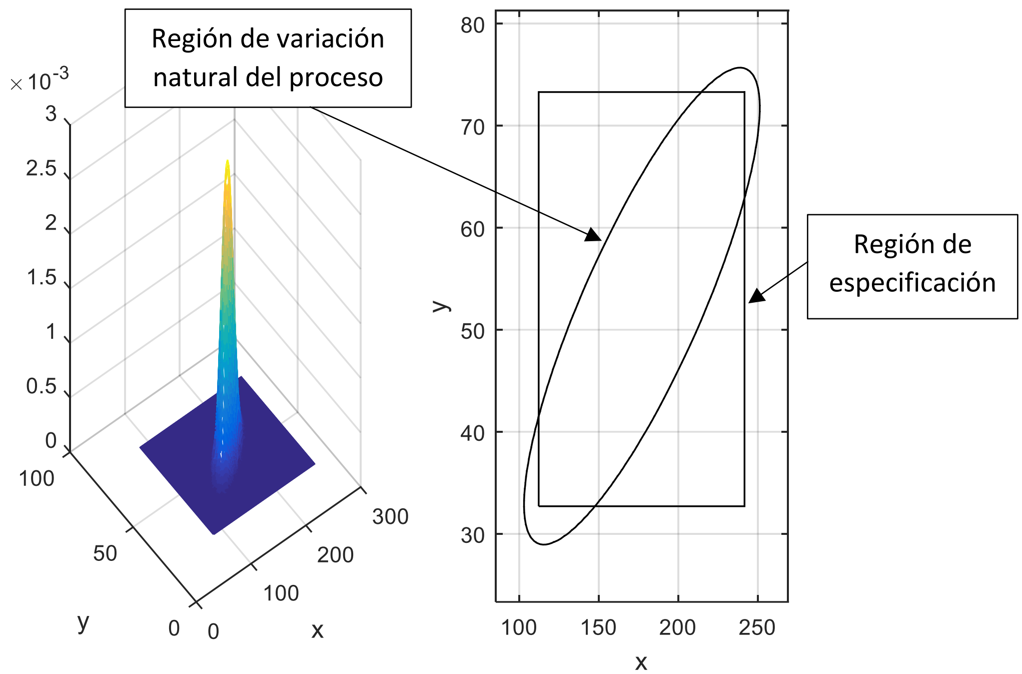
The value of C_{pkM} is calculated by:

$$C_{pkM} = min\sqrt{\frac{T_{m,n-m}^2}{T_{(0.9973),2,23}^2}} = min\left[\sqrt{\frac{12.4609}{16.1396}}, \sqrt{\frac{12.3078}{16.1394}}, \sqrt{\frac{11.3836}{16.1394}}, \sqrt{\frac{13.0268}{16.1394}}\right]$$

$$= min[0.8787, 0.8733, 0.8398, 0.8984] = 0.8398$$

Which shows that the process is not capable. This is because the process is slightly outside the specification region, as seen in Figure 2, where the process variation region is slightly outside the specification region.

Figure 2. Specification region and process region for H and S, when the process is at its current averages.



Source: self made

Discussion

The multivariate capacity indices developed in this work are an improvement to those presented by Cuamea and Rodriguez (2014), since they are more realistic in the sense that they use the information of the estimators contained in the sample and are consistent since they adequately measure the capacity of a process, through the indices C_{pM} *and* C_{pkM} proposed, which are obtained by correctly specifying which is the specification region and which is the variation region of the process for m quality characteristics analyzed jointly. Both regions are defined using the Mahalanobis distance, which allows them to be defined as ellipsoids, unlike other similar proposals that use, instead of the Mahalanobis distance, the volume ratio of the ellipsoids (Shahriari and Abdollahzadeh, 2009; Wang and Chen, 1998), which makes the calculation of the process capacity more complex. Other proposals use the main components to determine the capability of a process, but there is still no agreement on how many main components should be taken into account and how much of the variation should be explained (Barreto and Herrera, 2021; Shinde and Khadse, 2009). . In addition, these proposals do not take into account the information when the quality characteristics are correlated, information that is used in our proposal to calculate the capacity of a multivariate process. There are other proposals that obtain the capacity of a process from the calculation of the defective fraction that the process is producing (Kotz and Johnson, 2002), which can also be determined in our proposal using Monte Carlo simulation.

It can also be seen that the interpretation of the multivariate indices obtained through the proposal presented in this work is consistent with their univariate counterparts. In the calculation of the indexes proposed in this investigation, the value of the natural tolerance limits of the process is obtained for each of the quality characteristics, which allow us to know the region of variation of the process for each of the characteristics, considering a confidence of 99.73% (2700 ppm), and graphically establish whether or not a process is capable of meeting all specifications. On the other hand, our proposal has limitations, since it requires that the process can be modeled through a multivariate normal distribution and that the m quality characteristics have bilateral and symmetric specifications.

Conclusions

The multivariate capacity indices proposed in this research are consistent with the respective univariate capacity indices, since the concept of distance is applied for its calculation; they are also obtained as a quotient just like their univariate counterpart. In its calculation, the information of the specifications is used, as well as the information related to the process contained in the sample. And the interpretation is equivalent to the metric established for univariate capability indices, in other words, they are interpreted in the same way. On the other hand, in its calculation the quality characteristics are analyzed jointly or simultaneously, whether they are correlated or not, and they are easy to calculate and interpret.

Future lines of research

In future research, it is recommended that work continue on the definition of capability indices for bilateral specifications for processes that cannot be modeled using the multivariate normal. Additionally, the case of capability indices for processes with unilateral specifications should be investigated in more depth. In the same way, capability indices should be proposed for non-symmetric bilateral specifications.

References

Barreto, R., & Acosta Roberto, H. (2021). Application of a proposed reliability analysis multivariate capability index on manufacturing processes.

Bothe, D. R. (1999). Composite capability index for multiple product characteristics. Quality Engineering, 12(2), 253–258.

Castagliola, P., & Castellanos, J. V. G. (2008). Process capability indices dedicated to bivariate non normal distributions. Journal of Quality in Maintenance Engineering, 14(1), 87–101.

Chen, H. (1994). A multivariate process capability index over a rectangular solid tolerance zone. Statistica Sinica, 4(2), 749–758.

Cuamea, G., & Rodriguez, M. A. (2014). Propuesta para evaluar la capacidad de procesos de manufactura multivariados. Revista Ingeniería Industrial, 13(2), 35–47.

de-Felipe, D., & Benedito, E. (2017). A review of univariate and multivariate process capability indices. International Journal of Advanced Manufacturing Technology, 92(5–8), 1687–1705.

Foster, E. J., Barton, R. R., Gautam, N., Truss, L. T., & Tew, J. D. (2005). The process-oriented multivariate capability index. International Journal of Production Research, 43(10), 2135–2148.

Juran. (1974). Juran's Quality Control Handbook.

Kane. (1986). Process Capability Indices. Journal of Quality Technology, 18(1).

Khadse, K. G., & Shinde, R. L. (2009). Probability-Based Process Capability Indices. Communications in Statistics - Simulation and Computation, 38(4), 884–904.

Kotz, S., & Johnson, N. L. (2002). Process capability indices - A review, 1992-2000. Journal of Quality Technology, 34(1), 2–19.

Kotz, S., & Johnson, N. L. (2018). Process Capability Indices—A Review, 1992–2000. Https://Doi.Org/10.1080/00224065.2002.11980119, 34(1), 2–19.

Pana, J. N., & Lee, C. Y. (2010). New capability indices for evaluating the performance of multivariate manufacturing processes. Quality and Reliability Engineering International, 26(1), 3–15.

Peña, D. (2002). Análisis de Datos Multivariados. McGraw Hill.

Prokopenko, J. (1998). Globalización, competitividad y estrategias de productividad. Boletín Cinterfor, 39.

Shahriari, H., & Abdollahzadeh, M. (2009). A new multivariate process capability vector. Quality Engineering, 21(3), 290–299.

Shinde, R. L., & Khadse, K. G. (2009). Multivariate process capability using principal component analysis. Quality and Reliability Engineering International, 25(1), 69–77.

Taam, W., Subbaiah, P., & Liddy, J. W. (1993). A note on multivariate capability indices. Journal of Applied Statistics, 20(3), 339–351.

Wang, F. K., & Chen, J. C. (1998). Capability index using principal components analysis. Quality Engineering, 11(1), 21–27. https://doi.org/10.1080/08982119808919208

Wierda, S. J. (1994). Multivariate statistical process control - recent results and directions. Statistica Neerlandica, 48(2), 147–168.

Wu, C. W., Pearn, W. L., & Kotz, S. (2009). An overview of theory and practice on process capability indices for quality assurance. International Journal of Production Economics, 117(2), 338–359.

Rol de Contribución	Autor (es)
Conceptualización	Guillermo Cuamea Cruz
Metodología	Principal: Guillermo Cuamea Cruz Que apoya: Manuel Arnoldo Rodríguez Medina
Análisis Formal	Principal: Guillermo Cuamea Cruz Que apoya: Manuel Arnoldo Rodríguez Medina Que apoya: Eduardo Rafael Poblano Ojinaga Que apoya: Rafael García Martínez Que apoya: Diego Adiel Sandoval Chávez
Investigación	Guillermo Cuamea Cruz
Escritura - Preparación del borrador original	Guillermo Cuamea Cruz
Escritura - Revisión y edición	Principal: Guillermo Cuamea Cruz Que apoya: Manuel Arnoldo Rodríguez Medina Que apoya: Eduardo Rafael Poblano Ojinaga Que apoya: Rafael García Martínez Que apoya: Diego Adiel Sandoval Chávez
Adquisición de fondos	Guillermo Cuamea Cruz