

## INDICADORES DE CAPACIDAD APLICADOS A LA DESERCIÓN EN LAS UNIVERSIDADES COLOMBIANAS

### CAPACITY INDICATORS APPLIED TO THE DROP IN THE COLOMBIAN UNIVERSITIES

José Daniel Mosquera Artamónov

Ing. Industrial., estudiante de maestría en ingeniería de calidad, Facultad de Ingeniería, División de Estudios de Posgrado, Universidad Autónoma de Querétaro, Querétaro, México  
xoce15@ingenieros.com

Julio César Mosquera Mosquera

Físico, Ph.D., Profesor Asistente, Facultad de Ciencias Básicas,  
Investigador Grupo Optoelectrónica, Universidad del Quindío, Armenia, Colombia  
jcmosquera@uniquindio.edu.co

Irina Artamónova

Física, M.Sc., Profesora Asistente, Facultad de Ciencias Básicas, Investigadora Grupo Estudio y aplicación de herramientas estadísticas modernas en la solución de problemas del entorno.  
Universidad del Quindío, Armenia, Colombia  
artiri@gmail.com

**Fecha de recepción:** 15 de abril de 2011

**Fecha de aprobación:** 8 de diciembre de 2011

### RESUMEN

En este trabajo, se presenta una propuesta para implementar los Índices de Capacidad de un Proceso como herramienta para medir la efectividad de un proceso social. Estos índices son ampliamente empleados en la industria de productos de consumo, pero no es clara la forma como pueden ser usados en el control de procesos sociales. El trabajo se centra en el análisis de aquellos procesos que no están normalmente distribuidos y que por lo tanto, en una primera aproximación teórica, no son susceptibles de ser controlados con el uso de Índices de Capacidad. Mediante un doble ejemplo aplicado al caso de la deserción en las universidades colombianas, se ofrece una metodología detallada de cómo solucionar el problema de la no normalidad de los datos con el uso de la transformación de Box-Cox, la distribución de Weibull y el método de percentiles de Clements. De esta manera, se demuestra la aplicabilidad de los métodos ya contrastados en otros ámbitos, a casos concretos de deserción en las universidades colombianas. El método de índices de capacidad puede usarse para evaluar el proceso de deserción y definir las metas reales de deserción, posibles de cumplir. Se propone además, que esta metodología sea usada en el control de otros procesos sociales.

**Palabras clave:** índice de capacidad, transformada de box-cox, distribución weibull, deserción académica, normalidad, método de percentiles de Clements

### ABSTRACT

In this paper, a proposal to implement Process Capability Indices as a tool to measure the effectiveness of a social process is presented. These indices are widely used in consumer

products industry, but it is not clear how these can be used to control social processes. The work focuses on the analysis of processes that are not normally distributed and therefore, in first theoretical approximation, are not likely to be controlled by using capability indices. Through a double example applied to the case of desertion in Colombian universities, a detailed methodology of how to solve the problem of non-normality of the data using the Box-Cox transformation, the Weibull distribution and method based on Clements Percentiles are presented. In this way, applicability of different methods already proven in other areas is demonstrated for specific cases of desertion in Colombian universities. The capacity indices method can be used to evaluate the desertion process and set real abandonment goals, possible to achieve. This methodology is proposed in order to control other social processes.

**Keywords:** indexes process capabilities, box-cox transformation, weibull distribution, academic dropout, normality, method based on Clements percentiles.

## INTRODUCCIÓN

El desarrollo industrial ha sido acompañado de una incesante búsqueda de métodos para controlar los atributos de calidad de los productos fabricados. Las industrias desean tener un sello de diferenciación de sus productos que sea evidente a los clientes potenciales, en el momento de tomar la decisión de comprar el bien o servicio [1]. Este desarrollo ha llevado a enfocar las políticas de calidad, no en el producto o servicio, sino en el control de los procesos con los cuales se producen. Esto último ha potencializado el desarrollo de diferentes métodos que permitan controlar los factores que inciden en la calidad; entre los más conocidos están las diferentes certificaciones ISO y las cartas de control estadístico, siendo estas últimas, las que tienen mayor validez en el momento de afirmar, si un proceso industrial está bajo control o en caso contrario, fuera de control.

Las cartas de control permiten conocer si la distribución subyacente de un proceso o sus parámetros, tales como el promedio y la dispersión, están en un nivel deseado [2]. Estos parámetros son dos estimadores muy importantes para obtener los diferentes indicadores o índices de capacidad conformados por los cocientes entre la variación natural del proceso y el nivel de variación especificada. Para que un proceso sea considerado capaz, la variación actual no deberá representar más del 75% de la variación permitida [3]. Para realizar una valoración objetiva de la capacidad del proceso, es necesario contar con indicadores que permitan saber si un servicio o producto, puede o no cumplir con unos límites de especificación [4], que permitan aumentar paulatinamente la capacidad de los procesos evaluados. Estos indicadores reciben el nombre de *índices de capacidad*. El índice de capacidad del proceso (ICP), es el método utilizado para calcular la habilidad del proceso, para cumplir con las especificaciones, y se expresa [5], con la ecuación 1:

$$ICP = C_p = \frac{LSE - LIE}{6\hat{\sigma}} \quad (1)$$

Donde:

LSE: Límite superior especificado

LIE: Límite inferior especificado

$\hat{\sigma}$  : Desviación estándar de los datos

Los índices de capacidad de un proceso suministran información numérica de cuánto se ajusta el proceso a los límites de especificación establecidos [6]. Para ser calculados, se debe obtener una muestra de datos tomados en el momento de la inspección, los cuales deben cumplir dos condiciones: 1) Que se hayan tomado de un proceso que esté bajo control. 2) Que se ajusten a una distribución normal.

Aunque en la actualidad, esta metodología de control de procesos tiene una amplia utilización en la industria de bienes, no es muy clara en el control similar sobre la industria de servicios y menos en el campo de control de procesos sociales. Un proceso social se puede definir como la interacción dinámica de grupos sociales. En muchos casos, este proceso es dirigido por una de las partes a fin de obtener un resultado medible en la otra parte del grupo social. De esta manera, se busca controlar y medir los resultados de acciones dirigidas a cambiar ciertos hábitos en una población objetivo. Sin embargo, no es muy claro cómo medir la efectividad de tal proceso de intervención. Ya los autores hicieron un primer trabajo y demostraron la importancia de identificar la normalidad de los datos analizados en procesos sociales [7]. Sin embargo, sólo se ha presentado un método específico para el caso de la Deserción en la UTP. Por esta razón, el objetivo central de este trabajo es demostrar que sí es posible usar los ICP para valorar la capacidad de un proceso social, es decir, utilizar los ICP para valorar el mejoramiento de los servicios y mostrar la metodología para su aplicación a partir de dos casos concretos.

## 1. METODOLOGÍA POR UTILIZAR

La metodología general que se propone utilizar en los procesos sociales, se presenta en la Figura 1:

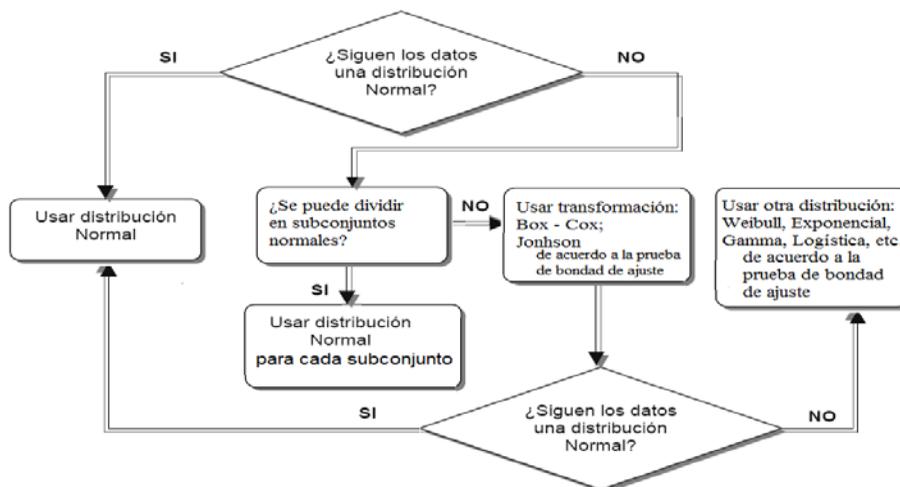


Figura 1. Toma de decisión en el análisis de datos

- Primero se deberá revisar la normalidad de los datos muestrales. Para este fin, se puede utilizar: coeficientes de curtosis y simetría como pruebas de bondad de ajuste a la distribución normal.
- La segunda propuesta, es dividir la muestra original en un conjunto de muestras cada una de cuales pueda ser ajustada a su propia distribución normal. Este método se utiliza en muestras de gran tamaño.
- Tercera propuesta: para cuando los datos experimentales no sigan una distribución normal, antes de buscar una distribución que se ajuste a los datos, se recomienda utilizar datos transformados para conseguir la normalidad, argumentando que los resultados son más simples de utilizar, porque si se encuentra alguna transformación adecuada, la forma de construir el índice no cambiaría. Una estrategia muy difundida actualmente, en especial en temas de calidad, es la presentada por Box y Cox (1964), explicada en Peña S.D.P [8]. Sin embargo, es posible aplicar otras transformaciones, como la de Johnson.
- Si los datos transformados se ajustan a la distribución normal, se utilizarán los mismos procedimientos para calcular los índices de capacidad del proceso.
- Cuando ninguno de los anteriores procedimientos resulta exitoso, se buscaría una distribución que se ajuste a los datos. Para tal fin, puede utilizarse las distribuciones de Weibull, Gamma, exponencial, logística, log-normal, distribución de Pareto Generalizada [9] u otras. ¿Cuál de todas se debe utilizar? Se define con la prueba de bondad de ajuste. En años recientes, se han sugerido distintos procedimientos para hacer estudios de Capacidad de Procesos en poblaciones no normales [10], [11], [12], [13] y [14], proponiendo diferentes fórmulas para calcular los índices de capacidad de procesos no normales [15].

Como ya se mencionó, los datos del proceso recolectados deben estar normalmente distribuidos. Si los datos obtenidos siguen una distribución no normal o asimétrica, el cálculo del índice de capacidad ( $C_p$ ), no sería un muy buen estimador de las reales probabilidades de producción de defectuosos. Para solucionar este problema, existen numerosos métodos estadísticos de los cuales en este trabajo, se proponen dos de ellos: transformación de Box-Cox o el modelo de Weibull [17], por ser los más utilizados en la industria. A continuación, se muestra cada uno por separado.

### 1.1 LA TRANSFORMACIÓN BOX-COX [8]

La familia de transformaciones de Box-Cox soluciona problemas de normalidad y heterocedasticidad (no homogeneidad de varianzas) y ofrece tres métodos para solucionar este problema:

**Método 1:** Una idea útil en muchas aplicaciones, es considerar transformar los datos de la respuesta  $Y$  en la potencia,  $Y^\lambda$ , por decir, y encontrar el mejor valor de  $\lambda$ . Si el mejor valor de  $\lambda$  fuera  $\lambda=0$ , entonces se tendría  $Y^0=1$ . Esta transformación simplemente reproduce los datos de entrada lo cual no es deseable. Para el caso  $\lambda=0$ , este método no es recomendable.

**Método 2:** Otro método consiste en transformar la variable  $Y$  en la variable  $W$  y así no se hay problema en el método 1 cuando  $\lambda=0$ , porque  $\ln(Y)$  es el límite apropiado, cuando  $\lambda$  tiende a cero, y así la familia es continua en  $\lambda$ .

$$W = \begin{cases} \frac{Y^{\lambda-1}}{\lambda}, & \text{si } \lambda \neq 0 \\ \ln Y, & \text{si } \lambda = 0 \end{cases} \quad (2)$$

**Método 3:** En este método, se transforma la variable  $Y$  en la variable  $V$  de la siguiente forma:

$$V = \begin{cases} \frac{Y^{\lambda-1}}{\lambda Y^{\lambda-1}}, & \text{si } \lambda \neq 0 \\ Y \ln Y, & \text{si } \lambda = 0 \end{cases} \quad (3)$$

Donde la cantidad  $Y$  es la media geométrica de los  $Y_i$ ,

$$Y = \sqrt[n]{Y_1 Y_2 Y_3 \dots Y_n} \quad (4)$$

La cual es una constante y debe ser calculada al inicio de los procedimientos de cálculo de  $\lambda$ , usualmente por la exponencial de la fórmula

$$\ln Y = \frac{1}{n} \sum_{i=1}^n \ln Y_i \quad (5)$$

## 1.2 TRASFORMACIÓN POR DISTRIBUCIÓN DE WEIBULL [18]

Esta distribución se aplica en los análisis de confiabilidad para establecer, por ejemplo: el período de vida de un componente hasta que presente una falla. La ecuación para la función de distribución acumulada de Weibull es:

$$F(x, \alpha, \beta) = 1 - e^{-(x/\beta)^\alpha} \quad (6)$$

La función de densidad de probabilidad es:

$$f(x, \alpha, \beta) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} \quad (7)$$

La apariencia de la curva Weibull varía notablemente en función del valor de  $\beta$  y su interpretación está dada en Tamborero P. J. M. [19]:

- Si :
- $\beta < 1$  La tasa de fallas está disminuyendo;
  - $\beta = 1$  Se tiene una tasa constante y se asemeja a la Distribución Exponencial;
  - $\beta = 2$  Distribución lineal (Raleigh);
  - $\beta > 2$  La tasa de fallas se está incrementando; Para valores entre 3 y 4 esta distribución se asemeja a la distribución normal.

Este modelo genera índices de capacidad de largo plazo, a diferencia cuando se emplea la distribución normal que permite conocer los índices de capacidad de corto plazo (Cp, Cpk, CPU y CPL), y los de largo plazo (Pp, Ppk, PPU y PPL). En la práctica, un valor de 1,33 es el valor mínimo aceptable para un índice de capacidad. Valores por debajo de este umbral 1,3 y por encima de 1, permiten afirmar que aunque esté bajo control estadístico, el proceso no cumple con las especificaciones deseadas. En la Tabla 1, se muestran algunas referencias sobre cuándo usar cada índices [20], en caso de normalidad.

**Tabla 1.** Descripción de los índices de capacidad de los procesos

Índice	Uso	Fórmula
Cp o Pp	El proceso está centrado en los límites de especificación	$(LES-LEI)/6\sigma$
Cpk o Ppk	El proceso no está centrado en los límites de especificación, pero está contenido en ellos	$\text{Min}\{(LES-\mu)/3\sigma, (\mu-LEI)/3\sigma\}$
CPU o PPU	El proceso sólo tiene un límite de especificación superior	$(LES-\mu)/3\sigma$
CPL o PPL	El proceso sólo tiene un límite de especificación inferior	$(\mu-LEI)/3\sigma$

En el modelo Weibull, el programa MINITAB calcula los índices de capacidad a largo plazo, Pp, Ppk, PPU y PPL. Dichos cálculos se basan en estimadores de máxima verosimilitud para los parámetros de la distribución Weibull.

Otra metodología es utilizar los índices de capacidad de procesos específicos para el caso de distribuciones no normales. El cálculo de tales índices, está bien explicado en Quaglino M. B., et al [15].

### 1.3. MÉTODO DE PERCENTILES DE CLEMENTS

Clements J.A. [16], propuso un índice generalizado de capacidad de procesos. El procedimiento consiste en calcular los índices de capacidad de procesos como una modificación de los índices de capacidad de procesos normales, utilizando los valores de puntos percentiles no normales y

considerando la mediana como medida central. En la ecuación (1), se reemplaza el valor  $6\sigma$  por la longitud del intervalo entre el límite superior e inferior de los puntos percentiles 99,865 y 0,135 de una distribución X, esto es respectivamente  $U_p - L_p$ . Se define el  $C_p$  de la siguiente manera:

$$C_p = \frac{U - L}{U_p - L_p} \quad (8)$$

Para el  $C_{pk}$ , la media del proceso  $\mu$  es estimada por la mediana M, y los valores de  $3\sigma$  son estimados por  $M - L_p$  y  $U_p - M$ , respectivamente. Así, se tienen las siguientes expresiones.

$$CC_{pl} = \frac{M-L}{M-L_p} \quad (9)$$

$$CC_{pu} = \frac{U-M}{U_p-M} \quad (10)$$

$$CC_{pk} = \text{Min}(CC_{pu}, CC_{pl}) \quad (11)$$

Donde, M es la mediana de la distribución del proceso, L y U son los Límites de Especificación Inferior y Superior, respectivamente. Los valores  $L_p$  y  $U_p$  son los percentiles 0,135% y 99,865% de la distribución utilizada.

## 2. FORMULACIÓN DEL PROBLEMA

La deserción es el proceso de abandono, voluntario o forzoso de la carrera en la cual se matricula un estudiante, por la influencia positiva o negativa de circunstancias internas o externas al estudiante. Hasta el año 2004, las instituciones de educación superior venían realizando aisladamente las acciones para mitigar la deserción en sus establecimientos: flexibilización curricular, acompañamiento académico, articulación con la educación media, asistencia a nuevos estudiantes, apoyo económico y programas integrales.

En el año 2005, se uso el diseño de una metodología de seguimiento a la deserción que respondiera a la necesidad de disponer de una visión nacional sobre el tema, y facilitara el análisis, la discusión, el intercambio de experiencias, el diagnóstico, el diseño y la evaluación de estrategias, la medición y la referenciación. Este fue el inicio del Sistema de Prevención de la Deserción en Educación Superior (SPADIES) [21]. Hoy, el sistema SPADIES, es pionero en América Latina, está instalado en el 94% de las instituciones y hace seguimiento a más de 2.900.000 estudiantes matriculados entre 1998 y 2009, y permite que cada institución cuente con un perfil completo de los estudiantes y con alertas tempranas sobre los factores que los hacen vulnerables para direccionar de forma más eficiente, apoyos y políticas.

En la actualidad, la tasa de deserción por cohorte, se define como la proporción de alumnos que abandonan la carrera en el transcurso de una cohorte educativa y se calcula por la relación entre el número de desertores durante el ciclo y el número de alumnos inscritos en este ciclo. De acuerdo con la medición del SPADIES, un desertor es el estudiante que lleva dos períodos consecutivos sin matricularse y no se ha graduado. Otra medida cuantitativa de deserción, es la tasa de deserción bruta anual.

En este trabajo, se presentan dos ejemplos reales con los cuales se muestra la aplicabilidad de esta herramienta estadística (los índices de capacidad de los procesos), a evaluación de procesos de deserción. El primer ejemplo, es el proceso de deserción académica de la Universidad Tecnológica de Pereira, donde se hace una explicación minuciosa de cómo usar la transformada Box- Cox o la distribución de Weibull.

Para el segundo caso, se muestra el efecto del uso de los ICP en la deserción académica de Colombia. Estos dos ejemplos permiten considerar una generalización en el tratamiento de datos de procesos sociales al momento de medir su capacidad.

## 2.1 CASO 1. UNIVERSIDAD TECNOLÓGICA DE PEREIRA

Durante cerca de 5 años, la gerencia para disminuir la deserción en la Universidad Tecnológica de Pereira, ha planteado e implementado diferentes estrategias para disminuir la deserción estudiantil [22], [23] y [24], es decir, fomentar la continuidad de los estudios universitarios en jóvenes con alto riesgo de deserción.

En este estudio, se hace necesario valorar la capacidad de tales estrategias en la solución real del problema de deserción. La propuesta es aplicar los ICP para el proceso de control de la deserción. Se sabe que los indicadores de capacidad fueron diseñados para ser utilizados en la industria de manufactura, pero no hay alguna oposición de emplearlos para evaluar procesos sociales cuando los estudiantes desertores se asemejan a los “productos defectuosos”. Es posible, hacer una analogía entre la industria en donde se pretende controlar el número de productos defectuosos generados, y las universidades donde se trata de controlar el número de desertores por período académico.

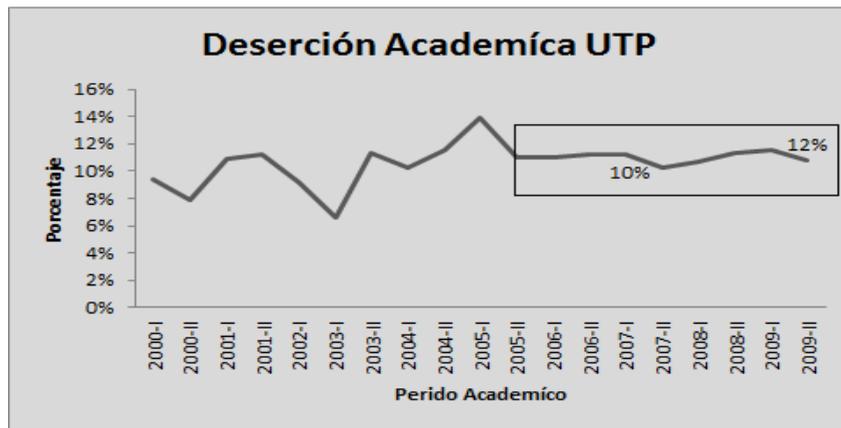
Para generar los indicadores de capacidad (ICP), se debe validar las hipótesis para control a un proceso, siendo una de ellas, la normalidad de los datos, para lo cual se prueba que los datos sigan una distribución normal.

Al someter los datos a dos pruebas de normalidad, la de Shapiro-Wilk [25] y la de Anderson-Darling [26] y [27], con un umbral de significancia  $\alpha = 0,05$ , se rechaza la hipótesis nula, y se acepta la no-normalidad de los datos (Tabla 2). Una de las soluciones del problema de la no-normalidad, es dividir la muestra en diferentes subconjuntos que sigan la distribución normal, pero se descarta este procedimiento por no ser la muestra de gran tamaño.

**Tabla 2.** Prueba de la hipótesis de normalidad

Prueba de Shapiro-Wilk		Prueba de Anderson-Darling	
W (valor observado)	0,872	A <sup>2</sup> de Anderson-Darling	1,191
p-value unilateral	0,015		0,003
Alpha	0,05		

Conocido el porcentaje de desertores por semestre en la Universidad Tecnológica de Pereira, se diseñó la Figura 2.

**Figura 2.** Serie de tiempo para deserción académica

**Fuente:** Oficina de Deserción I-440 UTP

Este caso fue analizado en Mosquera M.J.C. et al [7], y llegó a la conclusión de que la transformación de Johnson es la óptima para este caso (por tener mayor probabilidad  $P = 0,941$  y menor  $AD = 0,158$ ). Pero en este artículo, se hace una explicación minuciosa de cómo usar la transformada Box - Cox ( $AD = 1,027$ ) y la distribución de Weibull ( $AD = 1,077$ ), por ser otros dos procedimientos que tienen mejor ajuste de acuerdo con la prueba de bondad de ajuste realizada (Tabla 3), al compararlos con la transformación de Johnson expuesta en Mosquera M.J.C. et al [7].

**Tabla 3.** Prueba de bondad del ajuste

Distribución	AD	P
Normal	1,191	<0,005
Transformación Box-Cox	1,027	0,008
Lognormal	1,481	<0,005
Lognormal de 3 parámetros	1,184	*
Exponencial	6,693	<0,003
Exponencial de 2 parámetros	4,275	<0,010
Weibull	1,077	<0,010

Weibull de 3 parámetros	1,079	<0,005
Valor extremo más pequeño	1,195	<0,010
Valor extremo más grande	1,786	<0,010
Gamma	1,369	<0,005
Gamma de 3 parámetros	1,304	*
Logística	0,942	0,008
Loglogística	1,119	<0,005
Loglogística de 3 parámetros	0,942	*
Transformación de Johnson	0,158	0,941

Para emplear la transformación de Box-Cox, se debe conocer el coeficiente  $\lambda$  óptimo. En el mercado, existen diversos software que permiten realizar la transformación y calcular este coeficiente con gran facilidad, siendo uno de estos Minitab, que genera la Figura 3 para  $\lambda$ .

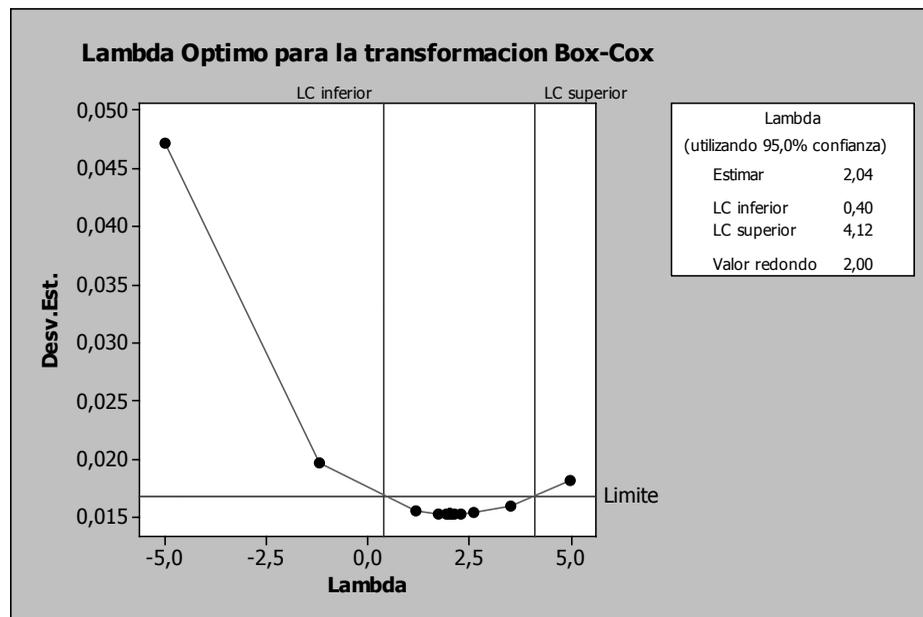
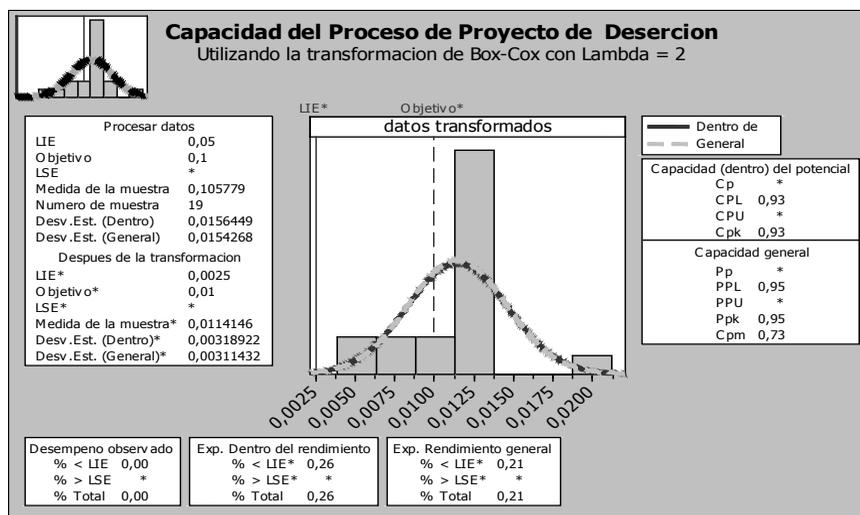


Figura 3.  $\lambda$  óptimo para la transformación Box-Cox

Con la Figura 3, se encuentra que el  $\lambda$  óptimo es igual a 2. Minitab presenta otra opción en donde el sistema es el que elige a  $\lambda$ . Determinada  $\lambda$ , se puede emplear la transformación de Box-Cox para calcular los índices de capacidad.

Como límite inferior de especificación, se proyecta una deserción académica de 5% del total de estudiantes de la UTP (información suministrada por la gerente del proyecto). Teniendo como referencia, el primer semestre de 2005, cuando se registró un total de 13,9%, la deserción más alta reportada en la última década.

Para el análisis, se ingresó como objetivo del proyecto, un 10% de deserción para conocer cómo reacciona el Cpm a esta meta, y esperando tener un valor superior a 1, lo cual quiere decir que el proceso cumple con las especificaciones y la medida del proceso está dentro de la tercera parte media de la banda de las especificaciones [28].



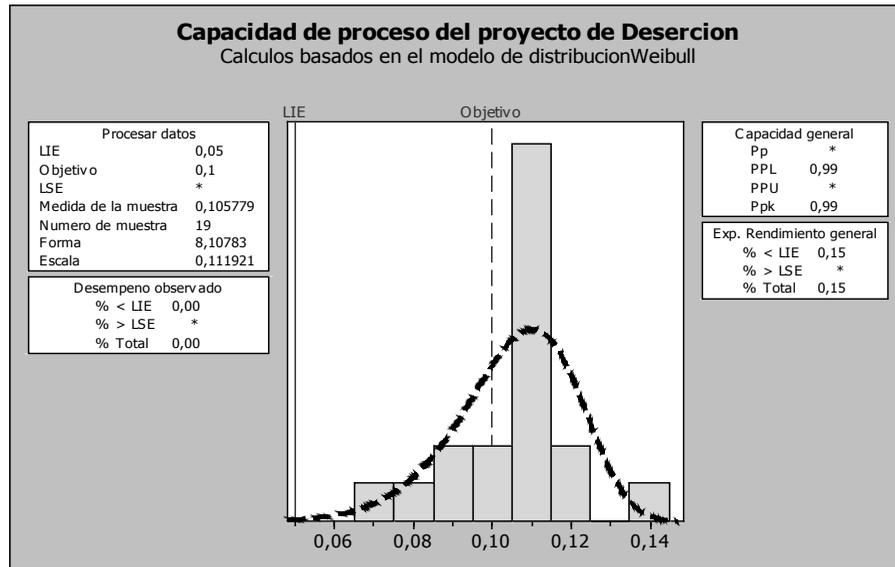
**Figura 4.** Transformación Box-Cox

En la Figura 4, desarrollada por el sistema, se puede apreciar que la deserción académica en la UTP tiene una media de 10,5%, que no alcanza el objetivo propuesto en esta investigación, y está ubicada a la derecha del objetivo. Se prefiere que el Cpk sea mayor que 1, con el cual se puede afirmar que el proceso es capaz de cumplir con las especificaciones. En el caso de la Universidad Tecnológica de Pereira, es de 0,93 lo cual implica reducir más la variabilidad del proceso y centrar el proceso en el objetivo.

Los diferentes índices del proceso están directamente relacionados con la desviación estándar del proceso, lo cual permite mostrar que aunque la deserción académica de los estudiantes puede estar bajo control (Figura 2), el proceso como tal, no está cumpliendo con las especificaciones, dado que los índices generados están por debajo de 1 (Figura 4), lo cual evidencia que el proyecto de disminuir la deserción en la UTP, necesita continuidad para cumplir con las metas propuestas, pues como menciona uno de los investigadores del proyecto, la disminución de un punto porcentual, requiere mucho esfuerzo monetario para atender los estudiantes que tengan dificultades para seguir el proceso académico.

El Cpm y el Ppk permiten determinar si el proceso está centrado en el objetivo. Este resultado se puede extraer de la gráfica de los datos transformados [29], en donde la media del proceso está al lado derecho del objetivo  $Cpm=0,73$  y  $Ppk=0,95$ , con una diferencia considerable. El Ppk expresa que la media está sesgada hacia la derecha del objetivo. Si los indicadores fueran contrarios, entonces se afirmaría que la media está a la izquierda del objetivo. Ahora, se

confrontará este análisis con el desarrollado por la distribución de Weibull, para determinar similitudes o diferencias que haya entre cada método (Figura 5).



**Figura 5.** Distribución de Weibull para la capacidad del proyecto de Deserción

Los indicadores calculados en esta investigación, están estructurados bajo el intervalo de confianza de  $6\sigma$  para el caso de transformación de Box-Cox. Las Figuras 4 y 5, muestran la implementación de dos metodologías diferentes para solucionar el problema de no normalidad de la variable que se está analizando. Para determinar cuál de los dos métodos tiene mejor ajuste a la distribución normal, se debe realizar la prueba de bondad del ajuste.

La Tabla 2 muestra información de las Pruebas estadísticas usando la prueba propuesta por Anderson-Darling ( $AD$ ), valores  $p$  correspondientes ( $P$ ) y valores  $p$  de pruebas de relación de verosimilitud, para evaluar si una distribución ajusta los datos, se elegirá entre la transformación de Box-Cox y la distribución de Weibull, y se tomará la que tenga el valor de  $p > 0,05$ , pues en este caso, se puede afirmar que tiene un ajuste adecuado. Entre los dos métodos propuestos, la distribución Weibull presenta un  $p < 0,010$  y la transformación Box-Cox  $p 0,008$ , lo cual deja una incertidumbre sobre cuál de las dos elegir. Se descartó la distribución exponencial de dos parámetros que también tiene  $p < 0,010$ , como la distribución de Weibull, pero el índice el de Anderson-Darling  $AD = 4,275$  es mayor para la distribución de Weibull que tiene  $AD = 1,077$ .

Dado que no se determinó el valor real para la distribución Weibull, además de usar el valor de  $p$ , se puede apoyar también en el valor de la prueba de Anderson-Darling ( $AD$ ), en la cual, el valor más cercano a cero, presenta mejor ajuste. Según este criterio, la transformación de Box-Cox es la que mejor ajusta los datos entre los dos métodos propuestos contra la deserción

académica de la UTP, pero no lo suficiente para concluir bajo esta transformación, sobre el proceso que se está evaluando.

Hay que tener cautela con cuál transformación o distribución concluir, pues una elección errónea de la herramienta usada, aumenta el error al momento de sobre calificar un proceso que no cumple con las especificaciones o en caso contrario, afirmar que no cumple con las especificaciones, cuando sí las está cumpliendo. Por esta razón, se recomienda realizar siempre pruebas de bondad del ajuste al momento de utilizar diferentes modelos de regresión, transformación, buscando moldearlos a la distribución normal.

En el análisis de este caso, los dos métodos elegidos para solucionar el problema de normalidad de los datos, no hicieron un buen ajuste de los datos, lo cual impide concluir sobre la capacidad de este proceso social de cumplir con la meta propuesta de tener 10% de deserción académica. Algunos se pueden ver tentados a concluir en este caso, que el proceso no es capaz de cumplir con la meta propuesta. Sin embargo, se debe analizar nuevamente los datos obtenidos, pero esta vez con una amplitud mayor de métodos estadísticos. Minitab usa para el análisis de la distribución individual, 16 métodos diferentes con los cuales se puede seguir esta misma metodología y obtener resultados que garanticen una valoración adecuada del proceso (Tabla 2).

Por último, se aplicó el método de Clements, para lo cual se calcularon para los datos de la deserción de la UTP: Mediana:0,1099; Límite Inferior:  $L_p = 0,0660$  y Límite Superior:  $U_p = 0,1384$ . Utilizando la fórmula  $C_p = (U-L)/(U_p-L_p)$ , se calcularon los  $C_p$  para diferentes Límites de Especificación Superior - U e Inferior - L, y se llegó a los resultados de la Tabla 4.

**Tabla 4.** Índice  $C_p$  para los Límites de Especificación L y U

<b>Cp para:</b>	<b>L</b>						
<b>U</b>	0,04	0,045	0,05	0,055	0,06	0,065	0,07
0,135	<b>1,312</b>	<b>1,243</b>	<b>1,174</b>	<b>1,105</b>	<b>1,036</b>	0,967	0,898
0,13	<b>1,243</b>	<b>1,174</b>	<b>1,105</b>	<b>1,036</b>	0,967	0,898	0,829
0,125	<b>1,174</b>	<b>1,105</b>	<b>1,036</b>	0,967	0,898	0,829	0,76
0,12	<b>1,105</b>	<b>1,036</b>	0,967	0,898	0,829	0,76	0,691
0,115	<b>1,036</b>	0,967	0,898	0,829	0,76	0,691	0,621
0,11	0,967	0,898	0,829	0,76	0,691	0,621	0,552
0,105	0,898	0,829	0,76	0,691	0,621	0,552	0,483
0,1	0,829	0,76	0,691	0,621	0,552	0,483	0,414

En la Tabla 4, para  $L = 0,050$  y  $U = 0,100$ , el índice  $C_p$  es igual a 0,691. En este caso, el proceso no tiene capacidad. En el caso de que el límite inferior de deserción fuera 6,0%, el

proceso tendría la capacidad de obtener una deserción máxima de 13,5% con un índice  $C_p = 1,036$ , que permite afirmar que el proceso tiene capacidad pero hay que vigilarlo con cuidado. Se observa con claridad, que el proceso no es capaz de obtener la meta de deserción máxima de 10%, con alguno de los límites inferiores desde 4% hasta 7%. En el mejor de los casos se llegaría a tener una deserción con límite superior de 11,5% y límite inferior de 4% ( $C_p = 1,036$ ). Llegar a pensar que la deserción puede bajar menos de 10%, es una realidad poco probable por el momento. Se considera que el índice que asegura que el proceso cumplirá la meta es  $C_p = 1,312$ , es decir, para este caso, el proceso sería capaz de cumplir la meta de tener una deserción entre 13,5% y 4%. En otras palabras, usar un límite superior o igual a 10%, significa que el proceso no es capaz.

En Mosquera et al [7], se hizo la transformación de Johnson que tiene  $AD = 0,158$  y el resultado fue que el proceso no es capaz de cumplir con la meta propuesta de tener una deserción del 10%, dado que tanto el Ppk como el PPL son iguales a 0,57 y por lo tanto, menores a uno. La transformación de Box – Cox desarrollada en este artículo, no fue concluyente. La aplicación de la distribución de Weibull no presentó un buen ajuste, y entonces, la única conclusión que se puede hacer sobre el proceso de deserción en la Universidad Tecnológica de Pereira, es que la meta de bajar la deserción a 10%, no es posible de cumplir (método de percentiles de Clements), a menos que se hagan esfuerzos mayores: mayor retención de estudiantes con los subsidios de alimentación, subsidios de transporte, descuentos en las matrículas o mayor cantidad de tutorías para estudiantes con bajo rendimiento académico. Debido a que estas alternativas pueden implicar que la Universidad invierta mayor cantidad de recursos económicos y si no los tiene, la mejor opción tal vez sea proponer otra meta de deserción alcanzable.

**Tabla 5.** Metas de deserción para las universidades públicas colombianas

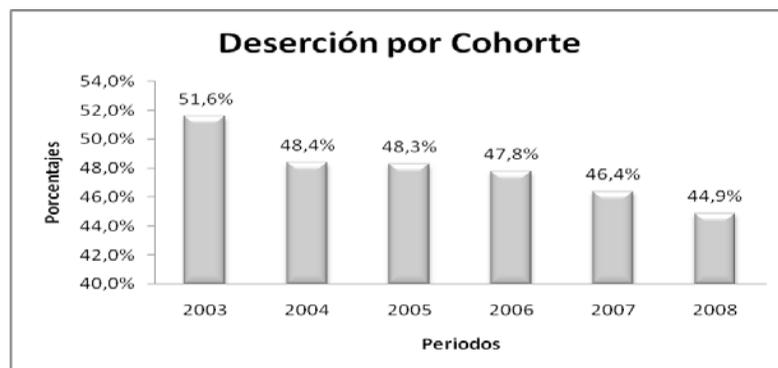
Eje de política: COBERTURA		Ministerio de Educación Nacional	
Nombre del Proyecto: Permanencia Estudiantil en la Educación Superior		República de Colombia	
		Libertad y Orden	
INDICADOR PROYECTO	META CUATRIENIO	META CUMPLIDA 31-DIC-08	META 2009
Tasa de deserción por cohorte 2006: 47.5%	40%	44.9%	43%
Tasa bruta anual de deserción 2006: 12.9%	9%	12.1%	10%
Número de IES con programas apoyados de retención estudiantil	30	11	10
<b>Presupuesto asignado</b>			<b>\$ 3.043</b>

El Gobierno Nacional en apoyo a Instituciones de Educación Superior, tiene como meta disminuir la tasa bruta anual de deserción estudiantil de 2006 que era de 12,9%, y pasar a 9% en 2014 (Tabla 5), según el Ministerio de educación [30].

Esta meta de 9% se propuso para todas las universidades de educación superior que conforman el Sistema de Universidades Estatales (SUE). Como se mostró, para el caso de la UTP, la meta de 9% es difícil de cumplir. Para otras universidades, debe ser tema de revisión a las políticas institucionales de disminución de deserción, análisis de costos, análisis de sus propias causas de deserción y sus propias posibilidades; incluso, tal vez cada universidad debería definir sus propias metas alcanzables de deserción.

## 2.2 Caso 2. Deserción Académica en las Universidades de Colombia

Desde el momento cuando el Ministerio de Educación Nacional de Colombia comenzó a hacer una medición nacional por cohortes de manera regular y homogénea, las tasas de deserción bruta anual se han ubicado entre el 45% y el 50%, lo cual significaría que cerca de uno de cada dos estudiantes que ingresan en el sistema, no culminan sus estudios. Según información publicada por el MEN, se estableció una meta de tener la tasa bruta anual de deserción estudiantil en las universidades del País en el 40% para 2014 [31], y se conoce la deserción acumulada entre los periodos 2003-2008 [32] y [33], con la cual se diseñó la Figura 6. No se encontraron los valores para el periodo del 2009-2010.



**Figura 6.** Deserción por Cohortes para las Universidades Colombianas  
Fuente: Ministerio de Educación

Con los datos obtenidos en el Ministerio de Educación con su programa SPADIES y otras fuentes, se analizó el proceso para establecer si puede cumplir con la meta de tener la deserción al 40%. Teniendo un límite inferior del 25% como meta propuesta por el Gobierno para el 2019 [31]. El primer paso que se debe dar, es determinar si los datos siguen una distribución normal o en el caso contrario la no normalidad de los datos (Tabla 6).

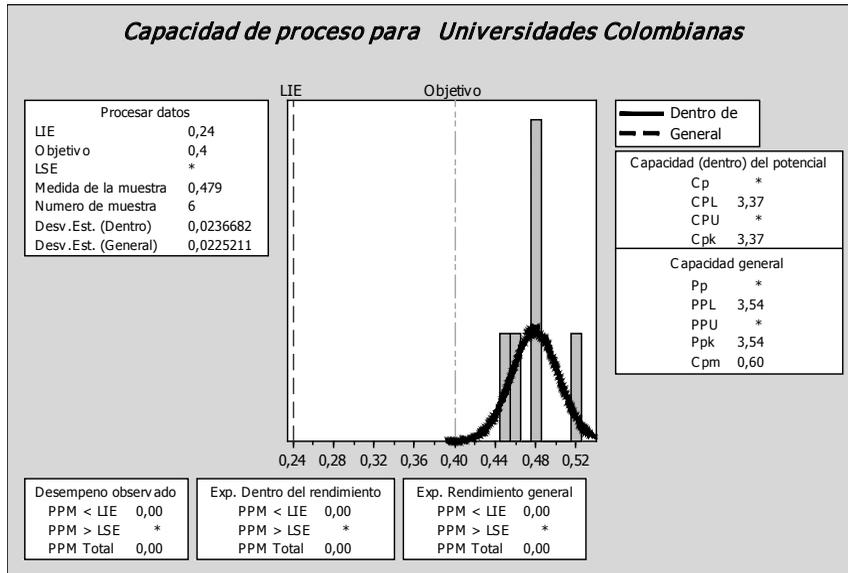
**Tabla 6.** Prueba de Normalidad para las Universidades Colombianas

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig
VAR00001	,245	6	,200*	,952	6	,756

\*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

Al someter los datos a dos pruebas de normalidad, la de Shapiro-Wilk y la de Kolmogorov-Smirnov, se rechaza la hipótesis alterna, y se acepta la normalidad de los datos, y en este último caso específico no es necesario utilizar alguno de los procedimientos para buscar la transformación o ajuste a otra distribución, para analizar los datos bajo la distribución normal, la cual da índices de capacidad a corto y a largo plazo.



**Figura 7.** Capacidad del proceso de Deserción de las Universidades Colombianas

En este caso, se puede afirmar que el proceso de deserción de las universidades colombianas está bajo la capacidad de una deserción académica del 40%, teniendo como límite inferior para este análisis, una deserción del 25% ( $Cpk=3,37$ ;  $Ppk=3,54$ ). El  $Cpm(0,60)$ , indica que el proceso no está centrado en el objetivo, análisis que se puede apoyar también por la Figura 7, donde la curva de normalidad no está centrada en el objetivo de este proceso.

Para el Gobierno colombiano, la deserción de los estudiantes de carreras profesionales para el año 2009, tuvo un costo de \$ 778 mil millones [34]. Como se mostró, la meta de llegar a la deserción por cohorte a 40%, parece ser alcanzable. Con el mejoramiento de este proceso social, no habría una pérdida de capital (costo), tan alto. Esta es la principal fortaleza que tiene la aplicación de los ICP al proceso de deserción en las universidades colombianas: poder definir las metas alcanzables a corto y largo plazo.

La disminución de la deserción de las universidades colombianas, se reflejará directamente en una disminución de pérdidas de capital para el Gobierno. Por otra parte, el gobierno creó un incentivo que se distribuye entre las universidades públicas colombianas según los indicadores SUE. Desde el año 2004, en Colombia se creó un incentivo para las Universidades que integran

el SUE y que estén acreditadas, destinando anualmente un 5% del total de los recursos del artículo 87 de la Ley 30 de 1992. Frente al modelo de indicadores SUE, se miden los resultados de eficiencia del sector en cuatro variables; Formación, Producción, Bienestar y Extensión. El índice de retención estudiantil (1- la deserción), está incorporado dentro de los índices de Bienestar. En 2010, la UTP fue la universidad que ganó mayores recursos adicionales (\$ 688 millones) de acuerdo con los indicadores SUE [35].

### 3. CONCLUSIONES

Mediante la transformación Box-Cox, se obtienen los índices de capacidad del proceso a corto plazo, mientras que por medio de la distribución Weibull, se estructuran los ICP a largo plazo. Los métodos propuestos deben pasar la prueba de bondad del ajuste y seleccionar el mejor ajuste. En caso contrario, se puede concluir sobre los ICP que se presentan de forma errónea, afirmando que un proceso está en capacidad de cumplir con las metas propuestas, cuando en realidad no tiene la capacidad de cumplir el objetivo propuesto.

Como se ha mostrado, es posible implementar diferentes métodos que permitan solucionar el problema de la no – normalidad de los datos de un proceso que serán usados para generar indicadores de capacidad del proceso. En los ejemplos mostrados, la Transformación Box-Cox presenta un mejor ajuste a la distribución normal en comparación con la Distribución Weibull.

Desde el punto de vista estadístico, esta transformación de los datos permite construir de manera correcta, los índices de capacidad de proceso que sin embargo, en el ejemplo analizado, no son suficientes para concluir sobre la eficiencia del proceso para mantenerse dentro de los parámetros de diseño.

Dos metodologías adicionales se pueden usar para solucionar este impase. La primera es realizar un nuevo ajuste con otra transformación, lo cual se realizó exitosamente en [7], o utilizar los métodos para distribuciones no normales, como el método de percentiles de Clements. En ambos casos, se concluye que las metas propuestas son demasiado ambiciosas y deben ser rediseñadas para garantizar que el proceso se mantenga dentro de los límites establecidos.

Para el caso de la deserción en las universidades colombianas, los datos obtenidos no tienen problema de normalidad. En este caso, los ICP encuentran un lugar importante en su aplicación, afirmando que el proceso está en capacidad de cumplir con la meta propuesta por el Gobierno nacional, de tener la deserción nacional en el 40%.

Se puede concluir por lo tanto, que los ICP también pueden ser usados para el control de procesos de deserción (que son procesos sociales), con excelentes resultados para los casos de datos normalmente distribuidos, y para datos no – normales. Con estos ejemplos concretos, se demuestra que es posible utilizar esta herramienta estadística de control de procesos

industriales, para estudiar la capacidad de procesos que pueden denominarse procesos sociales. En particular, la utilización de los ICP permite hacer el seguimiento a los procesos de deserción y fijar metas alcanzables de deserción de corto o largo plazo.

### REFERENCIAS BIBLIOGRÁFICAS

- [1] Sokolov Yuriy. (2010). Capability and Performance. En: [http://www.isixsigma.com/index.php?option=com\\_k2&view=item&id=1479:capability-and-performance&Itemid=199](http://www.isixsigma.com/index.php?option=com_k2&view=item&id=1479:capability-and-performance&Itemid=199) Consultada el 14 de febrero del 2011.
- [2] Paternina C.D., Lerín E., y Márceles E.V., (2001). Un enfoque de redes neuronales hacia el control de calidad para procesos multivariados basados en estimaciones de densidad. En: *Ingeniería & desarrollo*, Vol. 009, Uninorte, pp 78-87.
- [3] Proyecto E-MATH (2002). Uso de las TIC en asignaturas cuantitativas aplicadas. En: [http://www.uoc.edu/in3/emath/docs/SPC\\_6.pdf](http://www.uoc.edu/in3/emath/docs/SPC_6.pdf) Consultada el 1 de marzo de 2011.
- [4] Valdiviezo M.M., y Fermin j., (2009). Evaluación del proceso de laminación del endospermo de maíz mediante índices de capacidad de procesos, En: *Producción y Gestión*, Vol. 12, Versión digital., pp. 47-54.
- [5] Guevara R.D, y Vargas J.A. (2006). Intervalos de confianza los índices de capacidad Cpm y Cpmk en procesos Gaussianos. En: *Revista Colombiana de Estadística*, Vol. 002, pp. 153-162.
- [6] Guevara R.D., and Vargas J.A. (2007). Comparison of Process Capability Indices under Autocorrelated Data. In: *Revista Colombiana de Estadística*, Vol. 002, pp. 301-316.
- [7] Mosquera M.J.C, Mosquera A.JD., y Medina V.P.D (2010). Evaluación del índice de capacidad del proyecto de deserción académica en la universidad tecnológica de Pereira (UTP). En: *Revista en Educación en Ingeniería*. Vol. 009., pp. 96-103.
- [8] Peña S.D.P., y Peña S.J.I, (1986). Un contraste de normalidad basado en la transformación Box-Cox. En: *Revista Estadística Española*, Vol. 110, pp.33-46.
- [9] Valdiviezo Márquez M.R., y Fermín J.S. (2010). Estimación de índices de capacidad de procesos usando la Distribución Generalizada de Pareto. En: *Ingeniería Industrial*, Año 9 N° 2: 93-106.
- [10] Boethe D.R. (2002). Comments on Process Capability Indices-A review 1992-2000 by Johnson and Kotz. En: *Journal of Quality Technology*, Vol.34, p.p.45-50

- [11] Joglekar A.M. (2003). Statistical Methods for Six Sigma. In: R&D and Manufacturing. Wiley Interscience.
- [12] Lamprechet J. (2005). Applied data Analysis for Process Improvement. In: ASQ Quality Press.
- [13] Kotz S., and Johnson N.L. (2002). Process Capability Indices - A review, 1992-2000. (With discussion). In: Journal of Quality Technology, Vol. 34, p.p.32-37
- [14] Pearn W.L., and Kotz S., (2006). Encyclopedia and Handbook of Process Capability Indices. World Scientific.
- [15] Quaglino M. B., Lupachini E., Pagura J.A., y Dianda D. Capacidad de procesos en situaciones no convencionales. En: [http://www.fcecon.unr.edu.ar/investigacion/jornadas/archivos/Quaglino, Lupachini,Pagura,Dianda\\_Capacidad%20de%20procesos.pdf](http://www.fcecon.unr.edu.ar/investigacion/jornadas/archivos/Quaglino,Lupachini,Pagura,Dianda_Capacidad%20de%20procesos.pdf) Consultada 01 noviembre del 2011.
- [16] Clements J.A. (1989). Process Capability Calculations for Non-Normal Distribution. En: Quality Progress, Vol.24 (9), p.p.95-100.
- [17] Muir Alastair (2010). Process capability calculations with non-normal data. En: [http://www.isixsigma.com/index.php?option=com\\_k2&view=item&id=484:process-capability-calculations-with-non-normal-data&Itemid=199#author](http://www.isixsigma.com/index.php?option=com_k2&view=item&id=484:process-capability-calculations-with-non-normal-data&Itemid=199#author) Consultada el 14 de febrero del 2011.
- [18] Universidad Nacional de Colombia. (2010). Introducción a la teoría de probabilidades En: [http://www.virtual.unal.edu.co/cursos/sedes/manizales/4030011/lecciones/cap3/cap\\_3\\_pag\\_14.html](http://www.virtual.unal.edu.co/cursos/sedes/manizales/4030011/lecciones/cap3/cap_3_pag_14.html) Consultada 01 de marzo del 2011.
- [19] Tamborero P.J.M., (2010). Fiabilidad: la distribución de Weibull. En: [http://www.jmcprl.net/ntps/@datos/ntp\\_331.htm](http://www.jmcprl.net/ntps/@datos/ntp_331.htm) Consultada el 10 de abril de 2011.
- [20] Universidad Oberta de Catalunya. (2003). Control Estadístico de la Calidad. En: [http://www.uoc.edu/in3/e-math/docs/SPC\\_6.pdf](http://www.uoc.edu/in3/e-math/docs/SPC_6.pdf) Consultada el 15 de abril de 2010.
- [21] Ministerio de Educación Nacional. Educación Superior. Boletín informativo No 10, febrero 2010. En: [http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articulos-254702\\_boletin\\_14.pdf](http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articulos-254702_boletin_14.pdf) Consultada el 01 de noviembre del 2011.
- [22] Artamónova Carvajal O.P., y Mosquera M.J.C. (2009). Modelos de predicción del rendimiento académico en matemáticas I en la universidad tecnológica de Pereira. En: Scientia et Technica. Vol.43, p.p.258-263.

- [23] Artamónova Carvajal O.P., y Mosquera M.J.C., (2009). Rendimiento en matemáticas. En: *Scientia et Technica*. Vol.41, p.p.379 – 383.
- [24] Carvajal O.P., Trejos C.A.A, y Caro. I.C.A., (2006). Estudio de causas de deserción de los estudiantes de la Universidad Tecnológica de Pereira enero/2000- diciembre/2004 utilizando la técnica de análisis de correspondencias simple. En: *Colombia Scientia Et Technica* Vol.30, p.p.130 – 135.
- [25] Shapiro S.S. and Wilk M.B. (1965). An analysis of variance test for normality (complete samples). In: *Biometrika*, Vol. 52, (3 and 4), p.p.591-611.
- [26] Anderson T. W., Darling D.A. (1952). Asymptotic theory of certain. goodness-of-fit criteria based on stochastic processes. In: *Annals of Mathematical Statistics* Vol.23. p.p.193–212.
- [27] Anderson T.W. and Darling D.A. (1954). A Test of Goodness-of-Fit. In: *Journal of the American Statistical Association*. Vol.49, p.p.765–769.
- [28] Curso de ingeniería de Calidad. (2007). En <http://www.itch.edu.mx/academic/industrial/ingcalidad/unidad2.html> Consultada el 01 de marzo del 2011.
- [29] Marzagão Daniela. (2010). Cp, Cpk, Pp and Ppk: Know How and When to Use them. En: [http://www.isixsigma.com/index.php?option=com\\_k2&view=item&id=760:cp-cpk-pp-and-ppk-know-how-and-when-to-use-them&Itemid=199](http://www.isixsigma.com/index.php?option=com_k2&view=item&id=760:cp-cpk-pp-and-ppk-know-how-and-when-to-use-them&Itemid=199) Consultada el 14 de febrero del 2011.
- [30] Ministerio de Educación Nacional. República de Colombia. En: [http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702\\_visitas\\_ies.pdf](http://www.mineduacion.gov.co/sistemasdeinformacion/1735/articles-254702_visitas_ies.pdf) Consultada el 01 de noviembre del 2011.
- [31] Gobierno Nacional lanza ante los rectores su Política de Educación Superior. (2010). En: <http://noticias.universia.net.co/en-portada/noticia/2010/11/24/750674/gobierno-nacional-lanza-rectores-politica-educacion-superior.html> Consultada el 01 de marzo del 2011.
- [32] Guzmán R.C. (2007). Estrategias para disminuir la deserción en educación superior. Ministerio de Educación. Enero del 2007.
- [33] El observatorio de la universidad Colombiana. (2009). El Gobierno confunde con las cifras de deserción universitaria. En: [http://www.universidad.edu.co/index.php?option=com\\_content&view=article&id=945:el-gobierno-confunde-con-las-cifras-de-desercion-universitaria&catid=16:noticias&Itemid=198](http://www.universidad.edu.co/index.php?option=com_content&view=article&id=945:el-gobierno-confunde-con-las-cifras-de-desercion-universitaria&catid=16:noticias&Itemid=198) Consultada el 01 de marzo del 2011

- [34] Ministerio de educación. (2010). 778 mil millones de pesos le costó a Colombia la deserción universitaria en el 2009. En: <http://www.mineducacion.gov.co/observatorio/1722/article-256223.html> Consultada el 01 de abril del 2011.
- [35] Ministerio de Educación Nacional. Propuesta metodológica para la distribución de recursos artículo 87 de la Ley 30 de 1992 Vigencia 2011. En: <http://www.utp.edu.co/cms-utp/data/bin/UTP/web/uploads/media/comunicaciones/documentos/Distribucion-de-Recursos-Articulo-87-Ley-30-Vigencia-2011.pdf> Consultada el 19 de octubre del 2011.

