



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**Modelo de eficiencia y seguimiento de  
proyectos de software y analítica de datos**

**REPORTE DE TRABAJO  
PROFESIONAL**

**QUE PARA OBTENER EL TÍTULO DE:**

**A C T U A R I A**

**P R E S E N T A :**

**Vania González Jiménez**



**TUTOR:  
Dr. Daniel Trejo Medina  
2019**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



1. Datos del alumno  
González  
Jiménez  
Vania  
55 56 34 04 85  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Actuaría  
309180052
2. Datos del tutor  
Dr  
Daniel  
Trejo  
Medina
3. Datos del sinodal 1  
M en C  
José Salvador  
Zamora  
Muñoz
4. Datos del sinodal 2  
Ing  
César Oswaldo  
Mariscal  
Bello
5. Datos del sinodal 3  
Mat  
Margarita Elvira  
Chávez  
Cano
6. Datos del sinodal 4  
Act  
Mauricio  
Aguilar  
González
7. Datos del trabajo escrito  
Modelo de eficiencia y seguimiento de proyectos de desarrollo de software y  
analítica de datos  
81 p  
2019

## **Agradecimiento y dedicatoria**

Agradezco a los profesores que ayudaron a toda mi formación académica, por su tiempo y enseñanzas compartidas.

Doy gracias a mi tutor y a mis sinodales por su labor al revisar y guiar la conclusión de mi trabajo profesional.

Agradezco a mis padres, familiares y amigos que durante todo mi desarrollo profesional me apoyaron.

Dedico todo mi esfuerzo y este trabajo a mis padres por guiarme durante toda mi vida, desde la instancia académica hasta la personal, por desvelarse conmigo y apoyarme siempre en todo lo que hago.

También dedico este trabajo a mi novio por todo su apoyo durante la carrera en las diversas materias que compartimos y por apoyarme en todo momento.

Finalmente dedico mi escrito al resto de mis familiares y amigos que he conocido a lo largo de mi vida y que han estado conmigo en los buenos y malos momentos.

## Índice

Introducción.....	9
Antecedentes.....	11
Modelos lineales .....	11
Método cualitativo .....	11
Modelo cuantitativo .....	12
Modelos deterministas .....	12
Modelo probabilístico .....	12
Objetivo general .....	14
Objetivo 1 .....	14
Objetivo 2 .....	14
Requerimiento de proyecto.....	15
Análisis de datos .....	15
Regresión lineal .....	17
Objetivo 1:.....	17
Objetivo 2:.....	23
Análisis de correlación.....	27
Objetivo 1 .....	29
Objetivo 2.....	33
Análisis de Pareto .....	35

Aplicación a los datos del proyecto.....	36
Capacidad de procesos.....	40
Índice de capacidad potencial del proceso $C_p$ .....	41
Gráficas del índice $C_p$ y $C_{pk}$ .....	44
Gráficas de los datos del proyecto.....	45
Gráficos de control.....	46
Análisis de puntos críticos.....	48
Gráficos de control del proyecto .....	49
Diseño de solución .....	61
Modelo propuesto .....	61
Modelo de rendimiento .....	62
Modelo de calidad .....	64
Pronóstico de rendimiento.....	66
Pronóstico de calidad .....	66
Softwares candidatos .....	67
Implementación de la solución.....	69
Programación del modelo .....	69
Pruebas del modelo.....	70
Implementación del modelo.....	70
Resultados.....	71

Ejemplo.....	71
Resultado “Modelo de rendimiento” .....	72
Resultado “Modelo de calidad”.....	72
Conclusiones.....	73
Bibliografía.....	76
Anexos .....	78
SCMPI nivel 4 de SVC / CMMI – SVC ML4 .....	78
Mejores prácticas.....	78
Riesgo .....	78
Plan de contingencia.....	79
Plan de mitigación.....	79
Prueba de Anderson-Darling .....	79
Glosario.....	81



## Introducción

La empresa donde se realizó este proyecto es una organización de investigación y desarrollo aplicado con más de trece años de experiencia y presencia en el mercado nacional e internacional, conformada por especialistas en tecnologías de la información y en la integración de soluciones utilizando plataformas tecnológicas de clase mundial. Durante el año 2017 comenzó un proceso de SCAMPI (*Standard CMMI-Based Appraisal Method for Process Improvement*, en español significa “Método estándar de evaluaciones basadas en CMMI para la mejora de procesos”) para nivel 4 de SVC (Services, es decir, dirigido al establecimiento y gestión de servicios), por lo cual se desarrolló un modelo para el seguimiento de procesos basado en herramientas estadísticas.

Las métricas de software (conjunto de medidas como, gráficas, tablas, indicadores, etc. para estimar o conocer alguna característica del software en cuestión) pueden realizarse desde un enfoque actuarial y aportar evidencias con la finalidad de mejorar la calidad de los procesos de una empresa y evitar los riesgos que se presenten en un futuro utilizando técnicas de reducción, supervisión y gestión de riesgos.

Uno de los objetivos principales de la ingeniería de software es la producción de un sistema, aplicación o producto de alta calidad. Para esto es necesaria la implementación de métodos efectivos junto con herramientas modernas para lograr un buen proceso durante el desarrollo del software, de tal forma que se pueda medir su calidad durante dicho proceso.

De acuerdo con las necesidades del negocio, con el modelo final construido se obtiene el pronóstico de éxito de un proyecto y en caso de ser necesario, sugerencias para corregir algún error y que éste se entregue en tiempo y forma al cliente.

Para lograr un óptimo desarrollo del modelo con los requerimientos necesarios para obtener la certificación, se aplicó previamente un análisis tomando en cuenta todos los datos históricos y éstos se trataron bajo herramientas estadísticas con el objetivo de justificar matemáticamente cada resultado y parte del modelo, dichos resultados se presentan durante el desarrollo del trabajo y al final podrá verse el resultado numérico de la aplicación del

modelo a datos similares a los de un proyecto. Cabe mencionar que los resultados expuestos durante todo el trabajo están limitados a los datos permitidos a mostrar por la empresa, debido a políticas de privacidad.

## **Antecedentes**

Los modelos de optimización de procesos con los que la empresa contaba anteriormente al desarrollo del proyecto eran lineales y deterministas, además de utilizar métodos cualitativos; para beneficio del negocio se decidió cambiar a modelos cuantitativos y probabilísticos de corte actuarial. A continuación, se explica cada tipo de modelo mencionado para tener bases claras.

### **Modelos lineales**

Los modelos lineales predicen un objetivo basándose en relaciones lineales entre una variable exógena y una o varias variables endógenas, de ellos puede percibirse la naturaleza de problemas e interpretar soluciones, comunicándolas de manera inteligible a los directivos de la empresa.

Una ventaja de este tipo de modelos es la capacidad para manipularlos computacionalmente en comparación con otros, como redes neuronales, modelos estocásticos, árboles de decisión, etc. de un mismo conjunto de datos.

### **Método cualitativo**

Los métodos cualitativos hacen referencia a las características, es decir, cualidades de aquello que se estudia, por tanto, no pueden obtenerse datos numéricos, lo cual ocasiona una limitante al momento de diseñar modelos o mediciones de algún objetivo ya que, aunque son muy representativos, no pueden ser proyectados.

Las características principales del análisis cualitativo son que se basa en la intuición, se realiza en pequeña escala, no puede cuantificarse y no permite utilizar fundamentos estadísticos.

## **Modelo cuantitativo**

A diferencia del modelo cualitativo, éste se basa en datos numéricos para investigar, analizar, comprobar y hacer mediciones mientras se tenga un número mínimo de datos históricos, es decir, una muestra de tamaño suficiente dependiendo del modelo; a este tipo de modelos se les pueden aplicar métodos estadísticos y probabilísticos para obtener resultados, además, involucra valores desconocidos que pueden pronosticarse.

Algunas características de un modelo cuantitativo son que el análisis se ejecuta relacionando un objetivo con distintas variables numéricas, puede basarse en probabilidades, y va adaptándose a través del tiempo según las condiciones.

## **Modelos deterministas**

En los modelos deterministas se tiene certeza de los datos, es decir, cuando el modelo es analizado se tiene disponible toda la información necesaria para tomar decisiones.

Las buenas decisiones con estos modelos se basan en buenos resultados. Se consigue lo deseado de manera “determinista”, es decir, libre de riesgo ya que no se contempla la existencia del azar ni del principio de incertidumbre (probabilidad). Lo principal es el criterio de la persona que desarrolla el modelo; también depende de la influencia que puedan tener los factores no controlables en la determinación de los resultados de una decisión y de la cantidad de información que se tenga para controlar dichos factores.

## **Modelo probabilístico**

Cuando un modelo probabilístico se muestrea bajo las mismas condiciones en distintas ocasiones, el resultado será diferente, por lo tanto, pueden pronosticarse resultados y entonces tomar decisiones, la cual es una gran ventaja sobre los modelos deterministas.

Al momento de tomar decisiones con base en fenómenos con incertidumbre, puede incorporarse un modelo matemático y manejarse de forma cuantitativa. En este tipo de modelos no se busca una solución óptima ya que los estados futuros de los procesos o

sistemas a los que se aplica nunca serán los mismos, sin embargo, bajo diversos tratamientos estadísticos, inferenciales y experimentales se llega a una solución satisfactoria que generalmente está muy cerca de la solución óptima.

## Objetivo general

El objetivo primordial de la empresa fue alcanzar la certificación “CMMI-SVC ML4” (*Capability Maturity Model Integration for Services - Maturity Level 4*, en español significa “Integración de modelos de madurez de las capacidades para mejorar los procesos de servicio - Nivel de madurez 4”), por consiguiente, el objetivo principal del trabajo fue generar un modelo estadístico para el seguimiento de proyectos.

Como división del objetivo principal, al garantizar un buen seguimiento de los proyectos de la empresa, el proyecto se enfocó en los objetivos más relevantes obtenidos de un análisis de todos los procesos que llevaba a cabo el negocio, identificando cuáles pudieron controlarse estadísticamente y entonces pronosticarse.

### Objetivo 1

Mantener las estimaciones de procesos dentro de los límites establecidos por la Dirección General para evitar desviaciones y ocasionar pérdidas para la empresa. En caso de tener desviaciones tomar acciones respecto a ellas y documentar sus causas, para así evitarlas o pronosticarlas en procesos futuros.

### Objetivo 2

Mantener un servicio de buena calidad, es decir, continuar con los defectos mínimos dentro de los procesos que realiza la empresa para entregar un producto final satisfactorio para el cliente.

## Requerimiento de proyecto

Este apartado tiene como objetivo mostrar el desarrollo para llegar al modelo final solicitado por la empresa, es decir, el análisis realizado, los resultados numéricos, las pruebas aplicadas, gráficas y tablas obtenidas, etc., cada uno con sus respectivas conclusiones e interpretación.

### **Análisis de datos**

Previo a establecer el objetivo general y objetivos del proyecto se tomaron en cuenta varios aspectos importantes.

Se consideraron tres puntos principales:

- Asociación de objetivos de negocio vs. objetivos de medición y obtención de líneas base de dichas mediciones bajo la aprobación de la Dirección General.
- Medición del rendimiento histórico de los procesos y subprocesos seleccionados, estableciendo límites aprobados por la Dirección General.
- Identificación de los atributos del proceso, es decir, qué influye para que funcione adecuadamente.

Fue de suma importancia revisar que cada objetivo fuera SMART: Specific (específico), Measurable (medible), Achievable (alcanzable), Realistic (realista) y Time-bound (con logro calendarizable).

Al seleccionar los objetivos se buscó tener el control de cada uno de ellos y fue necesario, tanto identificar si afectaba a un grupo o área de trabajo específica, como ver la contribución de cada individuo a cada objetivo. Fue relevante también asignarles un nivel de importancia: Crítico, importante o deseable.

Con los objetivos SMART seleccionados, se evaluó su aportación al modelo de pronóstico y se aseguró que cubrieran al menos dos de los siguientes aspectos: Calidad, productividad, desempeño, costo y desviación.

Posteriormente se realizó la asociación de objetivos a cada proceso, y se concluyó que, a mayor cantidad de procesos asociados con un objetivo específico, mayor peso tenía dicho objetivo.

En cuanto al primer objetivo, se produce un “riesgo” cuando se detecta una desviación o alerta de desviación de una estimación. El análisis de riesgos tiene como principal objetivo ayudar al equipo del proyecto a desarrollar una estrategia para tratar las eventualidades desfavorables, por lo cual, una estrategia eficaz debe considerar los siguientes tres aspectos: Evitar el riesgo, supervisar el riesgo y la aplicación de planes de contingencia o mitigación.

Si el equipo de software adopta un enfoque proactivo frente al riesgo, es decir, pronosticar o identificar su aparición, evitarlo es en general la mejor estrategia, esto se consigue desarrollando un plan de reducción de aparición del riesgo. Para reducir el riesgo, la gestión del proyecto debe exponer una estrategia para reducir la movilidad, así a medida que el proyecto va progresando las actividades de supervisión comienzan; el jefe del proyecto supervisa factores que puedan proporcionar un indicador de la probabilidad de presencia del riesgo, además de supervisar la efectividad de los pasos para su reducción.

La gestión del riesgo y los planes de contingencia aceptan que los esfuerzos de reducción pueden fracasar y el riesgo se presente, en cuyo caso se documenta lo sucedido (causas, efectos y acciones aplicadas) para evitarlos o enfrentarlos de la mejor manera en un futuro, creando planes de contingencia y/o mitigación.

En el caso del segundo objetivo, para enfocarse a “calidad” es importante saber que, durante el desarrollo de un proyecto de software, puede existir un juego de factores que afectan de primera instancia hacia la preparación de métricas de la calidad de dicho software, con estos factores se evalúa el proyecto desde tres puntos de vista diferentes:

1. Operación del producto (cuando se utiliza).
2. Revisión del producto (cuando se cambia alguna especificación).
3. Transición del producto (cuando es modificado para que funcione en un entorno diferente).

Dichos factores tienen una relación llamada marco de trabajo, el cual proporciona al administrador del software la capacidad de identificar en el proyecto lo que desde su punto de vista considera importante (facilidad de mantenimiento, atributos del software, corrección y rendimiento funcional, etc.) y provee una evaluación cuantitativa del progreso en el desarrollo de software, teniendo relación con los objetivos de calidad establecidos y facilitando la interacción con el personal de calidad, el cual utiliza indicadores para identificar estándares “mejores” que se podrán verificar en un futuro.

## Regresión lineal

En el análisis de varianza se observa cómo un factor (variable) influye en el resultado de una variable de respuesta. La regresión lineal simple mejora el entendimiento de la dependencia funcional entre una variable y otra. Específicamente el modelo de regresión lineal simple tiene una relación de la siguiente manera:

$$Y_i = \alpha + \beta x_i + \varepsilon_i \quad ;$$

**Fórmula 1. “Ecuación de regresión lineal simple”.**

donde  $Y_i$  es una variable aleatoria y  $x_i$  una variable observable,  $\alpha$  es la intersección y  $\beta$  la pendiente de la regresión que se suponen parámetros fijos y desconocidos, finalmente  $\varepsilon_i$  es una variable aleatoria con distribución  $N(0, \sigma^2)$ .

La representación de las variables en este tipo de ecuación se adaptó a los procesos de la empresa para tener un mejor entendimiento de las métricas y los procesos que se aplicaron en el desarrollo del proyecto, por ello se expresaron las relaciones de regresión lineal simple para cada uno de los objetivos establecidos, dando lugar a modelos y submodelos, donde se especifica la naturaleza de las variables en los modelos y a qué factor se refiere cada una.

### Objetivo 1:

A continuación, se muestran las ecuaciones de regresión lineal simple que representan cada proceso y subproceso relacionados con el primer objetivo.

## Proceso 1

$$Z_1 = \gamma_0 + \gamma_1 Y_1 + \gamma_2 Y_2 + \gamma_3 Y_3 + \varepsilon \quad ;$$

donde:

$Z_1$  = Variable aleatoria 1.  $Z_1$  v. a. continua,  $\mathbb{R}^+/\{0\}$

$Y_1$  = Variable observable 1.1.  $Y_1$  v. a. continua,  $\mathbb{R}^+/\{0\}$

$Y_2$  = Variable observable 1.2.  $Y_2$  v. a. continua,  $\mathbb{R}^+/\{0\}$

$Y_3$  = Variable observable 1.3.  $Y_3$  v. a. continua,  $\mathbb{R}^+/\{0\}$

**Fórmula 2. “Regresión lineal del Proceso 1”. Fuente: Adaptación del autor.**

## Submodelos de pronóstico

### Subproceso 1-1:

$$Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon \quad ;$$

donde:

$Y_1$  = Variable aleatoria 1.1.  $Y_1$  v.a. continua,  $\mathbb{R}^+/\{0\}$

$X_1$  = Variable observable 1.1.1.  $X_1$  v.a. continua,  $\mathbb{R}^+/\{0\}$

$X_2$  = Variable observable 1.1.2.  $X_2$  v.a. continua,  $\mathbb{R}^+/\{0\}$

$X_3$  = Variable observable 1.1.3.  $X_3$  v.a. discreta,  $\mathbb{Z}^+$

$X_4$  = Variable observable 1.1.4.  $X_4$  v.a. discreta,  $\{3, 4, 5, 6, 7\}$

$X_5$  = Variable observable 1.1.5.  $X_5$  v.a. discreta,  $\{1, 2\}$

donde:

1. Tipo 1
2. Tipo 2

**Fórmula 3. “Regresión lineal del Subproceso 1-1”. Fuente: Adaptación del autor.**

### Subproceso 1-2:

$$Y_2 = \zeta_0 + \zeta_1 X_1 + \zeta_2 X_2 + \zeta_3 X_3 + \zeta_4 X_4 + \zeta_5 X_5 + \varepsilon \quad ;$$

donde:

$Y_2$  = Variable aleatoria 1.2.  $Y_1$  v. a. *continua*,  $\mathbb{R}^+/\{0\}$

$X_1$  = Variable observable 1.2.1.  $X_1$  v. a. *continua*,  $\mathbb{R}^+/\{0\}$

$X_2$  = Variable observable 1.2.2.  $X_2$  v. a. *continua*,  $\mathbb{R}^+/\{0\}$

$X_3$  = Variable observable 1.2.3.  $X_3$  v. a. *discreta*,  $\mathbb{Z}^+$

$X_4$  = Variable observable 1.2.4.  $X_4$  v. a. *discreta*,  $\{3, 4, 5, 6, 7\}$

$X_5$  = Variable observable 1.2.5.  $X_5$  v. a. *discreta*,  $\{1, 2\}$

donde:

1. Externo
2. Interno

Fórmula 4. "Regresión lineal del Subproceso 1-2". Fuente: Adaptación del autor.

### Subproceso 1-3:

$$Y_2 = \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \varepsilon \quad ;$$

donde:

$$Y_2 = \text{Variable aleatoria 1.3.} \quad Y_2 \text{ v. a. } \textit{continua}, \mathbb{R}^+/\{0\}$$

$$X_1 = \text{Variable observable 1.3.1.} \quad X_1 \text{ v. a. } \textit{continua}, \mathbb{R}^+/\{0\}$$

$$X_2 = \text{Variable observable 1.3.2.} \quad X_2 \text{ v. a. } \textit{continua}, \mathbb{R}^+/\{0\}$$

Fórmula 5. "Regresión lineal del Subproceso 1-3". Fuente: Adaptación del autor.

### Proceso 2

$$Z_2 = \delta_0 + \delta_1 W_1 + \varepsilon \quad ;$$

donde:

$$Z_2 = \text{Variable aleatoria 2.} \quad Z_2 \text{ v. a. } \textit{continua}, \mathbb{R}^+/\{0\}$$

$$W_1 = \text{Variable observable 2.1.} \quad W_1 \text{ v. a. } \textit{continua}, \mathbb{R}^+/\{0\}$$

Fórmula 6. "Regresión lineal del Proceso 2". Fuente: Adaptación del autor.

## Submodelos de pronóstico

### Subproceso 2-1:

$$W_1 = \alpha_0 + \alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_3 + \alpha_4 V_4 + \alpha_5 V_5 + \alpha_6 V_6 + \alpha_7 V_7 + \varepsilon \quad ;$$

donde:

$W_1$  = Variable aleatoria 2.1.

$W_1$  v. a. *continua*,  $\mathbb{R}^+/\{0\}$

$V_1$  = Variable observable 2.1.1.

$V_1$  v. a. *continua*,  $\mathbb{R}^+/\{0\}$

$V_2$  = Variable observable 2.1.2.

$V_2$  v. a. *continua*,  $\mathbb{R}^+/\{0\}$

$V_3$  = Variable observable 2.1.3.

$V_3$  v. a. *discreta*,  $\{1,2,3\}$

donde:

1. Experto
2. Competente
3. Junior

$V_4$  = Variable observable 2.1.4.

$V_4$  v. a. *discreta*,  $\{1,2,3\}$

donde:

1. Experto
2. Competente
3. Junior

$V_5$  = Variable observable 2.1.5.

$V_5$  v. a. *discreta*,  $\{1,2,3,4,5,6,7\}$

$V_6$  = Variable observable 2.1.6.

$V_6$  v. a. *discreta*,  $\{1,2,3,4\}$

donde:

1. Herramienta 1
2. Herramienta 2
3. Herramienta 3
4. Herramienta 4

**Fórmula 7. “Regresión lineal del Subproceso 2-1. Fuente: Adaptación del autor.**

## Objetivo 2:

Para el segundo objetivo, también se muestran las ecuaciones de regresión lineal simple representativas de cada proceso y subproceso asociados.

### Proceso 3

$$Z_3 = \tau_0 + \tau_1 U_1 + \varepsilon \quad ;$$

donde:

$Z_3$  = Variable aleatoria 3.

$Z_3$  v.a. *continua*, (0, 1)

$U_1$  = Variable observable 3.1.

$U_1$  v.a. *continua*, (0, 1)

**Fórmula 8. “Regresión lineal del Proceso 3”. Fuente: Adaptación del autor.**

## Submodelos de pronóstico

### Subproceso 3-1:

$$U_1 = \varphi_0 + \varphi_1 K_1 + \varphi_2 K_2 + \varphi_3 K_3 + \varphi_4 K_4 + \varepsilon \quad ;$$

donde:

$U_1$  = Variable aleatoria 3.1.  $U_1$  v. a. continua, (0,1)

$K_1$  = Variable observable 3.1.1.  $K_1$  v. a. discreta, {1,2,3}

donde:

1. Experto
2. Competente
3. Junior

$K_2$  = Variable observable 3.1.2.  $K_2$  v. a. discreta, {1,2,3}

donde:

1. Experto
2. Competente
3. Junior

$K_3$  = Variable observable 3.1.3.  $K_3$  v. a. continua, (0,1)

$K_4$  = Variable observable 3.1.4.  $K_4$  v. a. discreta,  $\mathbb{Z}^+$

**Fórmula 9. “Regresión lineal del Subproceso 3-1”. Fuente: Adaptación del autor.**

## Proceso 4

$$Z_4 = \omega_0 + \omega_1 S_1 + \varepsilon \quad ;$$

donde:

$Z_4$  = Variable aleatoria 4.  $Z_4$  v. a. *continua*, (0, 1)

$S_1$  = Variable observable 4.1.  $S_1$  v. a. *continua*, (0, 1)

**Fórmula 10. “Regresión lineal del Proceso 4”. Fuente: Adaptación del autor.**

## Submodelos de pronóstico

### Subproceso 4-1:

$$S_1 = \vartheta_0 + \vartheta_1 L_1 + \vartheta_2 L_2 + \vartheta_3 L_3 + \vartheta_4 L_4 + \vartheta_5 L_5 + \varepsilon \quad ;$$

donde:

$S_1$  = Variable aleatoria 4.1.

$S_1$  v. a. *continua*, (0,1)

$L_1$  = Variable observable 4.1.1.

$L_1$  v. a. *discreta*, {1,2,3}

donde:

1. Experto
2. Competente
3. Junior

$L_2$  = Variable observable 4.1.2.

$L_2$  v. a. *discreta*, {1,2,3,4}

donde:

1. Herramienta 1
2. Herramienta 2
3. Herramienta 3
4. Herramienta 4

$L_3$  = Variable observable 4.1.3.

$L_3$  v. a. *discreta*, {1,2,3,4,5,6,7}

$L_4$  = Variable observable 4.1.4.

$L_4$  v. a. *discreta*, {1,2,3}

donde:

1. Experto
2. Competente
3. Junior

$L_5$  = Variable observable 4.1.5.

$L_5$  v. a. *continua*, (0,1)

**Fórmula 11. “Regresión lineal del Subproceso 4-1. Fuente: Adaptación del autor.**

## Análisis de correlación

El propósito de las ecuaciones de los procesos y subprocesos fue analizar la función de cada uno de ellos, mostrando desde la parte general de cada proceso, es decir, sus actividades, hasta el detalle de cada subproceso, es decir, sus tareas.

Con las variables observables y variables respuesta definidas, el siguiente paso fue realizar un análisis de correlación de las variables en las ecuaciones de subprocesos antes presentadas de acuerdo con los datos históricos numéricos que se tenían de las variables de cada proceso estudiado. Se analizaron los resultados obtenidos de cada análisis de correlación y se documentó su interpretación, para posteriormente decidir conjuntamente con la Dirección General las variables a considerar.

Es importante saber antes de mostrar los análisis de correlación, lo que significa el signo del valor de coeficiente obtenido:

Correlación	Significado
Correlación positiva:	Dos variables tiene correlación positiva cuando al aumentar o disminuir el valor de una de ellas entonces el valor correspondiente a la otra aumentará o disminuirá respectivamente, es decir, cuando las dos variables aumentan en el mismo sentido.
Correlación negativa:	Dos variables tiene correlación negativa cuando al aumentar o disminuir el valor de una de ellas entonces el valor de la otra disminuirá o aumentará respectivamente, es decir, una variable aumenta y otra disminuye o viceversa.

Tabla 1. “Análisis del coeficiente de correlación”. Fuente: Adaptación del autor.

Otro aspecto importante al momento de obtener conclusiones del valor de coeficiente de correlación es saber la interpretación del valor según el rango en el que se encuentre, cuyos rangos fueron definidos de la siguiente manera para poder tomar decisiones sobre las variables a tomar en cuenta para el desarrollo del modelo:

Valor	Significado
-1	Correlación negativa grande y perfecta.
(-0.99,-0.9]	Correlación negativa muy alta.
(-0.9,-0.7]	Correlación negativa alta.
(-0.7,-0.4]	Correlación negativa moderada.
(-0.4,-0.2]	Correlación negativa baja.
(-0.2,-0.01]	Correlación negativa muy baja.
0	Correlación nula.
[0.01,0.2)	Correlación positiva muy baja.
[0.2,0.4)	Correlación positiva baja.
[0.4,0.7)	Correlación positiva moderada.
[0.7,0.9)	Correlación positiva alta.
[0.9,0.99)	Correlación positiva muy alta.
1	Correlación positiva grande y perfecta.

**Tabla 2. “Interpretación del coeficiente de correlación”. Fuente: Adaptación del autor.**

## Objetivo 1

Teniendo conocimiento de la interpretación del valor de coeficiente de correlación, a continuación, se mostrarán los análisis realizados para el primer objetivo.

### Proceso 1

Cada proceso tiene subprocesos, así que fue necesario realizar un análisis de correlación de Pearson por cada subproceso.

#### *Subprocesos*

	Variable aleatoria 1.1.1	Variable aleatoria 1.1.2	Variable aleatoria 1.1.3	Variable aleatoria 1.1.4	Variable aleatoria 1.1.5
Variable aleatoria 1.1.1	1				
Variable aleatoria 1.1.2	0.23145043	1			
Variable aleatoria 1.1.3	-0.96731027	-0.09363350	1		
Variable aleatoria 1.1.4	0.44473315	-0.87941972	-0.49894476	1	
Variable aleatoria 1.1.5	-0.93932698	0.48553759	0.73866421	0.07240384	1

**Tabla 3. “Coeficientes de correlación Subproceso 1-1”. Fuente: Adaptación del autor.**

#### Conclusiones:

- La variable aleatoria 1.1.1 y 1.1.3 están muy altamente correlacionadas de manera negativa, mientras mayor sea el valor de la variable 1.1.1, menos periodos de trabajo tomará la variable 1.1.3.
- La variable aleatoria 1.1.1 y 1.1.5 están muy altamente correlacionadas de manera negativa, si la variable 1.1.5 indica que el proyecto es “Tipo 1”, la variable 1.1.1 tomará valores grandes, por el contrario, si la variable 1.1.5 fuera “Tipo 2”, la variable 1.1.1 tomará valores pequeños.
- La variable aleatoria 1.1.2 y 1.1.4 están altamente correlacionadas de manera negativa, mientras la variable 1.1.2 tome un valor más grande, menos unidades de la variable 1.1.4 serán utilizadas.
- La variable aleatoria 1.1.3 y 1.1.5 están altamente correlacionadas de manera positiva, si la variable 1.1.5 indica que el proyecto es “Tipo 1”, la variable 1.1.3

tendrá menos periodos de trabajo, y viceversa, si la variable 1.1.5 es “Tipo 2”, la variable 1.1.3 toma más periodos de trabajo.

NOTA: El resto de las variables tienen coeficientes de correlación de moderado a muy bajo, por lo tanto, no se tomaron en cuenta para el desarrollo del modelo ya que no muestran gran influencia para el proceso analizado.

	Variable aleatoria 1.2.1	Variable aleatoria 1.2.2	Variable aleatoria 1.2.3	Variable aleatoria 1.2.4	Variable aleatoria 1.2.5
Variable aleatoria 1.2.1	1				
Variable aleatoria 1.2.2	-0.59436294	1			
Variable aleatoria 1.2.3	-0.30000660	-0.68573479	1		
Variable aleatoria 1.2.4	0.72930666	0.04811058	-0.89737132	1	
Variable aleatoria 1.2.5	0.57368442	-0.42780749	0.69660737	0.52041260	1

**Tabla 4. “Coeficientes de correlación Subproceso 1-2”. Fuente: Adaptación del autor.**

Conclusiones:

- La variable aleatoria 1.2.1 y 1.2.4 están altamente correlacionadas de manera positiva, mientras mayor sea el valor de la variable 1.2.1, mayor número de unidades de la variable 1.2.4 se necesitan para el proyecto.
- La variable aleatoria 1.2.3 y 1.2.4 están altamente correlacionadas de manera negativa, conforme más unidades de la variable 1.2.4 se utilicen, menos periodos de trabajo de la variable 1.2.3 serán requeridos.

NOTA: El resto de las variables tienen coeficientes de correlación de moderado a muy bajo, por lo tanto, no se tomaron en cuenta para el desarrollo del modelo ya que no muestran gran influencia para el proceso analizado.

	Variable aleatoria 1.3.1	Variable aleatoria 1.3.2
Variable aleatoria 1.3.1	1	
Variable aleatoria 1.3.2	0.78571535	1

Tabla 5. “Coeficientes de correlación Subproceso 1-3”. Fuente: Adaptación del autor.

Conclusiones:

- La variable aleatoria 1.3.1 y 1.3.2 están altamente correlacionadas de manera positiva, mientras la variable 1.3.1 aumente, mayor será el valor de la variable 1.3.2.

## Proceso 2

El segundo proceso sólo tiene un subproceso, así que sólo fue necesario realizar un análisis de correlación de Pearson.

### *Subprocesos*

	Variable aleatoria 2.1.1	Variable aleatoria 2.1.2	Variable aleatoria 2.1.3	Variable aleatoria 2.1.4	Variable aleatoria 2.1.5	Variable aleatoria 2.1.6
Variable aleatoria 2.1.1	1					
Variable aleatoria 2.1.2	-0.93482868	1				
Variable aleatoria 2.1.3	0	0.48473617	1			
Variable aleatoria 2.1.4	0.26748676	0.04969400	0.66515057	1		
Variable aleatoria 2.1.5	-0.86782273	-0.67393625	-0.77503819	0	1	
Variable aleatoria 2.1.6	0.24575428	0.30585088	0.95990747	0.28930188	-0.39844458	1

Tabla 6. “Coeficientes de correlación Subproceso 2-1”. Fuente: Adaptación del autor.

Conclusiones:

- La variable aleatoria 2.1.1 y 2.1.2 están muy altamente correlacionadas de manera negativa, conforme disminuye la variable 2.1.1, el valor de la variable 2.1.2 será mayor.
- La variable aleatoria 2.1.1 y 2.1.5 están altamente correlacionadas de manera negativa, a mayor valor de la variable 2.1.1, menos revisiones de la variable 2.1.5 habrá.

- La variable aleatoria 2.1.3 y 2.1.5 están altamente correlacionados de manera negativa, mientras más alto sea el perfil de la variable 2.1.3, menos revisiones de la variable 2.1.5 habrá.
- La variable aleatoria 2.1.3 y 2.1.6 están muy altamente correlacionados de manera positiva, mientras más bajo sea el perfil de la variable 2.1.3, menor será el valor de la variable 2.1.6.

NOTA: El resto de las variables tienen coeficientes de correlación de moderado a muy bajo, por lo tanto, no se tomaron en cuenta para el desarrollo del modelo ya que no muestran gran influencia para el proceso analizado.

## Objetivo 2

A continuación, se muestra el análisis de cada subproceso asociado al segundo objetivo, es decir, cada ecuación de regresión lineal planteada en el capítulo anterior.

## Proceso 3

Se presenta el análisis de correlación de Pearson realizado al único subproceso que contiene el tercer proceso estudiado.

### *Subprocesos*

	Variable aleatoria 3.1.1	Variable aleatoria 3.1.2	Variable aleatoria 3.1.3	Variable aleatoria 3.1.4
Variable aleatoria 3.1.1	1			
Variable aleatoria 3.1.2	0.55979150	1		
Variable aleatoria 3.1.3	0.20149192	-0.55937267	1	
Variable aleatoria 3.1.4	0.93672879	-0.74407384	-0.94669341	1

Tabla 7. “Coeficientes de correlación Subproceso 3-1”. Fuente: Adaptación del autor.

### Conclusiones:

- La variable aleatoria 3.1.1 y 3.1.4 están muy altamente correlacionadas de manera positiva, mientras más bajo sea el perfil de la variable 3.1.1, menor será el número de recursos de la variable 3.1.4.
- La variable aleatoria 3.1.2 y 3.1.4 están altamente correlacionadas de manera negativa, mientras el perfil del coordinador de la variable 3.1.2 es más alto, menos recursos de la variable 3.1.4 se requerirán.
- La variable aleatoria 3.1.3 y 3.1.4 están muy altamente correlacionadas de manera negativa, a mayor número de recursos de la variable 3.1.4, menos porcentaje de error de la variable 3.1.3 existirá.

NOTA: El resto de las variables tienen coeficientes de correlación de moderado a muy bajo, por lo tanto, no se tomaron en cuenta para el desarrollo del modelo ya que no muestran gran influencia para el proceso analizado.

## Proceso 4

Finalmente, se presenta el subproceso analizado con correlación de Pearson para el último proceso estudiado.

### *Subprocesos*

	Variable aleatoria 4.1.1	Variable aleatoria 4.1.2	Variable aleatoria 4.1.3	Variable aleatoria 4.1.4	Variable aleatoria 4.1.5
Variable aleatoria 4.1.1	1				
Variable aleatoria 4.1.2	0.40426904	1			
Variable aleatoria 4.1.3	-0.53728056	-0.83510728	1		
Variable aleatoria 4.1.4	0.44988892	0.50206611	0.81733272	1	
Variable aleatoria 4.1.5	0.45398156	-0.91451053	-0.70368522	-0.09014881	1

**Tabla 8. “Coeficientes de correlación Subproceso 4-1”. Fuente: Adaptación del autor.**

### Conclusiones:

- La variable aleatoria 4.1.2 y 4.1.3 están altamente correlacionadas de manera negativa, mientras mayor sea el nivel de la herramienta de la variable 4.1.2, menos revisiones de la variable 4.1.3 se requerirán.
- La variable aleatoria 4.1.2 y 4.1.5 están muy altamente correlacionadas de manera negativa, mientras menor sea el nivel de la herramienta de la variable 4.1.2, más porcentaje de errores de la variable 4.1.5 existirá.
- La variable aleatoria 4.1.3 y 4.1.4 están altamente correlacionadas de manera positiva, mientras más alto sea el perfil del coordinador de la variable 4.1.4, mayor número de revisiones de la variable 4.1.3 habrá.

NOTA: El resto de las variables tienen coeficientes de correlación de moderado a muy bajo, por lo tanto, no se tomaron en cuenta para el desarrollo del modelo ya que no muestran gran influencia para el proceso analizado.

## Análisis de Pareto

El Principio de Pareto o regla del 80/20, muestra el 20% de las causas que generan el 80% de los problemas del objetivo estudiado.

Para realizar un Análisis de Pareto deben seguirse los pasos a continuación:

1. Identificación de problemas.- Se identifican y enlistan los problemas detectados en el proceso, entonces deben realizarse encuestas, lluvia de ideas o algún procedimiento semejante acerca de las causas con los considerados en cada problema. Cabe destacar que todos los problemas investigados deben ser dentro de un mismo periodo de tiempo.
2. Identificación de la causa raíz de cada problema.- Puede suceder que haya causas escritas de distinta manera pero realmente sean la misma, así que es importante agrupar, englobar y generalizar el listado de causas detectadas. Nótese que de ser necesario se puede realizar otra encuesta, lluvia de ideas o procedimiento en caso de no ser suficiente con la información obtenida en el paso anterior.
3. Ordenación de la lista de problemas.- Deben priorizarse los problemas identificados, es decir, ordenar del problema con mayor influencia o importancia sobre el objetivo que se está tratando, hasta el de menor relevancia.
4. Agrupación de problemas según causa raíz.- De acuerdo a las causa raíz detectadas, deben agruparse todos los problemas que tienen la misma causa.
5. Porcentaje.- Se obtiene una columna con el porcentaje de participación de cada causa raíz y se ordena de manera descendente. Posteriormente se incluye una columna con la suma de porcentaje de participación acumulado de cada causa raíz (es importante verificar que al final la suma sea 100%).
6. Desarrollo de la gráfica.- Se grafican los cálculos del paso anterior, identificando el punto donde la gráfica acumule el 80% de los problemas, reflejando el 20% de las causas. La forma de graficar cada columna es:
  - Porcentaje de cada causa raíz.- Histograma.
  - Suma de porcentaje acumulado de cada causa raíz.- Lineal.

7. Tomar acciones.- Con las principales causas identificadas, deben aplicarse estrategias o planes para solucionar los problemas enlistados al principio.

## Aplicación a los datos del proyecto

A continuación, se muestran los análisis de Pareto realizados a cada objetivo del proyecto.

### Objetivo 1

ID	Causa	Frecuencia	Frecuencia Acumulada	Porcentaje	Porcentaje Acumulado
1	Causa A.	52	52	29%	29%
4	Causa B.	46	98	25%	54%
3	Causa C.	45	143	25%	79%
5	Causa D.	10	153	6%	85%
2	Causa E.	9	162	5%	90%
6	Causa F.	7	169	4%	93%
7	Causa G.	6	175	3%	97%
9	Causa H.	4	179	2%	99%
8	Causa I.	1	180	1%	99%
10	Causa J.	1	181	1%	100%
		181		100%	

Tabla 9. “Análisis de Pareto Objetivo 1”.

donde:

Causa A.- Estimación errónea de horas para realizar alguna tarea.

Causa B.- Asignación incorrecta de alguna tarea a cierto recurso.

Causa C.- Subestimación de correcciones a realizar.

Causa D.- Configuración incompleta o con errores.

Causa E.- Mala planeación.

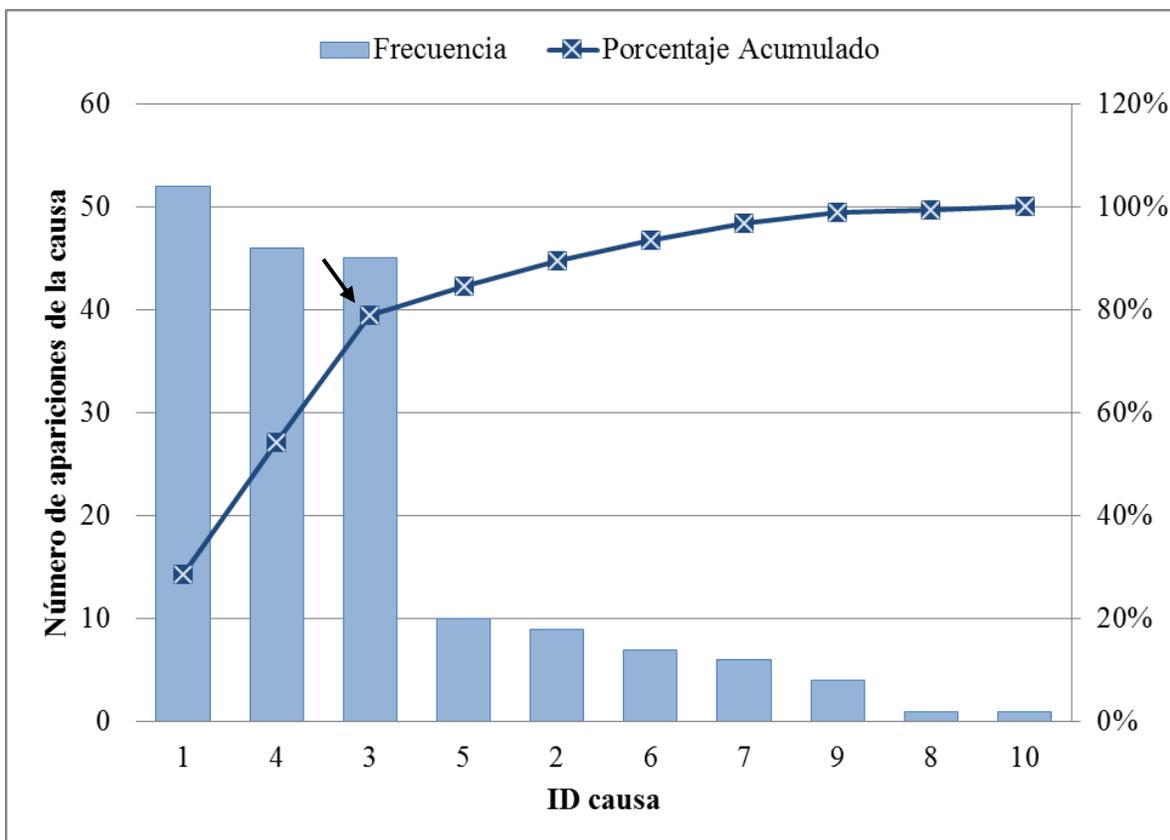
Causa F.- Tarea nueva a realizar.

Causa G.- Nueva tecnología a utilizar.

Causa H.- Versiones viejas o incompatibles de software.

Causa I.- Ajustes no contemplados.

Causa J.- Registro incorrecto por los recursos de sus horas invertidas.



Gráfica 1. "Causas de desviación Objetivo 1".

La flecha negra señala el punto que acumula el 80% de los problemas, ocasionados por el 20% de las causas. Después de analizar cada causa raíz se propuso un plan para evitar sus apariciones futuras.

ID causa	Causa	Propuesta de solución	ID propuesta
1	Causa A.	Solución I.	1
4	Causa B.	Solución II.	2
3	Causa C.	Solución III.	3

Tabla 10. “Propuestas de solución Objetivo 1”.

donde:

Solución I.- Analizar el proceso de estimación de tiempo asignado a cada tarea según su clasificación y recurso asignado.

Solución II.- Crear o revisar la matriz de capacidades de recursos de acuerdo a la clasificación de cada tarea para tomar la mejor decisión al momento de la asignación.

Solución III.- Analizar el proceso de pronóstico de correcciones para mejorarlo, de tal manera que el pronóstico tenga un margen de error mínimo.

## Objetivo 2

ID	Causa	Frecuencia	Frecuencia Acumulada	Porcentaje	Porcentaje Acumulado
6	Causa A.	12	12	27%	27%
4	Causa B.	12	24	27%	55%
8	Causa C.	11	35	25%	80%
1	Causa D.	4	39	9%	89%
5	Causa E.	2	41	5%	93%
7	Causa F.	1	42	2%	95%
3	Causa G.	1	43	2%	98%
2	Causa H.	1	44	2%	100%
		44		100%	

Tabla 11. “Análisis de Pareto Objetivo 2”.

donde:

Causa A.- Cambios en las especificaciones.

Causa B.- Falta de reporte de un error para su corrección.

Causa C.- Falta de ejecución de alguna actividad al momento de la revisión del producto.

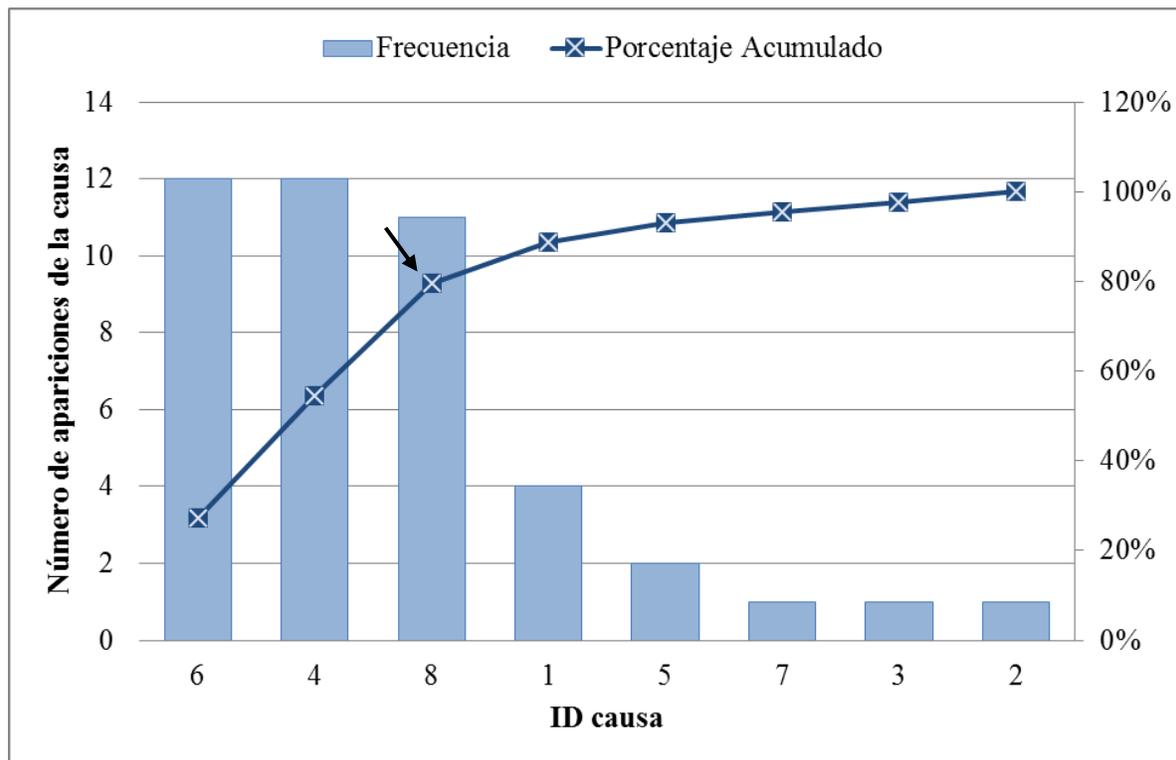
Causa D.- Requerimientos incompletos.

Causa E.- Equipo de cómputo del cliente no compatible con algún software a utilizar.

Causa F.- Cliente nuevo.

Causa G.- Cliente difícil de tratar.

Causa H.- El cliente no quedó satisfecho con el producto final.



Gráfica 2. "Causas de desviación Objetivo 2".

La flecha negra señala el punto que acumula el 80% de los problemas, ocasionados por el 20% de las causas. Después de analizar cada causa raíz se propuso un plan para evitar sus apariciones futuras.

ID causa	Causa	Propuesta de solución	ID propuesta
6	Causa A.	Solución I.	1
4	Causa B.	Solución II.	2
8	Causa C.	Solución III.	3

Tabla 12. “Propuestas de solución Objetivo 2”.

donde:

Solución I.- Analizar el proceso de manejo y/o clasificación de cambios.

Solución II.- Agregar a la lista de funciones a revisar una liga que envíe al sistema de reporte de errores.

Solución III.- Capacitar a los recursos encargados de la revisión del producto.

## Capacidad de procesos

“Una necesidad frecuente en los procesos consiste en evaluar la variabilidad y tendencia central de una característica de calidad, para así compararla con sus especificaciones de diseño. La capacidad de proceso es el grado de aptitud que tiene un proceso para cumplir con las especificaciones técnicas deseadas.” (“Capacidad de Proceso”, 2017)

Si se encuentra que “la capacidad de un proceso es alta, entonces se dice que el proceso es capaz, si se encuentra estable a lo largo del tiempo, se dice que el proceso está bajo control, en otro caso el proceso no es adecuado para el trabajo o requiere de modificaciones inmediatas” (“Capacidad de Proceso”, 2017).

Si los procesos no sufren modificaciones o reajustes, la forma de evaluar su capacidad es mediante histogramas, gráficos de control, planillas de inspección, entre otras herramientas, sin embargo, si el proceso se ve modificado, (reajuste de métodos, cambio de parámetros, reacomodo de recursos, etc.) entonces debe recurrirse a un estudio de índices de capacidad.

Para poder realizar un estudio de índices de capacidad se requiere que el proceso se encuentre estadísticamente estable, además de que el comportamiento de las mediciones individuales del proceso debe ajustarse a una distribución normal.

## Índice de capacidad potencial del proceso $C_p$

“Para considerar que un producto sea de calidad, las mediciones de sus características deben ser iguales a su valor ideal” (“Capacidad de Proceso”, 2017), sin embargo, estas mediciones deben estar al menos dentro de un intervalo dependiendo de las metas establecidas. “La medida de capacidad potencial del proceso para cumplir con estas especificaciones de calidad es proporcionada por el “índice de capacidad potencial  $C_p$ ”.” (“Capacidad de Proceso”, 2017) Este índice al no tener unidades específicas, permite comparar dos procesos completamente diferentes, por lo tanto, para poder realizar una estimación de capacidad del proceso es necesario que este se encuentre bajo control.

“El índice  $C_p$  compara el ancho de las especificaciones (tolerancia) con la amplitud de la variación del proceso. Si la variación del proceso es mayor que la amplitud de las especificaciones, entonces el índice  $C_p$  es menor que 1, evidencia de que no se está cumpliendo con las especificaciones; si el índice  $C_p$  es mayor que 1, hay evidencia de que el proceso es potencialmente capaz de cumplir con las especificaciones” (“Capacidad de Proceso”, 2017).

El índice de capacidad potencial del proceso se define por la siguiente fórmula:

$$C_p = \frac{E_s - E_i}{6\sigma} \quad ;$$

donde:

$E_s$ : Especificación superior. Medida de calidad que indica la desviación tolerada.

$E_i$ : Especificación inferior.  $-E_s$  por la simetría de la distribución normal.

$\sigma$ : Desviación estándar.

Fórmula 12. “Índice de capacidad potencial”. Fuente: Adaptación del autor.

El enfoque del índice  $C_p$  se centra en reducir o eliminar la variación en un proceso. Frente a una variación lo primordial es entender la causa del por qué sucede esto. Matemáticamente significa que dentro de un proceso pueden haber 6 desviaciones estándar entre el promedio y la especificación, lo cual hace que la variación sea diminuta. Este índice es utilizado para asegurar la calidad en cada puesto de trabajo y mejorar significativamente la calidad de los productos y/o servicios, es por ello que cuando se quiere reducir la variabilidad en los procesos, es decir, mejorar el nivel del cumplimiento de las especificaciones si presentan una variación que se ha salido de control, se utiliza dicho enfoque.

El índice  $C_p$  se utiliza para conocer la capacidad del proceso y tomar decisiones, dependiendo de su valor es la conclusión y posición que debe tomarse.

Valor de $C_p$	Interpretación y Acción
$C_p > 2$	Se tiene calidad $6\sigma$ , es decir, la variación es diminuta, así que no es necesario hacer modificaciones.
$1.33 \leq C_p \leq 2$	Más que adecuado, no es necesario modificar algo.
$1 \leq C_p < 1.33$	Adecuado para el trabajo, pero requiere de un control estricto conforme el $C_p$ se acerca a 1.
$0.67 \leq C_p < 1$	No adecuado para el trabajo, es necesario un análisis del proceso, realizando modificaciones serias para alcanzar una calidad satisfactoria.
$C_p < 0.67$	No adecuado para el trabajo, requiere modificaciones serias.

Tabla 13. "Interpretación del índice  $C_p$ ". Fuente: Adaptación de <https://www.ingenieriaindustrialonline.com/herramientas-para-el-ingeniero-industrial/gestión-y-control-de-calidad/capacidad-de-proceso/>

## Índice de capacidad real del proceso $C_{pk}$

“El índice  $C_p$  estima la capacidad potencial del proceso para satisfacer una tolerancia, sin embargo, una de sus desventajas es que no toma en cuenta el centrado del proceso, una solución a este problema es modificar el índice  $C_p$  de tal forma que se pueda evaluar también dónde se localiza la media del proceso respecto a las especificaciones, a esta modificación se le conoce como índice de capacidad real  $C_{pk}$ .” (“Capacidad de Proceso”, 2017)

El índice de capacidad real del proceso se define por la siguiente fórmula:

$$C_{pk} = \min \left\{ \frac{\mu - E_i}{3\sigma}, \frac{E_s - \mu}{3\sigma} \right\} ;$$

donde:

$\mu$ : Media o promedio. Respecto a la característica de calidad.

$E_s$ : Especificación superior. Medida de calidad que indica la desviación tolerada.

$E_i$ : Especificación inferior. -  $E_s$  por la simetría de la distribución normal.

$\sigma$ : Desviación estándar.

**Fórmula 13. “Índice de capacidad real”. Fuente: Adaptación del autor.**

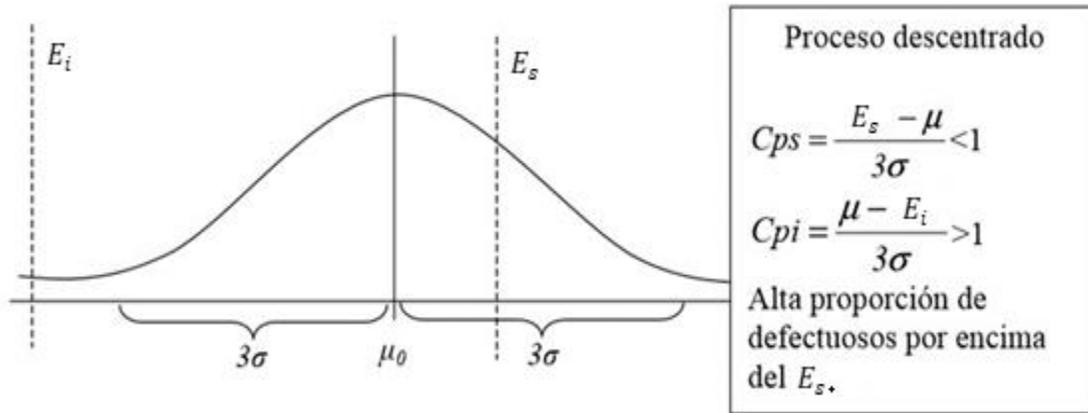
“El índice  $C_{pk}$  será igual al  $C_p$  cuando la media del proceso se ubique en el punto medio de las especificaciones, si el proceso no está centrado entonces el valor del índice de  $C_{pk}$  será menor que el  $C_p$ .” (“Capacidad de Proceso”, 2017) La interpretación del índice  $C_{pk}$  depende de su resultado.

Valor de $C_{pk}$	Interpretación
$C_{pk} > 1$	El proceso se orienta más al objetivo y tiene menos unidades defectuosas.
$C_{pk} = 1$	La variación del proceso está centrada dentro de los límites superior e inferior de especificaciones.
$C_{pk} < 1$	El proceso no cumple con las especificaciones.

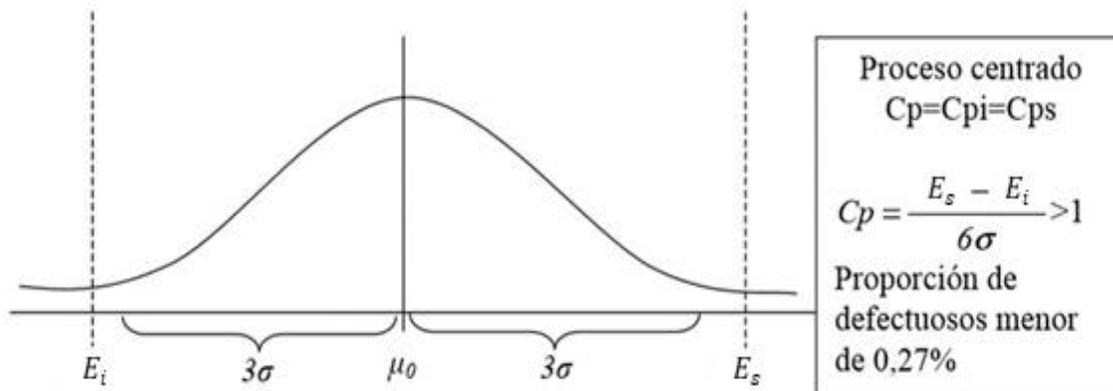
**Tabla 14. “Interpretación del índice  $C_{pk}$ ”. Fuente: Adaptación de <https://www.ingenieriaindustrialonline.com/herramientas-para-el-ingeniero-industrial/gestión-y-control-de-calidad/capacidad-de-proceso/>**

## Gráficas del índice $C_p$ y $C_{pk}$

Gráficamente se puede ver el comportamiento del índice  $C_p$  y  $C_{pk}$  de la siguiente forma:



Gráfica 3. “Ejemplo de proceso descentrado”. Fuente: Adaptación de <http://navarrof.orgfree.com/Docencia/Calidad/UT4/graficoscontrol2.pdf>

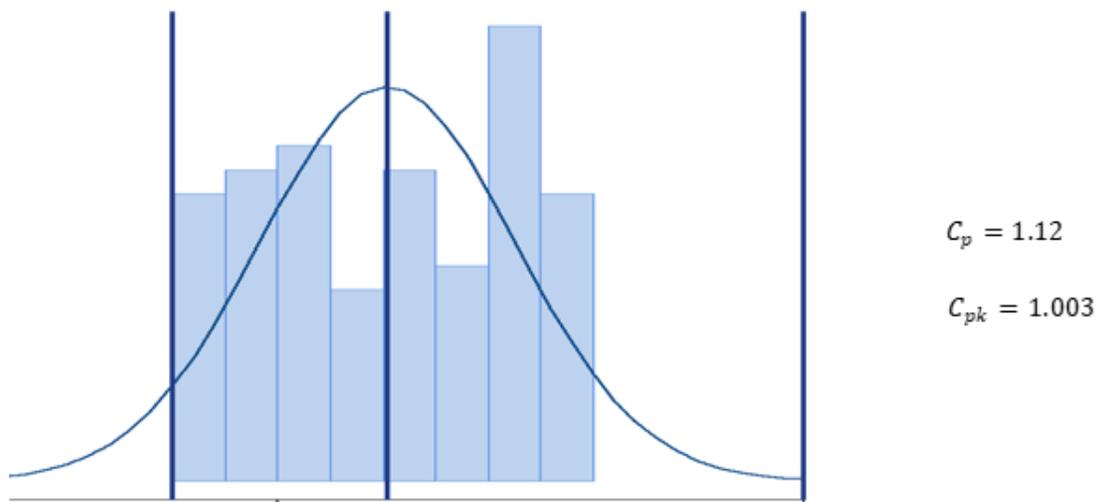


Gráfica 4. “Ejemplo de proceso centrado”. Fuente: Adaptación de <http://navarrof.orgfree.com/Docencia/Calidad/UT4/graficoscontrol2.pdf>

## Gráficas de los datos del proyecto

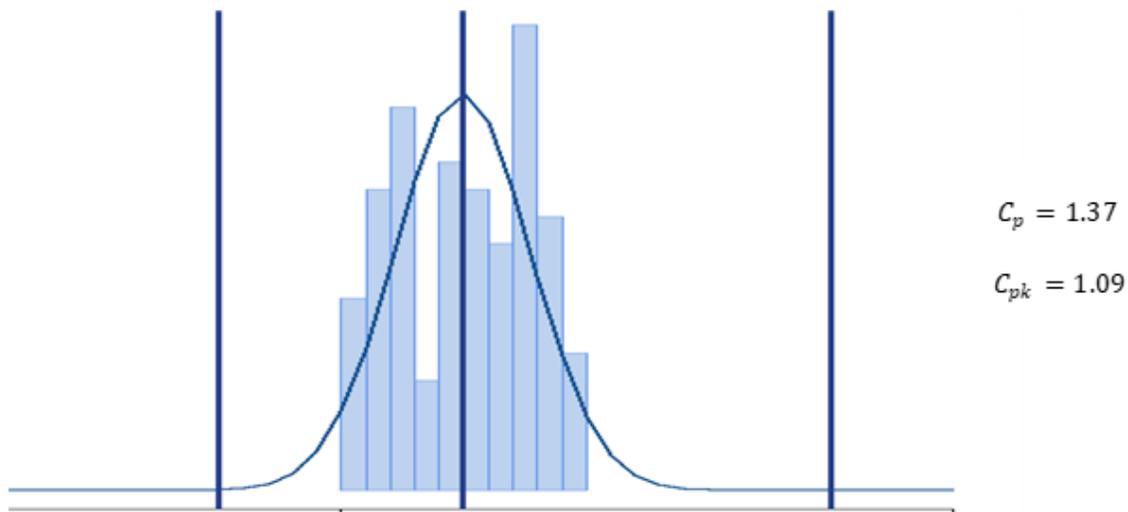
En esta sección se muestran las gráficas e índices de  $C_p$  y  $C_{pk}$  obtenidos de los datos de la empresa para cada objetivo estudiado.

### Objetivo 1



Gráfica 5. “ $C_p$  y  $C_{pk}$  – Objetivo 1”. Fuente: Adaptación del autor.

### Objetivo 2



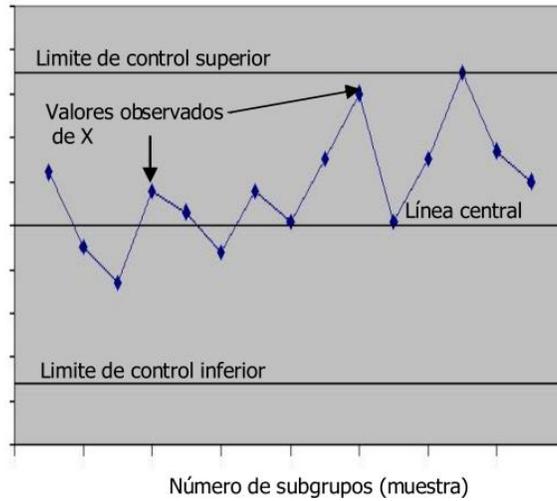
Gráfica 6. “ $C_p$  y  $C_{pk}$  – Objetivo 2”. Fuente: Adaptación del autor.

## **Gráficos de control**

Cuando una empresa decide incluir una métrica en sus procesos, se pueden encontrar grandes beneficios como poder evaluar el rendimiento general del proceso ya que las compañías evalúan los productos descubriendo sus capacidades, midiendo su efectividad e identificando brechas de rendimiento, dan prioridad a la inversión de mejora, monitorean a la industria con mejores prácticas y enriquecen con fiabilidad operacional definiendo claramente sus métricas en términos de objetivos organizacionales.

Todos los beneficios pueden ser expresados mediante gráficos de control, los cuales se utilizan como técnica de diagnóstico para supervisar procesos e identificar alguna inestabilidad y circunstancias anormales, comparando de manera gráfica los datos de desempeño del proceso con los “límites de control estadístico” dibujados como rectas limitantes sobre la gráfica. Cuando una gráfica indica una situación fuera de control se inicia una investigación para identificar las causas y tomar medidas de corrección.

El objetivo de un gráfico de control no es lograr un estado de control estadístico, sino reducir la variación. Un elemento básico de las gráficas de control es que las muestras del proceso de interés se seleccionan a lo largo de una sucesión de puntos en el tiempo; si todos los puntos de la gráfica se encuentran entre los dos límites de control, entonces se considera que el proceso ha sido controlado, una señal fuera de control aparece cuando un punto trazado cae fuera de dichos límites, lo cual se atribuye a alguna causa asignable y de esa manera se comienza su búsqueda.



Gráfica 7. "Ejemplo de gráfica de control". Fuente: Adaptación de [http://www.calidad.com.mx/docs/art\\_88\\_1.pdf](http://www.calidad.com.mx/docs/art_88_1.pdf)

Cabe mencionar que los límites de control pueden ser determinados de acuerdo a lo que requiera la empresa o de forma matemática, pues éstos se adaptan a las necesidades.

## Análisis de puntos críticos

En los gráficos de control, pueden identificarse puntos críticos existentes, los cuales hacen referencia a cierto comportamiento o a puntos que se salen de los límites de control establecidos. Existen diversos tipos de puntos críticos, dependiendo de su comportamiento éstos serán clasificados, de tal manera que al identificarlos se apliquen acciones correctivas y se aborden con una nueva estrategia o con un plan de mitigación previamente determinado basado en experiencias pasadas.

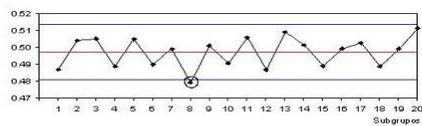
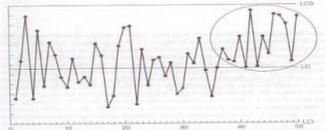
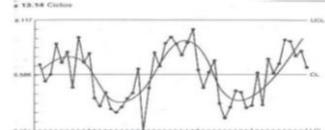
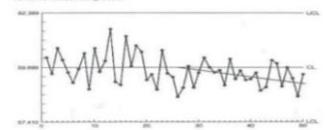
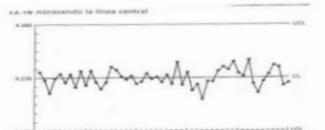
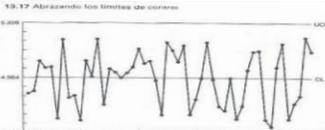
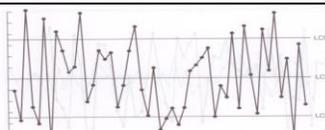
PUNTOS CRÍTICOS	EXPLICACIÓN	EXPLICACIÓN ILUSTRATIVA
Punto fuera de control	Punto atípico, es decir, punto único que se encuentra fuera de los límites de control.	 Un gráfico de control con 20 subgrupos. El eje vertical muestra valores entre 0.47 y 0.52. Una línea central está en 0.50. Hay líneas de control superior y inferior. El punto 8 está marcado con un círculo y una 'X' roja, indicando un punto fuera de control.
Corrida	Cuando varios puntos caen consecutivamente entre la línea central y uno de los límites de control.	 Un gráfico de control con 30 puntos. Una línea central está en 0.50. Hay líneas de control superior y inferior. Una zona de los datos está circunscrita con un círculo rojo, mostrando una serie de puntos que se acercan repetidamente a uno de los límites de control.
Ciclo	Es cuando se repite una misma figura en el comportamiento del gráfico varias veces, es decir, siguen cierto patrón.	 Un gráfico de control con 30 puntos. Una línea central está en 0.50. Hay líneas de control superior y inferior. El gráfico muestra un patrón claro de oscilación que se repite a lo largo de los datos.
Tendencia	Resultado de algún evento que va afectando gradualmente el comportamiento de los datos, de tal manera que se observa que la gráfica va hacia arriba o hacia abajo.	 Un gráfico de control con 30 puntos. Una línea central está en 0.50. Hay líneas de control superior y inferior. El gráfico muestra una clara tendencia ascendente o descendente a lo largo de los datos.
Abrazando la línea central	Ocurre cuando casi todos los puntos caen relativamente cerca de la línea central.	 Un gráfico de control con 30 puntos. Una línea central está en 0.50. Hay líneas de control superior y inferior. La mayoría de los puntos están muy cercanos a la línea central.
Abrazando los límites de control	Cuando la mayoría de los puntos están muy cercanos a los límites de control pero sin rebasarlos.	 Un gráfico de control con 30 puntos. Una línea central está en 0.50. Hay líneas de control superior y inferior. La mayoría de los puntos están muy cercanos a uno de los límites de control.
Inestabilidad	Fluctuaciones de los datos muy dispersas, caen dentro de los límites centrales o rebasándolos, después cerca de la línea central, etc, es decir, no siguen un patrón consistente.	 Un gráfico de control con 30 puntos. Una línea central está en 0.50. Hay líneas de control superior y inferior. El gráfico muestra fluctuaciones muy dispersas y no consistentes.

Tabla 15. “Análisis de patrones”. Fuente: Adaptación del autor.

## Gráficos de control del proyecto

Se presentan las gráficas obtenidas de los datos del proyecto para conocer la estabilidad de los procesos estudiados.

### Objetivo 1

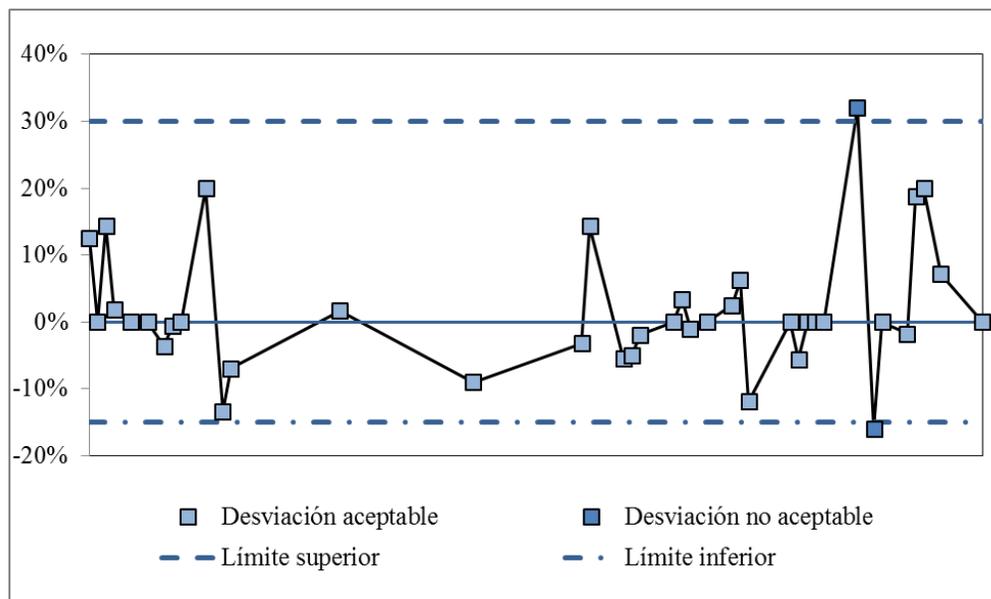
El gráfico de control mide el comportamiento de una variable a través del tiempo, así que fue necesario obtener gráficas de cada proceso del primer objetivo.

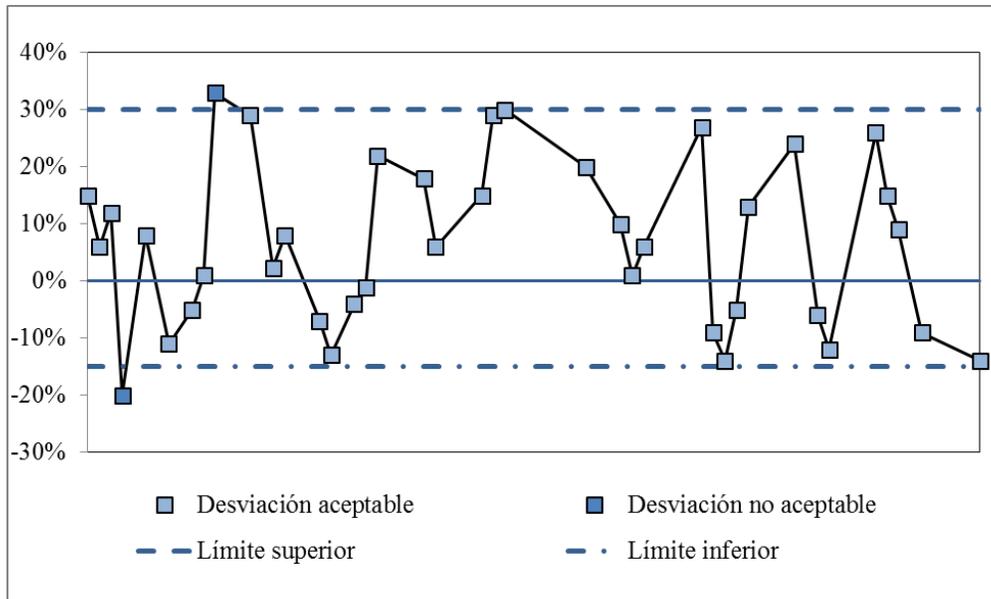
De las variables con mayor coeficiente de correlación analizadas anteriormente para cada subproceso, se acordó con la Dirección General las que se tomaron en cuenta.

#### Proceso 1

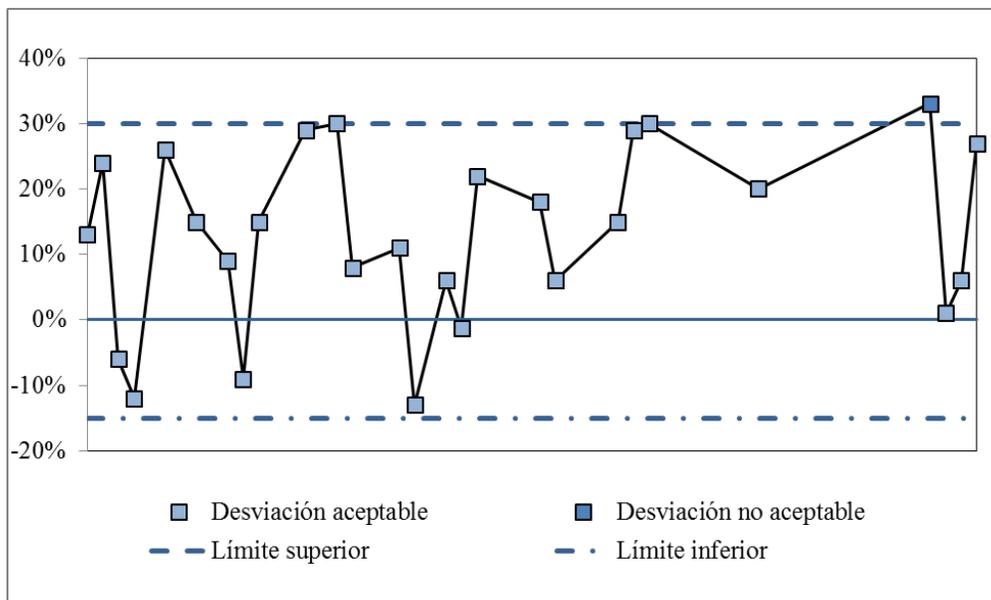
Se sabe que el primer proceso tiene varios subprocesos asociados y cada uno cuenta con tareas diferentes, por lo tanto, fue necesario obtener varios gráficos.

#### Subproceso 1-1





**Gráfica 9. “Gráfica de control – Variable 1.1.3”. Fuente: Adaptación del autor.**



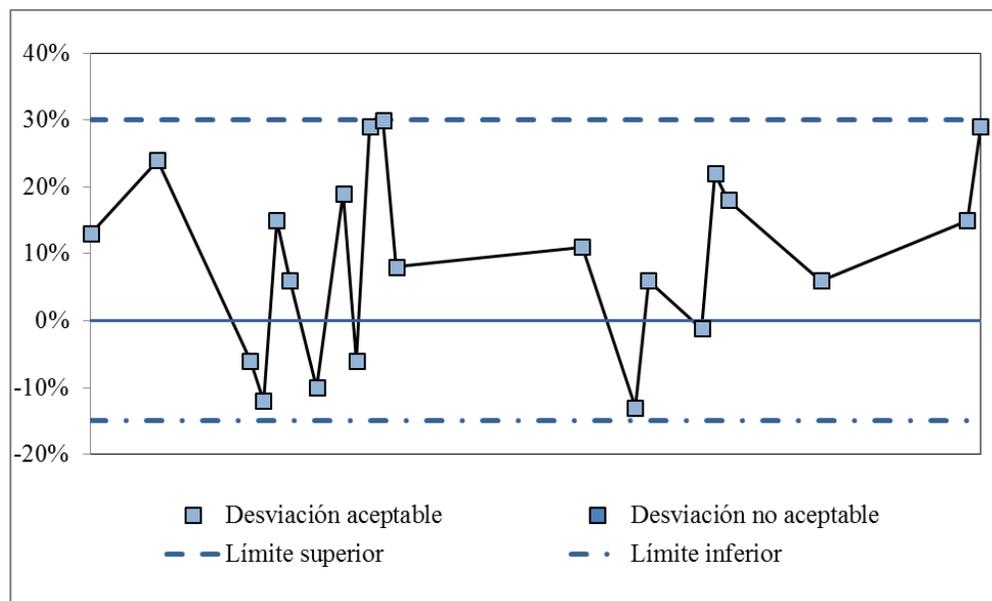
**Gráfica 10. “Gráfica de control – Variable 1.1.5”. Fuente: Adaptación del autor.**

En la “Gráfica 8” se observan dos puntos fuera de control seguidos al final del gráfico y luego se mantienen dentro de los límites de control, lo cual indica que la desviación se corrigió.

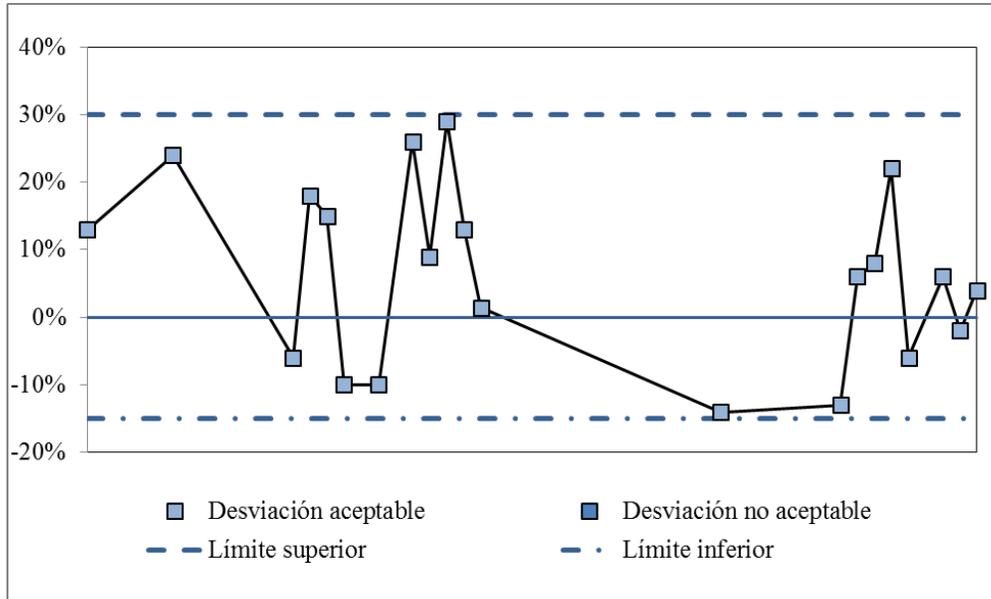
Se muestran dos puntos fuera de los límites de control al principio de la “Gráfica 9” pero después se mantienen dentro de ellos. Al centro del gráfico de control hay una corrida, lo cual indica que hubo sobreestimación de tiempo para las tareas, pues se dedicó menos de lo esperado, así que hay que revisar dichas tareas.

Finalmente, en la “Gráfica 10” hay una tendencia hacia arriba casi al final del gráfico de control, es decir, se pasa de subestimar a sobreestimar, por lo tanto, hay que cuidar que no continúe dicha tendencia. Se tiene un punto fuera de control, sin embargo, está cerca del límite, así que se toma como un punto atípico.

### Subproceso 1-2



Gráfica 11. “Gráfica de control – Variable 1.2.3”. Fuente: Adaptación del autor.

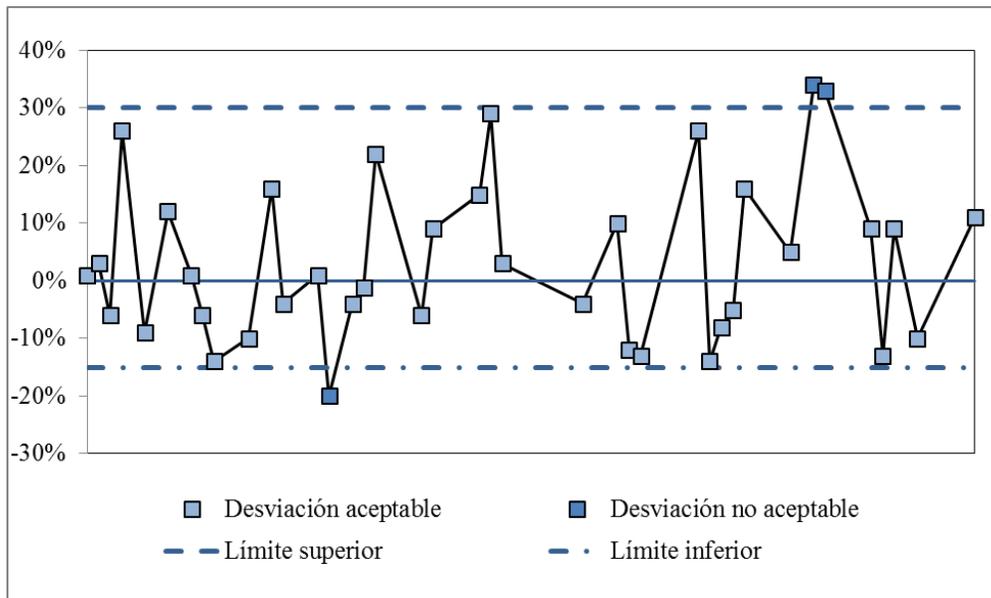


Gráfica 12. “Gráfica de control – Variable 1.2.4”. Fuente: Adaptación del autor.

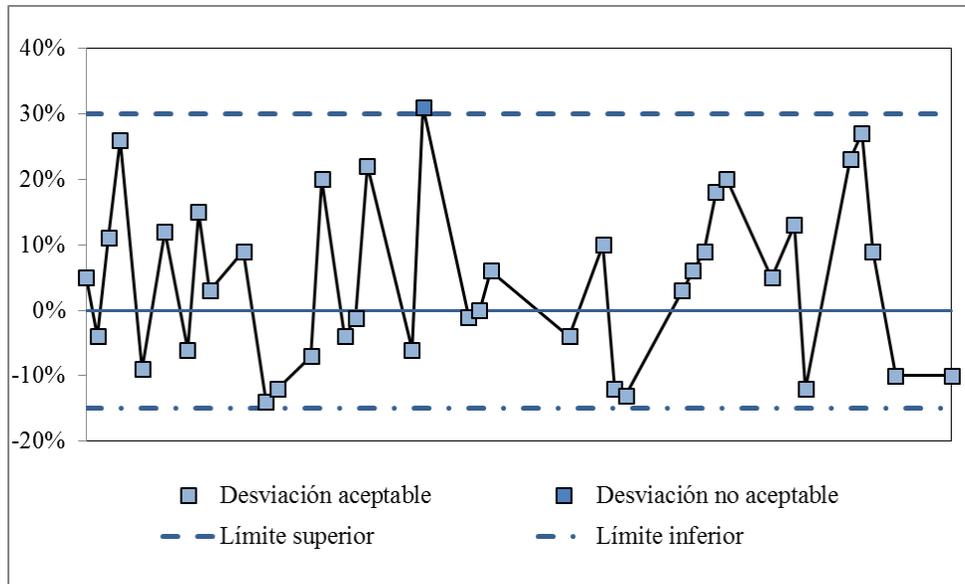
Se observan dos tendencias hacia arriba en la “Gráfica 11” y ningún punto fuera de los límites de control, así que no es necesario tomar acciones.

La “Gráfica 12” no muestra puntos fuera de control ni algún comportamiento específico.

### Subproceso 1-3



Gráfica 13. “Gráfica de control – Variable 1.3.1”. Fuente: Adaptación del autor.



**Gráfica 14. “Gráfica de control – Variable 1.3.2”. Fuente: Adaptación del autor.**

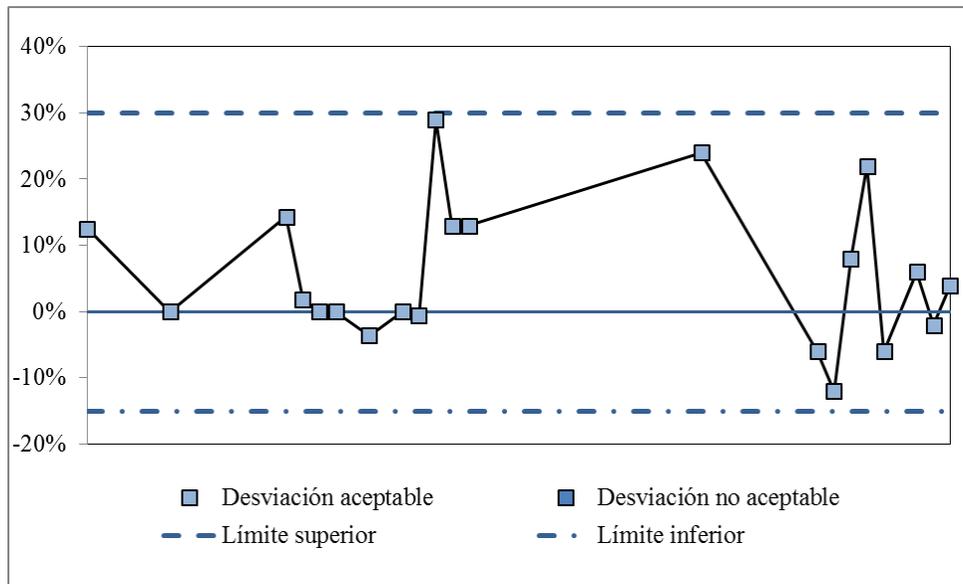
Los puntos de la “Gráfica 13” abrazan los límites de control y hay tres puntos fuera de dichos límites, debe revisarse la subestimación y la sobreestimación para tareas futuras, y cuidar que no sigan saliendo los puntos de desviación de tareas de los límites al estar tan cerca de éstos.

En la “Gráfica 14” se observa mucha inestabilidad, pero sólo un punto fuera de control cerca del límite superior, por lo tanto, se clasifica como un punto atípico, sólo hay que vigilar el subproceso.

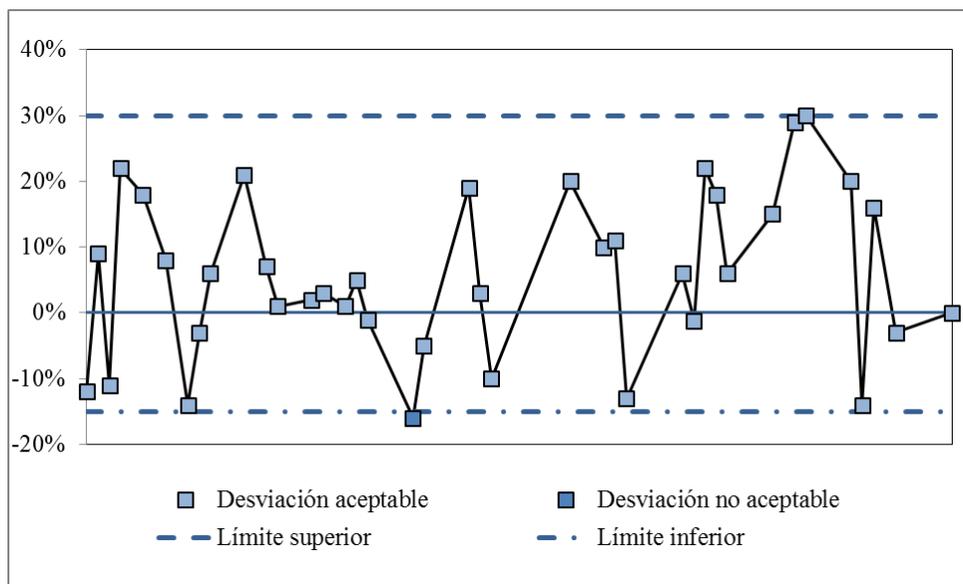
## Proceso 2

Se obtuvo un gráfico de control por cada variable asociada al subproceso del segundo proceso estudiado.

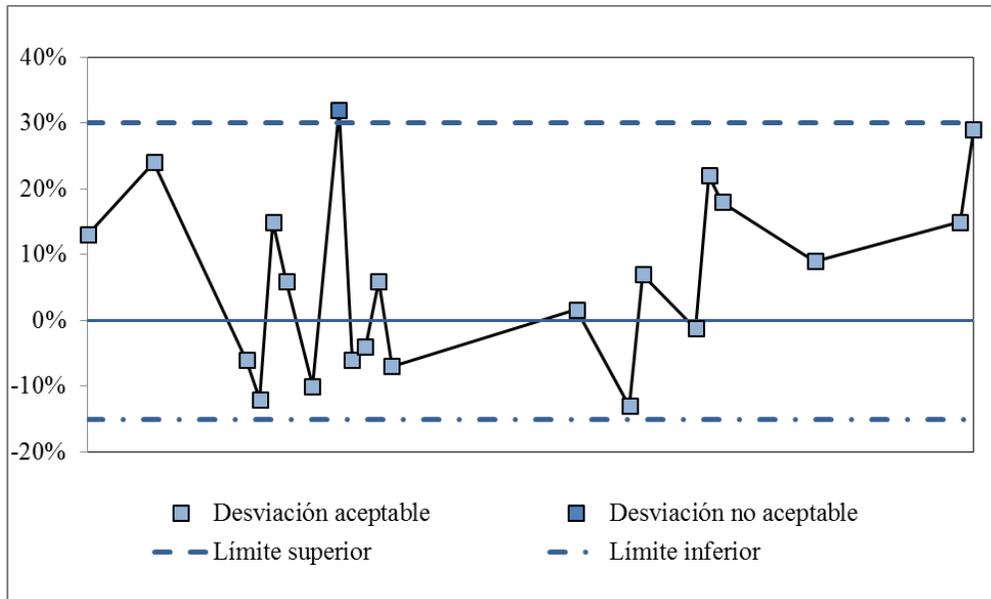
### Subproceso 2-1



Gráfica 15. "Gráfica de control – Variable 2.1.1". Fuente: Adaptación del autor.



Gráfica 16. "Gráfica de control – Variable 2.1.3". Fuente: Adaptación del autor.



**Gráfica 17. “Gráfica de control – Variable 2.1.5”. Fuente: Adaptación del autor.**

La “Gráfica 15” no muestra algún comportamiento específico y todos los puntos están dentro de los límites de control.

Se observa inestabilidad en la “Gráfica 16” y un punto superó por poco el límite inferior, es necesario vigilar el subproceso para evitar más puntos fuera de control.

Finalmente, en la “Gráfica 17” hay un punto fuera del límite superior y de la mitad del gráfico de control al final se observa una tendencia hacia arriba.

## Objetivo 2

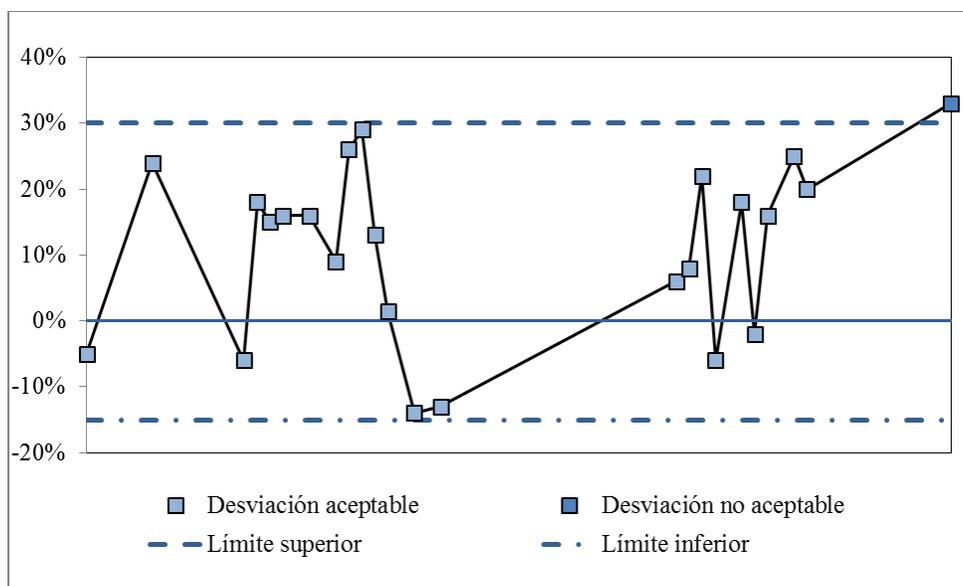
Se obtuvieron varios gráficos de control de cada proceso del segundo objetivo que, se muestran a continuación.

Es importante mencionar que, de acuerdo con las correlaciones obtenidas anteriormente por subproceso, se acordó con la Dirección General las que se tomaron en cuenta.

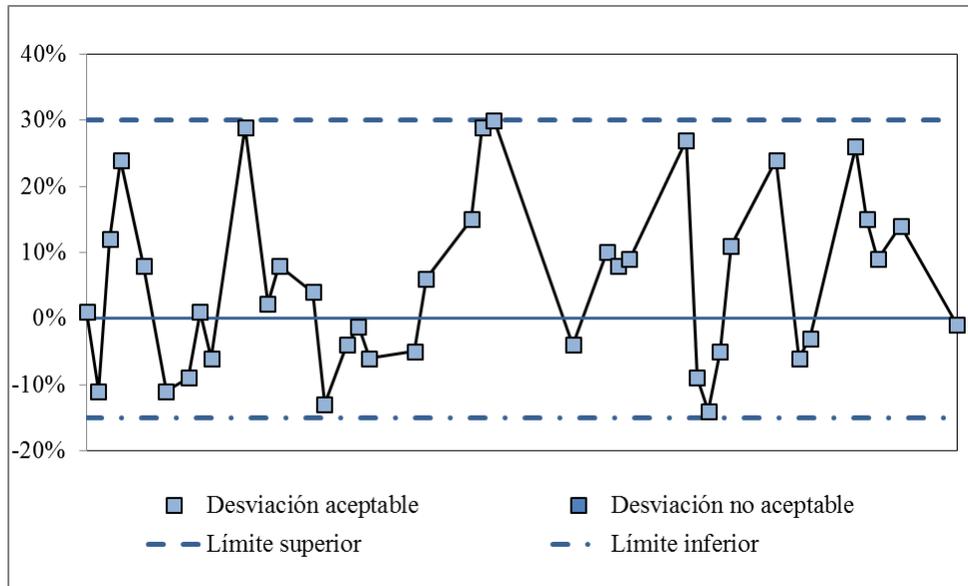
### Proceso 3

Se muestran los gráficos de control obtenidos del subproceso para observar el comportamiento de cada tarea.

#### Subproceso 3-1



Gráfica 18. “Gráfica de control – Variable 3.1.3”. Fuente: Adaptación del autor.



**Gráfica 19. “Gráfica de control – Variable 3.1.4”. Fuente: Adaptación del autor.**

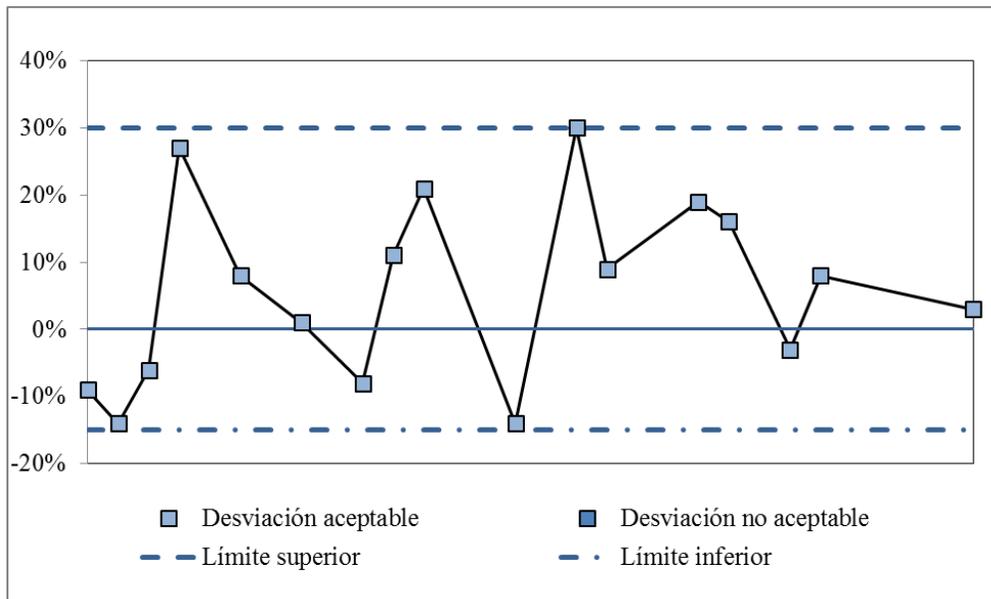
La “Gráfica 18” muestra una tendencia hacia arriba de la mitad del gráfico de control al final y se observa un punto superando el límite superior, es necesario vigilar que no sigan fuera del límite los puntos de desviación consecutivos.

Los puntos de la “Gráfica 19” están abrazando los límites de control, pero no hay alguno fuera de ellos, sólo hay que vigilar su comportamiento sucesivo.

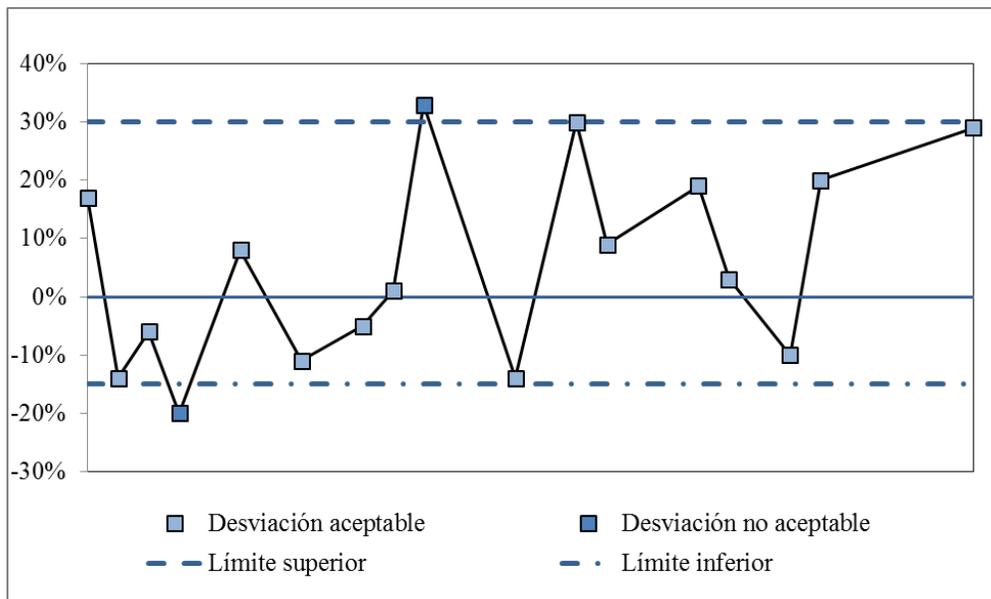
## Proceso 4

Se muestran los gráficos de control del subproceso para entender el comportamiento de las tareas respectivas.

### Subproceso 4-1



Gráfica 20. "Gráfica de control – Variable 4.1.2". Fuente: Adaptación del autor.



Gráfica 21. "Gráfica de control – Variable 4.1.3". Fuente: Adaptación del autor.

Todos los puntos de la “Gráfica 20” están dentro de los límites de control y no muestra algún comportamiento específico.

La “Gráfica 21” muestra un punto fuera de cada límite de control, pero cerca de ellos, así que se clasifican como puntos atípicos.



## Diseño de solución

Después de que se realizó el análisis de los datos (comportamiento y puntos críticos de la calidad de los procesos y productos) pudo construirse una solución de acuerdo con las necesidades de la empresa, la cual se apegó a los objetivos planteados al inicio del proyecto.

## Modelo propuesto

Debía demostrarse matemáticamente que los procesos y productos se encontraban bajo control, por lo cual fue necesario desarrollar dos modelos estadísticos para cada objetivo, uno para el control de rendimiento y otro para el control de calidad.

Tanto para el modelo de rendimiento como para el modelo de calidad los datos base de un proyecto son la fecha de inicio y fecha de entrega, así que para realizar un correcto análisis y estimación de éxito del mismo es fundamental conocer estos dos datos.

## Modelo de rendimiento

Los objetivos específicos del modelo son aquellos que brindan la información necesaria para dar inicio al análisis de rendimiento del proyecto, éstos se derivan de los objetivos del proyecto ya que proporcionan el camino a seguir para cumplir con lo requerido.

Es importante conocer la capacidad que se tiene para cada una de las tareas a realizar en los proyectos y saber las demás ocupaciones existentes.

Se tienen asignaciones para diferentes unidades, de acuerdo con ellas y las fechas dadas se obtiene una primera probabilidad que se utilizará en el modelo para posteriormente llegar al resultado final, esto es de la siguiente manera:

$$P_{1,j} = \begin{cases} F_i > F_a & ; & 1 \\ A_j = \emptyset & ; & 1 \\ \text{en otro caso} & ; & 1 - \% A_j \end{cases}$$

donde:

$F_i$  = Fecha de inicio del proyecto.

$F_a$  = Fecha actual.

$A_j$  = Asignación para la unidad  $j$ .

**Fórmula 14. “Probabilidad 1 – Modelo de rendimiento”. Fuente: Adaptación del autor.**

A su vez,  $\% A_j$  se calcula como sigue:

$$\% A_j = \frac{A_j}{A}$$

donde:

$A_j$  = Asignación para la unidad  $j$ .

$A$  = Asignaciones totales (suma de asignaciones de cada unidad).

**Fórmula 15. “Cálculo del porcentaje de  $A_j$ ”. Fuente: Adaptación del autor.**

De acuerdo con la capacidad de cada unidad asignada, se obtiene una segunda probabilidad utilizada en el modelo, como sigue:

$$P_2 = \begin{cases} 0.9 & ; & C \in (-.1, .1) \\ 0.7 & ; & C \in (-.3, -.1] \cup [.1, .3) \\ 0.4 & ; & C \in (-\infty, -.3] \cup [.3, \infty) \end{cases}$$

donde:

$C =$  Factor de capacidad.

**Fórmula 16. “Probabilidad 2 – Modelo de rendimiento”. Fuente: Adaptación del autor.**

A su vez,  $C$  se calcula de la siguiente manera:

$$C = \frac{A_j - A'_j}{A_j}$$

donde:

$A_j =$  Asignación para la unidad  $j$ .

$A'_j =$  Asignación estimada para la unidad  $j$ .

**Fórmula 17. “Cálculo del factor de capacidad  $C$ ”. Fuente: Adaptación del autor.**

Finalmente, con la lista de las probabilidades  $P_1$  y  $P_2$  de las unidades asignadas, deben ajustarse los datos a una distribución estadística para conocer su comportamiento. Las funciones de densidad de probabilidad consideradas fueron:

NOMBRE	PARÁMETROS	ESPERANZA
Normal	$\mu, \sigma$	$\mu$
Logística	$\mu, s$	$\mu$
Geométrica	$p$	$1/p$
Poisson	$\lambda$	$\lambda$
Exponencial	$\lambda$	$1/\lambda$

**Tabla 16. “Distribuciones estadísticas”. Fuente: Adaptación del autor.**

Para saber qué distribución siguen los datos se aplica la prueba de Bondad de Ajuste no paramétrica de Anderson-Darling, Kolmogorov Smirnov y Chi-cuadrado según sea el caso, seleccionando la distribución que tenga mayor p-value, pues esto significa que los datos se apegan más a dicha distribución.

## Modelo de calidad

Los objetivos específicos del modelo son aquellos que brindan la información necesaria para dar inicio al análisis de calidad del proyecto, éstos se derivan de los objetivos de calidad del proyecto ya que proporcionan el camino a seguir para cumplir con lo requerido.

Se tienen asignaciones para diferentes unidades y cada unidad tiene perfil propio, de acuerdo a dicho perfil, se obtiene un porcentaje promedio de errores en el proyecto y dependiendo del porcentaje, se asigna un porcentaje de éxito  $Q_{I,j}$ , entendiendo por  $j$  la unidad seleccionada, tal como se muestra a continuación:

Perfil Unidad	Porcentaje de éxito	Porcentaje promedio de errores
Experto	92 %	30 %
Competente	80 %	45 %
Junior	60 %	70 %

Tabla 17. “Porcentaje de éxito y errores por perfil”. Fuente: Adaptación del autor.

Se obtiene de este aspecto una primera probabilidad para ser utilizada en el modelo. Es importante mencionar que, de acuerdo con la relevancia de cada unidad dentro del proyecto, se le asigna cierto peso y luego se multiplica su ponderación por el porcentaje de éxito obtenido anteriormente, es decir:

$$Q_1 = r_j(Q_{1,j})$$

donde:

$r_j =$  *Peso asignado a la unidad j.*

$Q_{1,j} =$  *Porcentaje de éxito de la unidad j.*

**Fórmula 18. “Probabilidad 1 – Modelo de calidad”. Fuente: Adaptación del autor.**

Por otro lado, hay tareas fijas que deben realizarse en función de un listado previamente formado, se obtiene una segunda probabilidad para este aspecto del proyecto como se muestra a continuación:

$$Q_2 = \frac{m}{n}$$

donde:

$m =$  *Número de tareas realizadas.*

$n =$  *Número total de tareas que deben realizarse.*

**Fórmula 19. “Probabilidad 2 – Modelo de calidad”. Fuente: Adaptación del autor.**

## Pronóstico de rendimiento

De la distribución con mejor ajuste al comportamiento de los datos, se obtienen las mediciones principales para obtener un pronóstico de éxito del proyecto, como son: esperanza, varianza, desviación estándar, moda, mediana, simetría, etc., depende de las necesidades del negocio la medida utilizada.

Conociendo la probabilidad de éxito del proyecto  $E_1$  se tienen varias posibilidades de mensaje  $R_1$ :

$$R_1 = \begin{cases} E_1 > 0.9 & ; \text{ El proyecto será exitoso, no deben tomarse acciones} \\ 0.6 < E_1 \leq 0.9 & ; \text{ Habrá desviación, deben implementarse acciones} \\ E_1 \leq 0.6 & ; \text{ El proyecto no tendrá éxito, deben realizarse cambios} \end{cases}$$

Fórmula 20. “Resultado – Pronóstico de rendimiento”. Fuente: Adaptación del autor.

## Pronóstico de calidad

El pronóstico de calidad del proyecto se obtiene asignando un peso a cada aspecto evaluado en el proyecto ya que tienen diferente impacto, en este caso se asigna 0.7 para los perfiles de las unidades asignadas y 0.3 para las tareas fijas a realizar, así para obtener la probabilidad de éxito del proyecto  $E_2$  se tiene la siguiente fórmula:

$$E_2 = e_i(Q_i)$$

donde:

$e_i =$  Peso asignado al aspecto  $i$ .

$Q_i =$  Porcentaje de éxito del aspecto  $i$ .

Fórmula 21. “Resultado – Pronóstico de calidad”. Fuente: Adaptación del autor.

## Softwares candidatos

El paso final fue la programación del modelo, ya que son miles de datos a procesar y hubiera sido imposible generar el resultado a mano. Existen softwares especializados de acuerdo con las necesidades del usuario, por lo cual se realizó el análisis de algunos que se adaptan a lo requerido.

A continuación, se mencionan las principales características de los diferentes softwares propuestos:

- Forecast Pro.- Interfaz sencilla, fácil de usar, conjunto de gráficas limitado para representar adecuadamente los resultados, software comercial.
- R.- Interfaz vinculada con R Studio que es sencillo de utilizar, se necesita conocimiento de las paqueterías que contiene, el conjunto de gráficas permite representar de manera adecuada los resultados, software libre.
- Oracle Crystal Ball.- Interfaz sencilla, conjunto de gráficas adecuado para representar los resultados, software para realizar simulación y métricas, software comercial.

NOTA: Se trabajó con archivos de contenido histórico como, datos en Excel, archivos csv (*comma-separated values*, en español significa “valores separados por coma”) y bases de datos, los cuales podían leerse con todos los softwares analizados.

Después de analizar y discutir las ventajas y desventajas de cada software con la Dirección General, se eligió “R”. La decisión se basó en que el software es gratuito, las gráficas son buenas y se tienen las implementaciones requeridas para el modelo; además las personas asignadas para el desarrollo del modelo contaban con el conocimiento necesario para trabajar las diversas paqueterías del lenguaje y no se requería de una capacitación.

Cabe mencionar que la empresa ya contaba con una herramienta llamada “Statistica”, cuya funcionalidad está completamente enfocada al análisis estadístico. Esta herramienta no es gratis, pero la empresa ya la tenía, por lo cual, no tuvo que realizar alguna inversión para su uso, adicionalmente puede habilitarse la opción de permitir leer código en “R” desde

“Statistica”, dándole una mejor presentación (de mayor calidad) a las gráficas generadas con el código desarrollado en “R”.

## Implementación de la solución

Con el desarrollo teórico del modelo y el software seleccionado, se continuó con la programación en “R” para que el proceso fuera más veloz y dinámico, ya que dependiendo de ciertos datos que se introduzcan, el modelo realiza el análisis.

### **Programación del modelo**

Lo primordial fue identificar las paqueterías requeridas de acuerdo a las funcionalidades del modelo diseñado, posteriormente se programó cada paso para llegar al resultado de las fórmulas presentadas anteriormente.

Fue necesario solicitar una conexión directa a la base de datos de los registros de tareas, tiempo invertido, actividades fijas, etc., para poder realizar el análisis del proyecto en tiempo real, es decir, al momento que se ejecuta contar con todos los registros y dar un porcentaje de éxito correcto y actual.

Se revisó el tipo de dato de las variables obtenidas de la base de datos para poder trabajar con ellas y se omitieron datos nulos. Finalmente es importante fijarse si la fecha actual está entre las fechas de inicio y fin del proyecto, si es anterior al inicio o si sobrepasa el fin del proyecto.

Después de una revisión minuciosa sobre la funcionalidad del modelo ya programado, se solicitó la elaboración de una interfaz con Java, donde se pidiera al usuario sólo los datos necesarios para realizar el análisis y éste obtuviera al final el resultado de éxito del proyecto. Se implementó con la finalidad de hacer el modelo más amigable para los usuarios y evitar confusiones o errores al desconocer el lenguaje “R” y tener que realizar modificaciones al código.

## **Pruebas del modelo**

Con la interfaz terminada, los usuarios (realizadores de proyectos) comenzaron a utilizar el modelo en proyectos pasados y algunos en desarrollo para identificar errores, como son: si faltaba tomar en cuenta algún aspecto, si el proceso era lento, etc. De acuerdo con el análisis general del proyecto, gráficos y mediciones de control que se realizaban a parte de la implementación del modelo, los usuarios tenían una idea sobre si el proyecto sería exitoso, así que fue posible revisar que el resultado concordara con este análisis y, de no ser así, revisar qué estaba sucediendo.

Igualmente se revisó que la interfaz en Java funcionara correctamente y fuera fácil de usar, realizando las correcciones pertinentes.

## **Implementación del modelo**

Al finalizar la etapa de pruebas y correcciones se estableció oficialmente el uso de los modelos de pronóstico de rendimiento y de calidad para los dos objetivos planteados.

Para asegurarse que los usuarios entendieran el funcionamiento del modelo, se realizó una capacitación enfocada al análisis estadístico para explicar cómo se llegó al procedimiento de cálculo de los porcentajes de éxito.

## Resultados

Los modelos construidos fueron satisfactorios para las necesidades de la empresa y se logró obtener la certificación “CMMI-SVC ML4” que fue el objetivo principal del proyecto.

Como se menciona en la introducción del proyecto, a continuación, se muestra un ejemplo práctico de la aplicación del modelo, utilizando datos similares a los de un proyecto.

## Ejemplo

Se muestra una parte de la base de datos que se tiene:

Id	Actividad	Periodo	Categoría	Herramienta	Nivel	Estimación	Ajuste	Recurso
19	Instalación de software	1	Instalación	2	1	2	2	A1T
19	Conexión - Base de datos	1	Instalación	2	1	3	0.5	A1T
13	Atender falla 1	1	Pruebas	3	2		2	A1T
14	Soporte 1 para el cliente	1	Pruebas	3	2		2	A1T
15	Auditar proceso 1	1	Pruebas	3	5	10	12	A1T
18	Auditar calidad	1	Pruebas	3	2		3	A1T
22	Atender falla 2	1	Pruebas	3	2	1.5	2	A1T
22	Apoyo para preventa	1	Investigación	1	3		6	D2V
24	Compatibilidad de versiones de software	1	Investigación	1	3	15	11	D2V
10	Ajuste de riesgos	1	Desarrollo	3	2	1	2	D2V
11	Análisis inicial	1	Desarrollo	3	8	25	52	D2V
15	Ejecutar respaldos	1	Desarrollo	3	34		130	D2V
22	Visitas al cliente	1	Desarrollo	3	3	10.5	8	D2V
15	Capacitación sobre el tema	1	Investigación	1	5	13	14	D2V-V3G
23	Interfaces para el usuario	1	Diseño	2	13	30	32	J4R
26	Toma de mediciones	1	Desarrollo	3	5	19	30	J4R
25	Descripción de controles del sistema	1	Diseño	3	5	20	24	J4R
4	Reporte de fallas	1	Pruebas	3	1		1	J4R
5	Reporte de mejoras	1	Pruebas	3	1	1	1	J4R
14	Probar producto	1	Pruebas	3	2	3	3	O5M
17	Realizar una presentación de las mediciones	1	Documentación	1	3	18.5	21	O5M-A1T
17	Prueba visitas al cliente	1	Documentación	1	2	5	6	O5M-A1T
14	Revisión de corrección de fallas	1	Pruebas	3	2		2.5	R6L
18	Revisión de requerimientos	1	Pruebas	3	2		2	R6L
21	Soporte 2 para el cliente	1	Pruebas	3	2	0.5	1.5	R6L
21	Auditar proceso 2	1	Pruebas	3	2	1	1	R6L
17	Encuesta de satisfacción al cliente	1	Documentación	1	2	15	16	S7B
17	Retrospectiva del periodo	1	Documentación	1	3	20	24	S7B
1	Estimación de tareas	1	Desarrollo	3	5	20	18	O5M

Tabla 18. “Muestra de base de datos”. Fuente: Adaptación del autor.

A esta base de datos se le aplicaron los modelos desarrollados, adicionalmente se introdujeron la fecha de inicio y fin del proyecto que son 10/01/2018 y 01/06/2018 respectivamente.

### **Resultado “Modelo de rendimiento”**

Al aplicarse el modelo de rendimiento a los datos mostrados anteriormente, el mensaje obtenido fue el siguiente:

```
"El porcentaje de éxito de tu proyecto es: 83% "  
"Para asegurar el éxito del proyecto:"  
"1.- Establece una fecha de fin más lejana"  
"2.- Agrega un recurso al proyecto para terminar en la fecha de fin establecida"
```

**Resultado 1. “Resultado del modelo de rendimiento”. Fuente: Adaptación del autor.**

Nótese que el porcentaje de éxito obtenido es menor a 90%, por lo tanto, se considera necesario realizar algún cambio para asegurar el éxito, entonces el modelo proporciona dos opciones de solución al usuario.

### **Resultado “Modelo de calidad”**

Cuando se aplicó el modelo de calidad a la base de datos mostrada, el mensaje fue el siguiente:

```
"La calidad de tu proyecto es de: 95% "
```

**Resultado 2. “Resultado del modelo de calidad”. Fuente: Adaptación del autor.**

En este caso se tiene un porcentaje alto (mayor a 90%), así que el modelo no percibe que deba realizarse algún cambio y sólo proporciona el resultado al usuario.

## Conclusiones

Al dedicarse la empresa, donde se realizó el proyecto, al desarrollo de software y servicios, el ahorro de tiempo, la buena estimación de tareas y distribución adecuada de los trabajadores, puede mantener el ingreso establecido conforme a cierta planeación, buscando que los clientes queden satisfechos al recibir el producto o servicio a tiempo y con las funcionalidades que hayan solicitado al principio, lo cual conduce a poder mantenerlos como clientes para proyectos futuros.

Cuando comenzó el desarrollo del proyecto, se modificaron las mediciones de los procesos de la empresa buscando su mejora, esto fue, utilizar modelos cuantitativos y probabilísticos. Durante el desarrollo de los modelos se comprobó que las mediciones fueron más eficientes para el control de los procesos, ya que, al basarse en números históricos y obtener resultados estadísticos, los pronósticos fueron mucho más acertados que con las mediciones anteriores y pudieron evitarse diversas desviaciones de gran importancia para los proyectos y procesos generales de la empresa.

La optimización de las mediciones utilizadas por la empresa, además de beneficiarle en sus pronósticos, fue de un ahorro suficiente en tiempo para el recurso que las ejecutaba, pues ahora con sólo introducir algunos datos en la interfaz de Java y presionar un botón, obtiene el estado de su proyecto en tiempo real.

Gracias al desarrollo del proyecto se identificaron riesgos que antes no se tomaban en cuenta dentro de la lista ya existente, entonces, además de la construcción de los modelos, fue posible realizar planes de mitigación o contingencia que se adecuaron a situaciones desconocidas.

De acuerdo con la selección de variables a controlar obtenidas del análisis de correlación, cuando se realizaron las pruebas a los modelos, éstas fueron las variables necesarias a tomar en cuenta, es decir, no se reportó alguna variable sobrante o faltante para obtener un estado acertado del proyecto de acuerdo a la comparación con proyectos pasados y con la aplicación de mediciones en proyectos desarrollándose.

Es importante mencionar que durante el desarrollo de cada proyecto se aplican gráficos de control con el fin de ver si las tareas que se están realizando siguen dentro de los límites especificados establecidos por la Dirección General y, en caso de que estén a punto de desviarse aplicar un plan de mitigación para evitarlo. En caso de detectar algún punto crítico, éste se documenta para analizar sus causas y construir un plan de contingencia y/o mitigación según sea el caso.

Además de construir gráficos de control, se monitorean los índices de capacidad de procesos  $C_p$  y  $C_{pk}$  cuando un proyecto está en desarrollo, esto con el fin de mantenerlo bajo control y si se detecta gran variación aplicar medidas correctivas.

La implementación de los modelos fue de gran beneficio para el seguimiento del rendimiento y de la calidad de los procesos y productos o servicios del negocio, puede identificarse rápidamente ejecutando el modelo en qué estado se encuentra un proyecto y aplicar planes de contingencia o mitigación al identificar alguna anomalía o riesgo, tal como se mostró en el ejemplo.

En cuanto al modelo de rendimiento, es de gran ayuda saber si es necesario agregar un recurso al proyecto para terminarlo en tiempo y forma, detectar si es necesario agregar actividades o demás acciones y no esperar hasta que llegue la fecha de entrega del proyecto y darse cuenta que no está.

El análisis realizado en el modelo de calidad, permite identificar los errores cometidos durante las actividades realizadas en un proyecto, es importante notar el gran impacto que tienen sobre éste, ya que tal vez se terminen las tareas a tiempo, pero si no funciona correctamente el producto final, podrían tener que repetirse varias labores, así que la entrega del proyecto se retrasaría o se entregaría sin la funcionalidad requerida por el cliente. El modelo indica si se detectó alguna anomalía referente a la calidad del proyecto como, si se toman en cuenta todos los riesgos a los que está expuesto, si se realizaron las tareas fijas necesarias, el perfil de las unidades asignadas, etc., para al final entregar un proyecto que funcione correctamente y en la fecha estipulada.

Las herramientas y bases para la implementación de los modelos fueron satisfactorias para obtener la certificación “CMMI-SVC ML4”, ya que se revisa que los procesos del negocio estén controlados basándose en aplicaciones matemáticas, en este caso fueron probabilísticas, así que los modelos y procedimiento de construcción fueron satisfactorios para que la empresa alcanzara la meta establecida.

Posteriormente se espera dar seguimiento y mantenimiento al modelo para garantizar su efectividad. Por un lado, el modelo se actualiza sólo en cuanto a ajustes y sugerencias, así que bastaría revisar si es necesario añadir o eliminar variables de medición.

## Bibliografía

- Casella, G., Berger R. (2002). Statistical inference. Estados Unidos de América: Duxbury.
- Ross, S. (2014). Introduction to probability and statistics for engineers and scientists. Amsterdam: Elsevier.
- Samuel A. Broverman, Ph. D., ASA. (2014). ACTEX C Study Manual Volume I. Estados Unidos de América:ACTEX Publications, Inc.
- Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining. (2012). Introduction to linear regression analysis. Hoboken, New Jersey: J. Wiley.
- Rincón, Luis. (2012). Introducción a la teoría del riesgo. México:UNAM, Facultad de Ciencias.
- Documento “Modelos lineales” de apuntes de “IBM Knowledge Center”. Obtenido el 27 de febrero de 2019 del sitio [https://www.ibm.com/support/knowledgecenter/es/SS3RA7\\_sub/modeler\\_mainhelp\\_client\\_ddita/components/linear\\_models/idh\\_alm.html](https://www.ibm.com/support/knowledgecenter/es/SS3RA7_sub/modeler_mainhelp_client_ddita/components/linear_models/idh_alm.html)
- Documento “Modelos Deterministas: Optimización Lineal” de apuntes de “Espejo para América Latina”, 8ª edición. Obtenido el 27 de febrero de 2019 del sitio <http://home.ubalt.edu/ntsbarsh/opre640S/spanishD.htm>
- Artículo “¿Qué es el método cualitativo?” de “Artículos de Economía y Empresa”, escrito el 3 de mayo de 2017. Obtenido el 27 de febrero de 2019 del sitio <https://cursos.com/metodo-cualitativo/>
- Documento “Modelos probabilísticos” del curso “Matemáticas para la Toma de Decisiones”. Obtenido el 27 de febrero de 2019 del sitio <https://www.marcoteorico.com/curso/87/matematicas-para-la-toma-de-decisiones/719/modelos-probabilisticos>
- Documento “Análisis de Pareto” de apuntes de “Activa Conocimiento”. Obtenido el 28 de febrero de 2019 del sitio <http://activaconocimiento.es/analisis-de-pareto/>
- Documento “CMMI para Servicios, Versión 1.3” de CMMI Institute. (2013). Obtenido el 15 de marzo de 2017 del sitio <https://cmmiinstitute.com/getattachment/adc972f-5500-42b7-81cc-6c748a13e74d/attachment.aspx>

- Documento “Un ejemplo práctico de la participación del actuario en la implantación y desarrollo de determinados aspectos cuantitativos del proceso ORSA”, parte del libro “Dossier”, N° 32 n PRIMAVERA 2013. Obtenido el 20 de marzo de 2017 del sitio [https://www.fundacionmapfre.org/documentacion/publico/i18n/catalogo\\_imagenes/grupo.cmd?path=1075695](https://www.fundacionmapfre.org/documentacion/publico/i18n/catalogo_imagenes/grupo.cmd?path=1075695)
- Documento “Gestión de riesgos en ingeniería del software” de apuntes de la Universidad de Murcia de la asignatura “Informática Aplicada a la Gestión Pública”, Rev. 1.1 2002/12/13. Obtenido el 20 de marzo de 2017 del sitio [http://dis.um.es/~barzana/Informatica/IAGP/IAGP\\_riesgos.html](http://dis.um.es/~barzana/Informatica/IAGP/IAGP_riesgos.html)
- Documento “Estadística y control de calidad”. Obtenido el 11 de enero de 2017 del sitio [http://www.calidad.com.mx/docs/art\\_88\\_1.pdf](http://www.calidad.com.mx/docs/art_88_1.pdf)
- Documento “Capacidad de Proceso”. Obtenido el 04 de abril de 2017 del sitio <https://www.ingenieriaindustrialonline.com/herramientas-para-el-ingeniero-industrial/gestión-y-control-de-calidad/capacidad-de-proceso/>

## Anexos

### **SCMPI nivel 4 de SVC / CMMI – SVC ML4**

Los modelos CMMI (Capability Maturity Model Integration, en español significa “Integración de modelos de madurez de las capacidades”) son colecciones de buenas prácticas que ayudan a las organizaciones a mejorar sus procesos. Estos modelos son desarrollados por equipos de productos con miembros procedentes de la industria, el gobierno, y el SEI (Software Engineering Institute, en español significa “Instituto de Ingeniería de Software”). Este modelo, denominado CMMI-SVC (Capability Maturity Model Integration for Services, en español significa “Integración de modelos de madurez de las capacidades para mejorar los procesos de servicio”), proporciona un conjunto completo e integrado de guías para prestar mejores servicios.

#### **Mejores prácticas**

Las mejores prácticas del modelo se centran en las actividades para proveer servicios de calidad a los clientes y usuarios finales. CMMI-SVC integra conjuntos de conocimiento que son esenciales para proveedores de servicios.

#### **Riesgo**

Posibilidad de experimentar ciertos eventos de interés y las consecuencias derivadas de ellos. El objetivo es identificar los riesgos, ponderarlos con base en sus consecuencias, decidir la aceptación o no de los mismos y tomar provecho de su existencia. Es importante mencionar que los riesgos no necesariamente son negativos, en ocasiones ofrecen oportunidades.

Mientras la aceptación del riesgo conlleva a los planes de contingencia, evitar el riesgo requiere planes de mitigación.

## Plan de contingencia

Estrategia que debe aplicarse cuando se ha decidido aceptar un riesgo previsto y este se presenta. En términos generales se tienen, por ejemplo, los planes de evacuación en caso de incendios, temblores, inundaciones, etc., que no son de ocurrencia frecuente y sale de las posibilidades de la empresa evitarlos.

## Plan de mitigación

Acciones que deben implementarse para evitar un riesgo previsto, con la finalidad de evitar consecuencias graves derivadas de él en caso de suceder. Por ejemplo, que los contratos del negocio cuenten con una cláusula referente a evitar el robo de información de la empresa.

## Prueba de Anderson-Darling

Ésta prueba y algunas otras se utilizan para saber qué tan bien se ajusta una muestra de variables aleatorias a una distribución de probabilidad específica.

La prueba se basa en una muestra aleatoria de puntos de datos individuales  $x_1, x_2, \dots, x_n$ . Si existen valores repetidos se vuelven a etiquetar los distintos valores de la muestra. Suponiendo que se tienen  $k$  valores distintos en la muestra, se tiene  $y_1 < y_2 < \dots < y_k$ .

Si los datos no están censurados, entonces  $y_0 = 0$  y  $y_{k+1} = \infty$ . La prueba estadística de Anderson-Darling es:

$$A^2 = -n + n \left[ \sum_{j=0}^{k-1} (1 - F_n(y_j))^2 \cdot \ln \left( \frac{1 - F^*(y_j)}{1 - F^*(y_{j+1})} \right) + \sum_{j=1}^k F_n(y_j)^2 \cdot \ln \left( \frac{F^*(y_{j+1})}{F^*(y_j)} \right) \right]$$

donde:

$F$  = Función de distribución acumulativa de acuerdo a la muestra aleatoria.

$F^*$  = Función de distribución acumulativa obtenida sin datos repetidos.

**Fórmula 22. “Prueba estadística de Anderson-Darling sin censura”. Fuente: Adaptación del autor.**

En caso de que los datos estén truncados por la izquierda a  $d$ , entonces  $y_0 = d$  y si los datos están censurados por la derecha a  $u < \infty$ , entonces  $y_{k+1} = u$ . La prueba estadística es:

$$A^2 = -nF^*(u) + n \left[ \sum_{j=0}^{k-1} (1 - F_n(y_j))^2 \cdot \ln \left( \frac{1 - F^*(y_j)}{1 - F^*(y_{j+1})} \right) + \sum_{j=1}^k F_n(y_j)^2 \cdot \ln \left( \frac{F^*(y_{j+1})}{F^*(y_j)} \right) \right]$$

donde:

$F =$  Función de distribución acumulativa de acuerdo a la muestra aleatoria.

$F^* =$  Función de distribución acumulativa obtenida sin datos repetidos.

**Fórmula 23. “Prueba de Anderson-Darling con datos truncados por la izquierda y censurados por la derecha”.**  
**Fuente: Adaptación del autor.**

Los valores críticos de la prueba de hipótesis son:

Nivel de significancia:	.1	.05	.01
Valor crítico:	1.933	2.492	3.857

## **Glosario**

*Calidad:* Compromiso y estrecho vínculo de la empresa con los clientes, a fin de cumplir con sus expectativas y brindar un producto o servicio de acuerdo a los parámetros declarados.

*Estadísticamente estable:* Indica que los datos caen dentro de los límites de control y fluctúan de manera aleatoria, sin un orden o patrón de comportamiento.

*Línea base:* Estado inicial de los procesos de una empresa. Se toma como base para implementar mejoras o correcciones.

*Optimización:* Buscar la mejor manera de realizar algo. En términos de negocio, es obtener la mayor ganancia o producción, con el menor costo o desperdicio.

*Pronóstico:* Conocimiento anticipado de lo que sucederá en un futuro, basado en estudios realizados a datos históricos.

*Variabilidad:* Cuando hay una alta proporción de puntos cerca de los límites de control, en ambos lados de la línea central y pocos puntos en la parte central.

*Variable aleatoria:* Resultados posibles de un experimento (aún no sucede, pero se sabe qué resultados pueden obtenerse).

*Variable observable:* Indicador del posible comportamiento de las variables aleatorias (se conoce el resultado).