

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICAS**

PROYECTO DE TITULACIÓN

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

**“MAGÍSTER EN ESTADÍSTICA CON MENCIÓN EN CALIDAD Y
PRODUCTIVIDAD”**

TEMA:

APLICACIÓN DEL MÉTODO MONTECARLO EN EL ANÁLISIS DE
PRODUCTIVIDAD DE UN LABORATORIO BROMATOLÓGICO.

AUTOR:

STALIN BOANERGES PINARGOTE ZAVALA

Guayaquil - Ecuador

2020

RESUMEN

En una fábrica de la provincia del Guayas dedicada a la elaboración de alimentos, la primera etapa de la cadena productiva es la calificación de las materias primas. Bajo este concepto, el laboratorio bromatológico de esta empresa necesita de resultados confiables y a tiempo. Para cumplir con esto, el proceso al interior del laboratorio bromatológico comprende desde la recepción de materias primas hasta el envío de los resultados de las materias primas a las áreas involucradas, sin embargo, dentro del mismo, existen etapas del proceso que son más elaborados que otros y, por este motivo, existen retrasos en la entrega de resultados para que la materia prima calificada siga con el proceso de la cadena productiva. Dado a lo anteriormente descrito, se tiene la necesidad de mejorar la productividad de cada una de las etapas del proceso analítico para reducir los tiempos de entrega de los resultados. Los tiempos de cada proceso de análisis son tomados para posteriormente crear una o varias simulaciones mediante el método Monte Carlo y encontrar el mejor escenario para el incremento de la productividad dentro del laboratorio bromatológico.

Como resultado se obtuvo que la productividad se podría incrementar hasta en un 56% siempre y cuando se solucionen las causas que producen retrasos en la entrega de análisis.

ABSTRACT

In a factory in Guayas province dedicated to food manufactory, the first step of production chain is qualification of raw materials. Under this concept, the bromatological laboratory of this factory needs of reliable and in time results. In order to fulfill this, the process inside the bromatological laboratory takes from reception of trucks until the results were delivered to the involved areas, however, inside of it, there are steps of the process that takes more time than others and, for this reason, there are delays to get the results in order to follow the process of the production chain. Due this, there is a necessity to improve the productivity of each step of the analytical process to reduce delivery time of the results. The time of each analytical process are taken to get one or many simulations through Monte Carlo method and reach the best scenery to increase productivity inside the bromatological laboratory.

As result, we obtained that productivity could be increased to 56% as long as causes that produce delays on results delivery can be solved.

DEDICATORIA

A mis padres Stalin y Zenaida y a mis hermanos Edison y Christian los cuales me han tenido paciencia, confianza y sobre todo demostrado mucho amor durante todos estos años.

A mi familia los cuales me han apoyado en todo momento y están felices por esta nueva meta alcanzada, finalmente lo conseguí.

Stalin

AGRADECIMIENTO

A la vida que, pese a estar sufriendo de la situación actual, me ha permitido llegar a alcanzar esta nueva meta.

A mis padres Stalin y Zenaida los cuales me dieron la vida y me han ayudado mucho durante todos estos años, su paciencia e incondicionalidad fueron mi impulso para terminar una etapa más de mi vida profesional, son los mejores y los amo mucho. También a mis hermanos Edison y Christian ya que con su ayuda me hicieron posible llegar a esta meta.

A mi muy estimada familia que también me han dado muchos mensajes de apoyo a lo largo de este proyecto de titulación.

Finalmente, un agradecimiento muy especial es para la PhD. Sandra Acosta dado a que por su confianza pude proseguir con mis estudios de posgrado. De igual manera a la Mgtr. Nadia Cárdenas por su ayuda a lo largo de este proyecto de titulación.

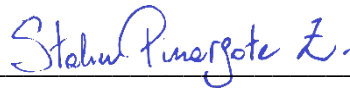
A todos ustedes mis más sinceros agradecimientos.

Stalin

DECLARACIÓN EXPRESA

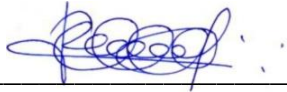
La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Titulación, me corresponde exclusivamente y ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría. El patrimonio intelectual del mismo, corresponde exclusivamente a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

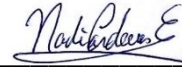


Stalin Boanerges Pinargote Zavala

TRIBUNAL DE GRADUACIÓN



M.Sc. John Ramirez Figueroa
PRESIDENTE



Mgtr. Nadia Cárdenas Escobar
DIRECTOR



Sandra García Bustos, Ph.D.
VOCAL 1



Omar Ruiz Barzola, Ph.D.
VOCAL 2

ABREVIATURAS O SIGLAS

FDP: Función de distribución de probabilidad.

TM: Toneladas métricas.

FPA: Función de probabilidad acumulada.

SMC: Simulación de Monte Carlo.

MC: Monte Carlo.

CNA: Cámara Nacional de Acuicultura.

USD: Dólares Americanos.

ppb: Partes por billón.

h: horas.

min: minutos.

m: metros.

mm: milímetros.

nm: nanómetros.

°C: grados centígrados.

μl: microlitros.

μg: microgramos.

g: gramos.

kg: kilogramos.

MMC: Método de Monte Carlo.

MT: Variable Mañana Temprana.

MM: Variable Media Mañana.

T: Variable Tarde.

ACF: Función de autocorrelación.

VIF ó FIV: Análisis de inflación de varianza.

ANOVA: Análisis de varianza.

H₀: Hipótesis nula.

H₁: Hipótesis alterna.

TABLA DE CONTENIDO

1.	INTRODUCCIÓN.....	1
1.1.	Antecedentes.....	3
1.2.	Descripción del problema	3
1.3.	Objetivo general.....	4
1.4.	Objetivos específicos.....	4
1.5.	Hipótesis.....	4
1.6.	Alcance.....	4
2.	MARCO TEÓRICO	6
2.1.	Espectrometría	6
2.2.	Quimiometría	7
2.2.1.	Técnicas Quimiométricas.....	9
2.2.2.	Aplicación de la Quimiometría.....	10
2.2.3.	Regresión.....	10
2.2.4.	Clasificación, reconocimiento de patrones y agrupamiento	11
2.2.5.	Resolución de Curvas Multivariadas.....	12
2.2.6.	Usos de la Quimiometría.....	12
2.3.	Análisis NIR	12
2.4.	Métodos Químicos de Referencia.....	13
2.4.1.	Determinación del contenido de humedad en alimentos.....	14
2.4.2.	Determinación del contenido de cenizas en alimentos	17
2.4.3.	Determinación del contenido de grasa en alimentos.....	19
2.5.	Proceso de Elaboración de Pellets	22
2.5.1.	Molienda	22
2.5.1.1.	La molienda y el proceso de peletización.....	23
2.5.1.2.	La molienda y su relación con el pre-acondicionamiento	23

2.5.1.3.	Los ingredientes y sus efectos en la molienda.....	24
2.5.2.	Mezclado.....	25
2.6.	Productividad.....	26
2.6.1.	Importancia de la medición:	27
2.6.2.	Pertinencia.	28
2.6.3.	Precisión.	28
2.6.4.	Oportunidad.	29
2.6.5.	Confiabilidad.	29
2.6.6.	Economía.	30
2.7.	Regresión Lineal Múltiple	30
2.7.1.	Condiciones para la Regresión Lineal Múltiple	31
2.7.1.1.	Multicolinealidad o No Colinealidad:	31
2.7.1.2.	Relación lineal entre los predictores numéricos y la variable respuesta34	
2.7.1.3.	Distribución normal de los residuos	35
2.7.1.4.	Variabilidad constante de los residuos (homocedasticidad):.....	35
2.7.1.5.	No autocorrelación (Independencia):	35
2.8.	Simulación	36
2.8.1.	Etapas de la Simulación.....	36
2.8.1.1.	Definición, descripción del problema. Plan.....	37
2.8.1.2.	Formulación o Construcción del modelo.	37
2.8.1.3.	Verificación del modelo.	38
2.8.1.4.	Validación del modelo.	38
2.8.1.5.	Análisis de resultados	39
2.8.2.	Tipos de Simulación.....	39
2.8.3.1.	Simulación continua:	39
2.8.3.2.	Simulación por eventos discretos:.....	40
2.8.3.3.	Simulación por autómatas celulares:	40

2.8.3.4. Simulación de Monte Carlo:	40
2.9. Investigaciones Recientes	41
2.9.1. Aplicación del Método Monte Carlo en Exploración y explotación de petróleo y gas natural	41
2.9.2. Diseño e implementación de un sistema de inventarios, aplicando simulación Montecarlo, en una empresa de Servicios Petroleros.....	42
3. METODOLOGÍA	43
3.1. Simulación de Monte Carlo.....	43
3.2. Pasos para Elaborar la Simulación de Monte Carlo	46
3.3. Descripción del Área de Trabajo.....	49
3.4. Pasos Previos a la Simulación de Monte Carlo	52
3.4.1. Identificación de las Variables de Entrada y de Salida.....	52
3.4.1.1. Variables de Entrada.....	52
3.4.1.2. Variable de Salida	59
3.4.2. Cálculo de las Correlaciones.....	61
3.4.3. Construcción del modelo.....	61
3.5. Verificación del modelo.....	67
3.6. Cálculo del número de simulaciones	71
3.7. Validación del Modelo de situación inicial.....	75
3.7.1. Estadística Descriptiva de la Simulación.....	75
3.7.2. Prueba Chi-cuadrado de Bondad de Ajuste.....	77
4. RESULTADOS	80
4.1. Análisis de Causa de los Retrasos	80
4.2. Aplicación del Método Monte Carlo en la resolución del Problema (Escenario Probable)	82
4.3. Aplicación del Método Monte Carlo en la resolución del Problema (Escenario Optimista).....	88

4.4. Aplicación del Método Monte Carlo en la resolución del Problema (Escenario Pesimista)	94
4.5. Diferencia entre Escenarios de Simulación	100
En la siguiente Tabla se observarán los valores de los diferentes escenarios como resumen:	100
5. CONCLUSIONES Y RECOMENDACIONES.....	101
5.1. Conclusiones	101
5.2. Recomendaciones	102
6. Referencias	103

LISTADO DE FIGURAS

Figura 1. Espectro electromagnético y tipos de radiación (Wikipedia, Espectro electromagnético, 2019).....	6
Figura 2. Ciencias relacionadas con la Quimiometría	7
Figura 3. Etapas de la Quimiometría.....	9
Figura 4. Diagrama de funcionamiento de los equipos NIR	13
Figura 5. Diagrama de Flujo de una Simulación	37
Figura 6. Diagrama de Proceso de los análisis de laboratorio	50
Figura 7. Histograma de la variable tiempo de ingreso de los datos al sistema T_{in}	53
Figura 8. Histograma de los datos de la Variable T_{pr}	54
Figura 9. Histograma de los datos de la variable T_{an}	56
Figura 10. Histograma de los datos de la variable T_{al}	57
Figura 11. Histograma de los datos de la variable T_{en}	58
Figura 12. Histograma de los datos de la variable de salida Mnts (minutos/análisis)	60
Figura 13. Matriz de Correlación de los datos de la situación inicial en el programa R	61
Figura 14. Medición de Akaike para determinar las variables significativas en el modelo de la situación inicial del laboratorio bromatológico	62
Figura 15. Modelo estadístico propuesto de la situación inicial del laboratorio bromatológico.....	63
Figura 16. Diagrama de dispersión entre los residuos del modelo propuesto y las variables de la situación inicial	64
Figura 17. Prueba de Durbin-Watson del modelo propuesto para la situación inicial	65
Figura 18. Prueba de homocedasticidad de Breusch-Pagan del modelo propuesto de la situación inicial	65
Figura 19. Prueba de normalidad de Shapiro-Wilk en los residuos del modelo propuesto en la situación inicial	66

Figura 20. Gráfico Q-Q de normalidad de los residuos del modelo propuesto del laboratorio en la situación inicial	66
Figura 21. Análisis de inflación de Varianza (VIF) del modelo propuesto en la situación inicial	67
Figura 22. Comparativo del Sistema Real vs. Simulación Piloto (Minutos/análisis)	71
Figura 24. Diagrama de cajas del comparativo de la distribución de la variable de salida (minutos/análisis) de los datos reales vs los datos de la simulación	76
Figura 25. Diagrama de cajas del comparativo del número de análisis diarios de los datos reales vs los datos de la simulación	76
Figura 26. Gráfica de Frecuencias Observadas y Esperadas	78
Figura 27. Diagrama de Causa-efecto en retraso de la entrega de resultados.....	80
Figura 28. Diagrama de Pareto de las frecuencias de ocurrencia de inconvenientes Septiembre-2019.....	81
Figura 29. Histograma de la Variable T_{in} – Simulación escenario probable	84
Figura 30. Histograma de la Variable T_{an} – Simulación escenario probable	85
Figura 31. Histograma de la Variable T_{al} – Simulación escenario probable	86
Figura 32. Histograma de la Variable T_{en} – Simulación escenario probable	87
Figura 33. Histograma de la Variable T_{in} – Simulación escenario optimista.....	90
Figura 34. Histograma de la Variable T_{an} – Simulación escenario optimista.....	91
Figura 35. Histograma de la Variable T_{al} – Simulación escenario optimista.....	92
Figura 36. Histograma de la Variable T_{en} – Simulación escenario optimista.....	93
Figura 37. Histograma de la Variable T_{in} – Simulación escenario pesimista.....	96
Figura 38. Histograma de la Variable T_{an} – Simulación escenario pesimista	97
Figura 39. Histograma de la Variable T_{al} – Simulación escenario pesimista.....	98
Figura 40. Histograma de la Variable T_{en} – Simulación escenario pesimista	99

LISTADO DE TABLAS

Tabla 1. Valor-p de la variable ingreso de los datos al sistema	53
Tabla 2. Valor-p de la variable preparación de la muestra	55
Tabla 3. Valor-p de la variable análisis de la muestra	56
Tabla 4. Valor-p de la variable almacenamiento de los datos en el sistema	57
Tabla 5. Valor-p de la variable tiempo de envío de resultados	59
Tabla 6. Valor-p de la variable de salida Mnts (minutos/análisis).....	60
Tabla 7. Números aleatorios generados junto con los valores de la variable aleatoria T_{in}	68
Tabla 8. Números aleatorios generados junto con los valores de la variable aleatoria T_{pr}	68
Tabla 9. Números aleatorios generados junto con los valores de la variable aleatoria T_{an}	69
Tabla 10. Números aleatorios generados junto con los valores de la variable aleatoria T_{al}	69
Tabla 11. Números aleatorios generados junto con los valores de la variable aleatoria T_{en}	70
Tabla 12. Minutos por análisis y número de análisis por día obtenidos de la simulación propuesta de la situación inicial del laboratorio	71
Tabla 13. Estadística Descriptiva del Sistema Real vs Simulación Piloto	Error!
Bookmark not defined.	
Tabla 14. Valor-p de la prueba de hipótesis de diferencia de medias.....	Error!
Bookmark not defined.	
Tabla 15. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo.....	75
Tabla 16. Resumen estadístico de los minutos por análisis y del número de análisis diarios obtenidos a través de los datos reales y de la simulación.....	75
Tabla 17. Tabla de Frecuencias Observadas en la Simulación comparados con las Proporciones del Sistema Real	78
Tabla 18. Prueba Chi-cuadrado de bondad de ajuste	78

Tabla 19. Tabla de frecuencia de los problemas encontrados durante Septiembre-2019	80
Tabla 20. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo en el Escenario Probable	83
Tabla 21. Valor-p de la variable tiempo de ingreso – Simulación escenario probable	84
Tabla 22. Valor-p de la variable tiempo de análisis – Simulación escenario probable	85
Tabla 23. Valor-p de la variable tiempo de almacenamiento de los datos en el sistema – Simulación escenario probable	86
Tabla 24. Valor-p de la variable tiempo de envío – Simulación escenario probable	87
Tabla 25. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo en el Escenario Optimista	89
Tabla 26. Valor-p de la variable tiempo de ingreso – Simulación escenario optimista	90
Tabla 27. Valor-p de la variable tiempo de análisis – Simulación escenario optimista	91
Tabla 28. Valor-p de la variable tiempo de almacenamiento de los datos en el sistema – Simulación escenario optimista	92
Tabla 29. Valor-p de la variable tiempo de envío – Simulación escenario optimista	93
Tabla 30. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo en el Escenario Pesimista.....	95
Tabla 31. Valor-p de la variable tiempo de ingreso – Simulación escenario pesimista	96
Tabla 32. Valor-p de la variable tiempo de análisis – Simulación escenario pesimista	97
Tabla 33. Valor-p de la variable tiempo de almacenamiento de los datos en el sistema – Simulación escenario pesimista.....	98
Tabla 34. Valor-p de la variable tiempo de envío – Simulación escenario pesimista	99

Tabla 35. Resumen de las Simulaciones en diferentes Escenarios.....100

CAPÍTULO 1

1. INTRODUCCIÓN

De acuerdo a la Cámara Nacional de Acuicultura (CNA), en Ecuador entre enero y noviembre del 2018 se exportaron 506 mil TM de camarón, un incremento de 15%, frente al mismo período del 2017 que equivalen a USD 3.234,8 millones. “Con este crecimiento de volumen necesitamos de todos los mercados posibles”. Eso dice Camposano cuando analiza la estrategia comercial para el 2019. Entre diciembre del 2017 y junio del 2018 se vendieron USD 1,5 millones a Brasil, según Pro Ecuador. En esa época el mercado se abrió luego de una batalla judicial de seis meses que permitió superar una barrera sanitaria de 18 años.

El sector tiene otros retos para el año 2019. Reabrir las ventas a México y a Corea del Sur, ambos con barreras sanitarias. Según (Pro Ecuador, 2013), a través de su oficina en Seúl se trabaja para “minimizar los impactos de este certificado”. En el 2017 ese país fue el quinto mayor destino y venía creciendo a un ritmo del 15 % al 20 % anual (CNA, 2019).

En lo que va del año (datos hasta septiembre 2019), la cifra de camarón exportado es de 472 miles de TM con un valor de USD 2.705,8 millones. Debido al incremento en años anteriores, ha aumentado la demanda y producción de alimentos para camarón que brinden soluciones al sector camaronero (CNA, 2019).

Consecuentemente, se ha incrementado la cantidad de materia prima que se reciben en la empresa en la cual se realiza el presente estudio, ubicada en Durán provincia del Guayas se dedica a la producción y

comercialización de balanceado de varias especies acuícolas y agrícolas.

El laboratorio bromatológico donde se aplicó esta simulación se dedica al análisis de materias primas para su calificación (aprobación – rechazo) antes de su ingreso al proceso, por lo cual se podría decir que, dentro del laboratorio se encuentra la primera parte del proceso productivo de la planta de balanceados.

El caso particular del presente trabajo, la simulación es utilizada para determinar los tiempos de proceso para el mejoramiento de la productividad, puesto que necesitamos ubicar, reubicar y mejorar los recursos que se tienen actualmente en el laboratorio para entregar resultados confiables y a tiempo. Según Rodríguez-Aragón (2011), en la Industria se utiliza la simulación para predecir las consecuencias que tendrá la toma de una decisión determinada como control de inventarios, planes de mantenimiento, localización de recursos, predicción de ventas o demanda, etc. La simulación permite resolver problemas complejos, aunque lo que obtendremos será una aproximación de la solución.

En el capítulo 2 se presenta temas teóricos relacionados a simulación y aspectos importantes como la revisión de literatura relevante para el presente trabajo con mención de trabajos anteriores relacionados.

En el capítulo 3 se muestra la metodología de la simulación Monte Carlo, la recolección de datos y el diseño de simulación correspondiente.

En el capítulo 4 se encuentra los resultados de las simulaciones realizadas bajo los escenarios probable, optimista y pesimista.

Y, por último, en el capítulo 5 se muestran las conclusiones y recomendaciones obtenidas de las simulaciones del presente trabajo.

El presente trabajo se enfoca en la elaboración de un modelo estadístico de simulación con el fin de reducir los tiempos de entrega de resultados y la mejora de la productividad en un laboratorio bromatológico a partir de una toma de decisiones.

1.1. Antecedentes

El laboratorio de análisis de alimentos de una empresa en la provincia del Guayas se encarga de proveer información de la composición de la materia prima que se recibe antes de ingresarla al proceso. Este laboratorio, entrega los resultados de análisis con el fin de proporcionar datos que son idóneos para el fin que están destinados que son la aprobación o rechazo según las especificaciones de calidad establecidas. Para esto, se debe mantener la garantía de la calidad en los resultados, los cuales deben ser confiables y oportunos. En este último punto se basará el presente trabajo, dado a que en la actualidad se tienen problemas de entregar los resultados de los análisis a tiempo.

1.2. Descripción del problema

Para realizar cada análisis en el laboratorio bromatológico es necesario seguir protocolos y/o procesos los cuales son necesarios para cumplir de mejor manera las actividades desempeñadas por la misma. Cada uno de estos procesos toman un tiempo más que otros, por lo cual se deben identificar cuáles son los procesos de vital importancia para implementar mejoras y crear un modelo que ayude a incrementar la productividad en el laboratorio.

En los últimos meses al incrementarse la demanda de alimento de camarón, también se ha incrementado la cantidad de materia prima recibida. Debido a esto, los resultados se han visto retrasados dado a

que se necesita una mejora en los procesos de recepción y análisis de las muestras.

1.3. Objetivo general

Reducir el tiempo de entrega de los resultados con el fin de mejorar la productividad en el laboratorio bromatológico.

1.4. Objetivos específicos

- Validar que los datos estén correctamente aplicados de acuerdo con la realidad de la empresa.
- Diseñar un modelo de simulación matemático aplicando la teoría de Monte Carlo.
- Encontrar el mejor escenario que reduzca el tiempo de la entrega de resultados.

1.5. Hipótesis

El diseño de un modelo de simulación, ayuda en la reducción de los tiempos de entrega de resultados que contribuya al incremento de la productividad, y de esta manera, aumenta el número de análisis realizados en el laboratorio bromatológico.

1.6. Alcance

Este estudio únicamente es aplicable en el laboratorio bromatológico ubicado en Durán, provincia del Guayas, Ecuador.

Los datos para este estudio fueron obtenidos del mismo sitio en el cual se ejecutan los procesos de análisis, se observaron cada una de las etapas de los análisis para la medición de los tiempos. Los datos fueron tomados durante el mes de septiembre del 2019.

A partir del 15 de octubre se procedió con la medición de las causas del problema de retrasos mediante análisis de causa-efecto y diagrama de Pareto que indican las frecuencias de las causas de los retrasos.

Finalmente, mediante datos históricos, se procede con las simulaciones bajo diferentes escenarios para observar el grado de mejora que se puede obtener en el proceso.

CAPÍTULO 2

2. MARCO TEÓRICO

2.1. Espectrometría

Es un método para realizar mediciones y/o cuantificación a partir de un espectro. Un espectro electromagnético es una distribución energética del conjunto de las ondas electromagnéticas. Un espectro electromagnético o simplemente espectro es una radiación electromagnética que emite (espectro de emisión) o absorbe (espectro de absorción) una sustancia. Esta relación sirve para identificar la sustancia tal como si fuera una huella dactilar. Los espectros se pueden observar mediante espectroscopios que, además de permitir ver el espectro, permiten realizar medidas sobre el mismo, como son la longitud de onda, la frecuencia y la intensidad de la radiación (Esbensen & Geladi, 2019).

El espectro electromagnético se extiende desde la radiación de menor longitud de onda, como los rayos gamma y los rayos X, pasando por la radiación ultravioleta, la luz visible y la radiación infrarroja, hasta las ondas electromagnéticas de mayor longitud de onda, como son las ondas de radio (Wikipedia, Espectro electromagnético, 2019).

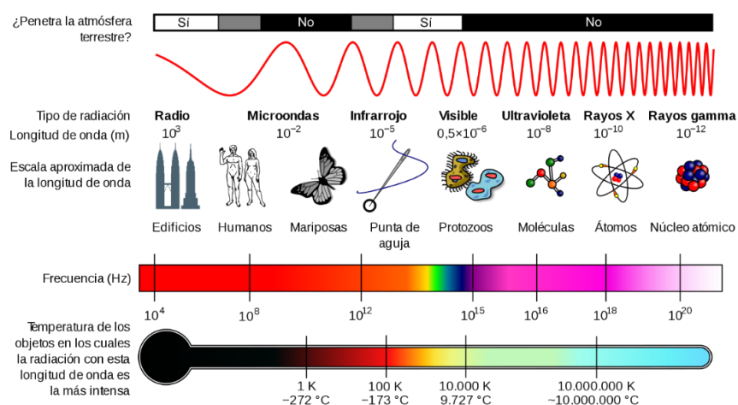


Figura 1. Espectro electromagnético y tipos de radiación (Wikipedia, Espectro electromagnético, 2019)

La radiación infrarroja, es un tipo de radiación electromagnética, de mayor longitud de onda que la luz visible, pero menor que la de las microondas. Por lo tanto, tiene menor frecuencia que la luz visible y mayor que las microondas. Su rango de longitudes de onda va desde unos 0,7 hasta los 1000 micrómetros (CRISP, 2019).

La clasificación de la luz infrarroja o IR, de acuerdo a su longitud de onda en nanómetros (nm), es:

- Infrarrojo cercano (de 800 nm a 2500 nm).
- Infrarrojo medio (de 2.5 μm a 50 μm).
- Infrarrojo lejano (de 50 μm a 1000 μm).

2.2. Quimiometría

Es una ciencia de la extracción de información de sistemas de datos químicos. Esta es de carácter interdisciplinaria, usa métodos utilizados en disciplinas de análisis de datos como ciencias computacionales, estadística multivariada y matemáticas aplicada para resolver problemas de química, biología, medicina, bioquímica, ingeniería química y correlacionarlos con parámetros de calidad o propiedades físicas (Wikipedia, Chemometrics, 2019).

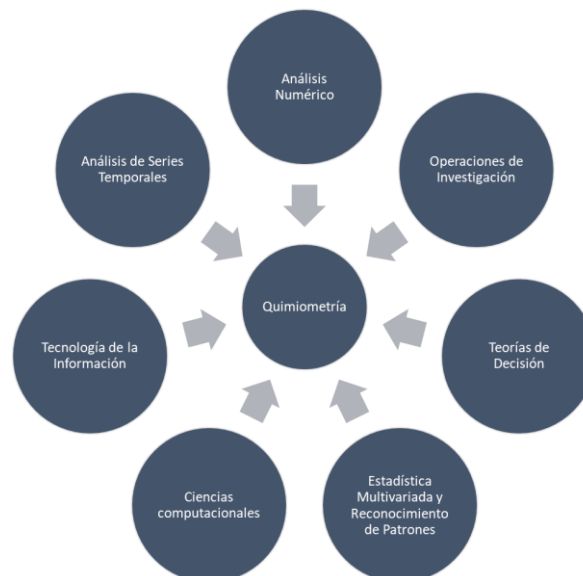


Figura 2. Ciencias relacionadas con la Quimiometría

La quimiometría se aplica para resolver problemas descriptivos y predictivos en experimentos de ciencias naturales, especialmente en química. Las propiedades químicas son modeladas; esto implica realizar cálculos en las mediciones de datos químicos. Por ejemplo, para calcular el pH desde la medición de la actividad de iones de hidrógeno mediante la transformación de una interpolación de Fourier de un espectro.

El uso de la quimiometría se refiere al uso de métodos de cálculos de álgebra lineal, para realizar medidas tanto cualitativas como cuantitativas a partir de datos de carácter químicos, principalmente espectros.

La quimiometría provee diferentes maneras de resolver un problema de calibración para el análisis de datos espectrales. Unos son sencillos de comprender, mientras que otros requieren una fuerte retrospectiva de álgebra lineal. Sin embargo, todos tienen un punto en común dado a que estos resuelven un problema individual pero no pueden solucionar todos los problemas posibles.

Algunos métodos tienen la ventaja de ser simples de entender, pero no son tan robustos para todas las muestras posibles. Otros son demasiado complejos de entender e implementar, pero las soluciones son muy estables y pueden manejar una variedad grande de “desconocidos”.

Esencialmente el entendimiento de la quimiometría no es necesariamente entender las matemáticas de todos los diferentes métodos; es conocer qué modelo utilizar en un problema analítico dado y aplicarlo de manera apropiada.

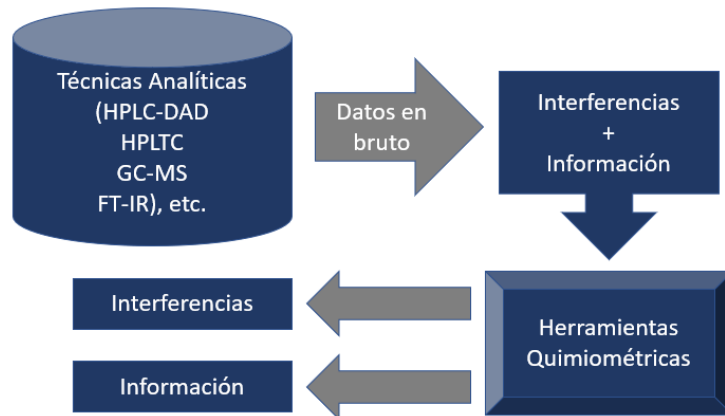


Figura 3. Etapas de la Quimiometría

2.2.1. Técnicas Quimiométricas

- Calibración multivariada: Algunos problemas químicos y aplicaciones quimiométricas aplican una calibración. El objetivo es desarrollar modelos que puedan ser usados para poder predecir propiedades de interés basadas en propiedades medibles de sistemas químicos, tales como temperatura, flujo, presión, infrarrojo, Raman, espectro de resonancia magnética nuclear, espectro de masa, entre otros. Este proceso necesita de una calibración o conjunto de datos de entrenamiento, estos incluyen valores de referencia para las propiedades de interés para la predicción, además de los atributos medibles que corresponden a estas propiedades. Por ejemplo, crear datos a partir de un número de muestras, inclusive concentraciones para un analito de interés para cada muestra como valor de referencia y su correspondiente espectro infrarrojo de esa muestra.

Existen técnicas de calibración multivariada como regresiones de mínimos cuadrados parciales (PLS), o regresiones de componentes principales (PCR) y otros métodos que usados para construir un modelo matemático que relaciona la respuesta multivariada o espectro a la

concentración de un analito de interés, por lo tanto, un modelo puede ser utilizado eficientemente para una predicción de las concentraciones de muestras nuevas.

El uso de técnicas de calibraciones multivariadas tiene una gran ventaja que son mediciones analíticas rápidas, económicas y en algunos casos son no destructivas. Además, es importante mencionar que la calibración multivariada también permite un análisis cuantitativo preciso cuando existe la presencia de fuertes interferencias por otros analitos.

2.2.2. Aplicación de la Quimiometría

- Explora posibles valores atípicos o también llamados outliers y determinar si existen patrones o tendencias en los datos.
- El análisis de componentes principales (PCA) es parte importante de la quimiometría y provee una compacta representación de toda la variación en una tabla de datos.
- Algoritmos exploratorios como PCA son utilizados para reducir un grande y complejo conjunto de datos a una serie de datos optimizados y de un tamaño interpretable.

2.2.3. Regresión

- Son utilizadas para predecir las propiedades relacionadas que son medibles de manera más fácil.
- Una regresión analítica de la quimiometría tiene por objetivo desarrollar un modelo que correlacione la información en el conjunto de medidas conocidas a partir de datos obtenidos a una propiedad deseada. Los algoritmos quimiométricos que se utilizan para elaborar regresiones incluyen mínimos

cuadrados parciales (PLS) y regresión de componentes principales (PCR).

- La regresión quimiométrica se utiliza extensamente para hacer decisiones relacionadas a la calidad de un producto en el monitoreo durante los procesos industriales donde se necesitan sistemas rápidos y costosos para ensayar.

2.2.4. Clasificación, reconocimiento de patrones y agrupamiento

Las técnicas de clasificación multivariadas supervisadas se encuentran muy relacionadas a las técnicas de calibración multivariada y son utilizadas para desarrollar modelos capaces de clasificar futuras muestras. Las técnicas empleadas en quimiometría son similares a las utilizadas en otros campos como los análisis regresiones logísticas, árboles de clasificación/regresión, redes neurales y discriminantes multivariados. La utilización de técnicas de reducción de rangos en conjunto con estos métodos de clasificación convencional es una rutina constante en quimiometría, por ejemplo, análisis discriminantes mediante PCR o PLS.

- Para asignar categorías predefinidas a las muestras analizadas para predecir una muestra desconocida y calificarla como perteneciente a uno de los diferentes grupos ya establecidos.
- Un modelo de clasificación se utiliza para predecir la clase de una muestra basando otras muestras anteriores que se asemejen a la muestra analizada.
- La quimiometría ayuda en la estandarización de los datos. Los modelos de clasificación también incluyen la habilidad de revelar posibles muestras inusuales en los datos generados.

2.2.5. Resolución de Curvas Multivariadas

En términos de quimiometría, la resolución de curvas multivariadas busca la deconstrucción de conjuntos de datos con información de referencia limitada o ausente. Algunos trabajos de estas técnicas fueron realizados por Lawton y Sylvestre a principios de los años 70. Estas aproximaciones también fueron llamadas análisis de mezcla, fuentes ciegas/separación de señales y desmezclado espectral.

2.2.6. Usos de la Quimiometría

- Calibraciones espectroscópicas.
- Modelos de proceso para su respectiva optimización.
- Modelos de proceso para monitoreo y detección de fallas.
- Modelos de identificación para el control de procesos.
- Control estadístico multivariado de procesos.
- Estandarización de instrumentos analíticos del proceso.
- Diseño y desarrollo de instrumentos analíticos.

2.3. Análisis NIR

La tecnología NIR (por sus siglas en inglés Near Infrared Reflectance, o Reflectancia Cerca del Infrarrojo) involucra un haz de luz interactuando con un material, donde una radiación electromagnética ocurre en forma de ondas. La longitud de onda es la distancia entre dos picos o puntos altos, se representa con el símbolo λ . Usualmente la longitud de onda en el espectro NIR se mide en nanómetros (nm) donde $1\text{nm} = 10^{-9}\text{ m}$ o $1000\text{ nm} = 0.001\text{ mm}$. La parte del espectro visible al ojo humano se extiende de 400nm a 800 nm, mientras que el espectro infrarrojo se extiende de 2,500 nm a 25,000 nm aproximadamente. El infrarrojo cercano es considerado la parte del espectro situada entre la región visible y la región infrarroja. El NIR cubre un rango de longitudes de onda que se encuentran entre los 750 nm a 2,600 nm.

En la práctica, la muestra analizada es bombardeada con rayos del haz de luz del NIR de diferentes longitudes de onda como se ilustra en la Figura 4. Por cada longitud de onda, algunos de los rayos serán entonces absorbidos por uniones químicas específicas presentes en la muestra. Al mismo tiempo, otros rayos serán diseminados y reflejados por otras uniones químicas (rebote). Este proceso se conoce como Reflectancia NIR. Por otro, algunos de los rayos pasarán a través de la muestra, denominado Transmisión NIR (NIT).

Los rayos esparcidos, reflejados y/o transmitidos de cada longitud de onda son concentrados dentro de una célula de medición. Un número de reflejos a diferentes longitudes de onda son medidos, para posteriormente ser convertidos en resultados analíticos por un microprocesador.

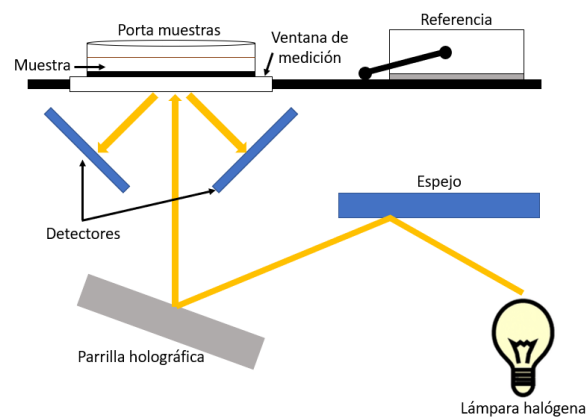


Figura 4. Diagrama de funcionamiento de los equipos NIR

2.4. Métodos Químicos de Referencia

Cada producto en los instrumentos NIR, posee una curva de predicción diferente, estas curvas son creadas a partir de los resultados obtenidos por los métodos de referencia explicados a continuación:

2.4.1. Determinación del contenido de humedad en alimentos

Los alimentos contienen agua en mayor o menor proporción, cualquiera que sea el método de industrialización a que hayan sido sometidos. El contenido de agua presente en los alimentos naturales varía entre un 60 y un 95%. Podría decirse que en los tejidos animales y vegetales existe en dos formas generales: “agua libre” y “agua ligada”. El agua libre o absorbida se encuentra en forma predominante, se pierde con gran facilidad dependiendo del alimento o del ambiente en el que se encuentra. El agua ligada se encuentra combinada o absorbida. Se encuentra en los alimentos como agua de cristalización (en los hidratos) o ligada a las proteínas y a las moléculas de sacáridos y absorbida sobre la superficie de las partículas coloidales (Hart, 1991).

La mayoría de las industrias de alimentos determinan la humedad por varios motivos, los principales son los siguientes:

- El comprador de materias primas no desea adquirir producto con agua en exceso.
- El agua, de estar presente por encima de ciertos niveles, facilita el desarrollo de microorganismos.
- Para la mantequilla, margarina, leche desecada y queso está señalado el máximo legal.
- Los materiales en polvo se aglomeran en presencia de agua, por ejemplo, sal, azúcar, café y leche en polvo.
- La humedad de ciertos granos como el trigo debe ajustarse adecuadamente para facilitar la molienda.
- La cantidad de agua presente en el alimento puede afectar a su textura.
- La determinación del contenido en agua es una forma sencilla para el control de la concentración en las distintas etapas de la fabricación de alimentos.

- La determinación de humedad es una de las técnicas más importantes y de mayor uso en el proceso, control y conservación de los alimentos, dado a que la mayoría de los productos alimenticios contienen mayoritariamente agua como ejemplo podemos citar la leche que posee un 88%, el yogurt, entre un 80 y 90%, las carnes frescas (60-75%) incluyendo los llamados productos secos como ciertas leguminosas o granos que alcanzan un contenido de humedad de hasta un 12%.
- El contenido de humedad en un alimento podría considerarse como un índice de estabilidad del producto, dado a que es importante controlar por ser un factor decisivo en muchos procesos industriales tales como la molienda de cereales, el mezclado de productos sólidos finos, en la elaboración de pan, etc. Además, es de mucha importancia conocer el contenido de agua de los productos o materias primas para formular el producto y evaluar las pérdidas durante el procesado.

a) Métodos de secado

Son los más comunes para valorar la concentración de humedad en los alimentos; se calcula el porcentaje en agua por la pérdida en peso debida a su eliminación por calentamiento bajo condiciones normalizadas. Aunque estos métodos dan buenos resultados que pueden interpretarse sobre bases de comparación, es preciso tener presente que en ocasiones es difícil eliminar por secado toda la humedad presente, dado a que, a cierta temperatura el alimento se vuelve susceptible de descomponerse, con lo que se volatilizan otras sustancias aparte de agua (Pearson, 1993).

Para la obtención de la concentración de humedad en productos alimenticios, existen varios métodos:

- Métodos directos:
 - Métodos de secado
 - En estufa de aire
 - Por radiación infrarroja
 - Método químico: Karl Fisher
- Método indirecto: Refractometría

1) Método de secado en termobalanza

Este método se basa en evaporar de manera continua la humedad de la muestra y el registrar la pérdida de peso, hasta que la muestra llegue a un peso constante.

Cuando la muestra no se expone constantemente al ambiente, el error de los pesos en este método se minimiza (Nollet, 1996).

b) Procedimiento

- Pesar de 2 a 3 g de muestra en un pesafiltro con tapa (previamente pesado después de tenerlo a peso constante 2 h a 130°C aprox.). Secar la muestra en la estufa 2 h a 100-110°C.
- Retirar de la estufa, tapar, dejar enfriar en el desecador y pesar tan pronto como se equilibre con la temperatura ambiente. Repetir hasta obtener un peso constante.
- Calcular el porcentaje de humedad, como pérdida por secado a 100-110°C.

$$\% \text{ Humedad} = \frac{P_i - P_f}{P_i} \times 100$$

Donde:

P_i: Peso inicial

P_f: Peso final

2.4.2. Determinación del contenido de cenizas en alimentos

Las cenizas de un alimento corresponden al residuo inorgánico que queda después de calcinar la materia orgánica. Las cenizas no son las mismas sustancias inorgánicas presentes en el alimento original, debido a las pérdidas por volatilización o a las interacciones químicas entre los constituyentes.

El valor principal de la determinación de cenizas supone un método sencillo para determinar la calidad de ciertos alimentos, por ejemplo, en especias y en gelatinas son un inconveniente un contenido en cenizas alto. Las cenizas de los alimentos deben estar comprendidas entre ciertos rangos, lo cual facilitará en parte su identificación (Pearson, 1993).

En los vegetales predominan los derivados de potasio y en las cenizas animales los del sodio. El carbonato de potasio se volatiliza aproximadamente a 700°C y se pierde casi por completo a 900°C. El carbonato de sodio permanece inalterado a 700°C, pero sufre pérdidas considerables a 900°C. Los fosfatos y carbonatos reaccionan además entre sí (Hart, 1991).

Notas:

- Los productos que poseen un alto contenido de agua se secan primero sobre un plato eléctrico caliente o al baño María.
- Se debe tomar como consideración principal que el producto no desprenda humos.

- Generalmente, la temperatura adecuada de la mufla son 500°C. Sin embargo, los cloruros, pueden volatilizarse a esta temperatura.
- Las cenizas se utilizan muchas veces para la determinación de otros constituyentes individuales, por ejemplo, cloruros, fosfatos, calcio y hierro. (Pearson, 1993).

a) Método de cenizas totales

En este método la materia orgánica se oxida en ausencia de flama a una temperatura entre los 550 – 600°C; el material inorgánico que no se volatiliza a esta temperatura se conoce como ceniza. (Nollet, 1996).

b) Procedimiento

- En un crisol, poner de 3 a 5 g de muestra; colocar el crisol con la muestra en una parrilla y quemar el material lentamente hasta que no desprenda humos, evitando que se proyecte fuera del crisol.
- Llevar el crisol a una mufla y efectuar la calcinación completa.
- Dejar enfriar en la mufla, transferirlo al desecador para su enfriamiento y encontrar la masa del crisol con cenizas.

$$\% \text{ Cenizas} = \frac{(P_f - P_c)}{(P_i - P_c)} \times 100$$

Donde:

P_i: Peso inicial

P_f: Peso final

P_c: Peso crisol

2.4.3. Determinación del contenido de grasa en alimentos

a) Método de Soxhlet

Es una extracción semicontinua con disolvente donde una cantidad de disolvente interacciona con la muestra y se calienta a ebullición, una vez que el líquido condensado llega a cierto nivel es sifonado de regreso al matraz de ebullición dentro del Soxhlet, la grasa se mide por pérdida de peso de la muestra o por cantidad de muestra removida. (Nielsen, 1998).

b) Procedimiento

- Una vez que se haya llevado a cabo el reconocimiento del material, se pesan unos 5g de muestra con una precisión de ± 1 mg (anotamos la cantidad exacta pesada en la balanza), en un cartucho de extracción.
- Tras cerrarse, se coloca en la pieza media del dispositivo de extracción Soxhlet o compartimento de muestra.
- El matraz redondo debe encontrarse en el desecador antes del análisis se provee del trozo de porcelana y se pesa exactamente sobre su soporte de corcho (llevaremos a cabo al menos tres pesadas). Se pesará con el mismo soporte de corcho a lo largo de toda la práctica para no introducir un error adicional debido a la variación de masa entre los distintos soportes de corcho.
- Se monta la parte inferior del dispositivo (con el pie de bureta, la manta calefactora, el matraz y el soxhlet, pero sin el reflujó). Es conveniente que la manta calefactora repose en un soporte estable pero removible para que al finalizar el enfriamiento sea más rápido.
- Se llena por la parte de arriba del soxhlet con una cantidad suficiente de disolvente (generalmente éter)

que en este caso serán unos 200 ml (es necesaria una cantidad tal que llene el asa de la parte intermedia para que durante el proceso de extracción sifonee y recircule, más las pérdidas eventuales) y se acopla al dispositivo.

- Se procede a completar el montaje del dispositivo de extracción en la campana. La parte superior del reflujo se tapona con desecante (sulfato sódico anhidrido) envuelto en algodón para evitar la entrada y condensación de vapor de agua.
- Tras el montaje se pone en marcha la manta calefactora y se regula el caudal de agua del reflujo. El éter, una vez que alcanza su temperatura de ebullición, se evapora y llega al refrigerante condensándose y cayendo en el compartimento del cartucho de muestra.
- Se tiene como precaución que un reflujo no se deja nunca sólo en el laboratorio, debido a la peligrosidad de los disolventes. También pueden ocurrir, por ejemplo, subidas de presión del agua de la red y que se suelten los tubos. Por lo tanto, tendrá que haber siempre una persona a cargo del reflujo.
- Durante la extracción, que en caso de que se complete totalmente dura entre 4 a 6 horas, se observará como se vacíe regularmente el espacio de extracción (compartimento de muestra), es decir, la pieza media del dispositivo, a través del conducto ascendente (asa) con lo que el disolvente va recirculando completándose lo que llamamos ciclos de extracción. En este caso, daremos por finalizada la extracción una vez que se han completado 4.
- Al finalizar el cuarto ciclo, se quita el calentamiento, por ejemplo, retirando la manta calefactora. Cuando el éter

deja de hervir, se quita el Soxhlet con cuidado y se extrae el cartucho.

- Se vuelve a colocar el dispositivo para calentar el matraz redondo y, cuando el Soxhlet se encuentre lleno, pero antes de que sifonee, se procede de forma análoga a anteriormente para recolectar el éter de la parte intermedia en un recipiente debidamente etiquetado. Se considera de pureza suficiente para servir para extracciones posteriores.
- Repetir este proceso una segunda vez hasta que la cantidad de éter en el matraz redondo sea muy poca.
- Dejar entonces el matraz redondo destapado unos 10 min en la campana en la manta calefactora puesta a potencia mínima y dejarlo enfriar sobre su soporte.
- Realizar una primera pesada del matraz con su soporte y trozo de porcelana y anotar su valor.
- Se considerará que corresponde al tiempo 0, volver a colocar el matraz redondo en la manta calefactora otros 10 min y dejarlo enfriar. Repetir la pesada y anotar el valor junto al tiempo total pasado en el calefactor desde el “tiempo 0”. Esta operación se repite hasta que la masa pesada deje de disminuir o, en su defecto, se hayan realizado cuatro pesadas.

$$\% \text{ Grasa} = \frac{(P_f - P_m)}{(P_i - P_m)} \times 100$$

Donde:

P_i: Peso inicial

P_f: Peso final

P_m: Peso matraz

2.5. Proceso de Elaboración de Pellets

Con ingredientes o materias primas de calidad es necesario que los procesos productivos se hagan con "Calidad". No se puede producir alimentos de calidad si los procesos utilizados no son controlados. Por eso el control de procesos debe ser parte de los controles de calidad. A continuación, se revisan los procesos productivos más importantes que se pueden emplear en la manufactura de alimentos balanceados en forma de pellets para peces y crustáceos.

2.5.1. Molienda

Es una reducción por medios mecánicos del tamaño de las partículas de un ingrediente o mezcla de ingredientes que conforman una formula completa.

El tamaño de partícula final depende del tipo de alimento que se esté elaborando. Es decir, que no es lo mismo fabricar un alimento para larvas de camarones que para camarones en la fase final de engorde. El tamaño de las partículas de la mezcla para los pellets ($< 100 \mu\text{m}$) debe ser mucho menor que para el camarón de engorde ($< 250 \mu\text{m}$). La molienda es el proceso más limitante en la producción de alimentos balanceados y representa aproximadamente el 50-60% de los costos de manufactura.

Al considerar la molienda también se debe considerar su impacto en el mezclado ya que afecta directamente la homogeneidad de la mezcla y del producto final (pellet o producto extrusado).

Cuando se consideran los procesos de aglomeración como la peletización o de extrusión el tamaño final de la partícula y la uniformidad de la mezcla tienen un efecto directo en estos

procesos. En el caso de la peletización en el que se utilizan dados con orificios de 2.2 mm de diámetro, requiere que el tamaño de la partícula no sea mayor a 260 μm (malla 60 US Tyler), pero es preferible tener un tamaño de partícula entre 150-180 μm . Es necesario tomar en consideración que una prensa Peletizadora debe cumplir su función de presionar y compactar la mezcla en un paquete llamado pellet y no utilizar los rodillos y la pista del dado como un molino para reducir el tamaño de la mezcla. Esto último no solo reduce la eficiencia del equipo, sino que también incrementa las mermas por finos, además de incrementar los costos energéticos.

2.5.1.1. La molienda y el proceso de peletización

Además de lo citado anteriormente, otra relación importante entre la molienda y la peletización es el efecto de compactación. Mientras el tamaño de partícula es mucho menor existen más puntos de contactos. A medida que aumentemos los puntos de contacto obtendremos mejor enlace entre los componentes de la mezcla – proteínas (gluten) y las moléculas de almidón, formando una estructura más sólida, compacta y resistente a la degradación en el agua.

2.5.1.2. La molienda y su relación con el pre-acondicionamiento

La molienda juega un papel importante en el pre-acondicionamiento. De igual manera que en la peletización a menor tamaño de partícula más superficie de área expuesta a la acción del vapor. Esto permite que el vapor se condense en más partículas, y al hacer esto transfiere su calor y el agua

es absorbida más rápidamente. Partículas más grandes requieren de mayor tiempo en el pre-acondicionador para lograr la gelatinización de los almidones que es muy importante en la hidroestabilidad de los pellets.

2.5.1.3. Los ingredientes y sus efectos en la molienda

Materias primas como harinas de pescado, calamar, krill, cabeza de camarón, almeja, frijol de soya, y las harinas de carne, carne y hueso se encuentran en mayor proporción de la ración de una fórmula para peces o camarones en las que los niveles de proteína oscilan entre el 25 y el 65%.

Muchos de estos ingredientes tienen un alto contenido de grasa. Esto tiende a tapar las mallas de orificios pequeños en los molinos de martillo. Debido a esto, es importante mantener el nivel de grasa por debajo del 13% para evitar estas obstrucciones, de preferencia el nivel de grasa no debería ser superior al 6%. Una forma de reducir el nivel total de grasa es mezclar los ingredientes con mayor contenido de grasa conjuntamente con un cereal (trigo, etc.) al moler. Por este motivo, hoy en día el método preferido es la post molienda.

Un sistema de post molienda, por lo general todos los ingredientes mayores son pesados y premezclados por mezcladoras estáticas o tolvas de cascada. Estas premezclas son pasadas posteriormente bien sea por una serie de molinos de martillo con cribas de selección, o por pulverizadores. Después de la

molienda, los ingredientes mayores molidos se depositan en la tolva de mezclado y luego pasan a la mezcladora donde los líquidos y las premezclas de vitaminas y otros ingredientes que no requieren ser molidos, son adicionados.

Con la post molienda se obtiene un tamaño de partícula más uniforme, se evitan que se tapen las mallas causadas por ingredientes con altos contenidos de grasa, se utilizan menos tolvas o silos de almacenamiento de ingredientes molidos y, por último, se mejora la homogeneidad de la mezcla y la calidad final del pellet.

2.5.2. Mezclado

El mezclado es una operación donde todos los ingredientes se incorporan con el objetivo principal de que la mezcla sea homogénea. Debido a esto, es una de las operaciones más importantes en la fabricación de alimentos balanceados para peces y crustáceos. Si se considera el gasto/inversión que se hace en adquirir los ingredientes o materias primas de calidad, almacenarlos, y pesarlos, se debe poner especial atención al proceso que se va a utilizar en poner todos estos ingredientes en una sola mezcla homogénea. El objetivo de este proceso es crear una mezcla homogénea que cubra todos los requerimientos nutricionales de la especie en la fase de desarrollo específica para la cual se creó la fórmula. En otras palabras, cualquier muestra que se tome de una mezcla debe ser idéntica en contenido nutricional a cualquier otra mezcla.

2.6. Productividad

La Productividad se define como la relación entre la cantidad de servicios realizados y la cantidad de recursos utilizados. En términos de servicios la productividad sirve para evaluar el rendimiento de los equipos de trabajo y los empleados.

En términos de empleados, la productividad es sinónimo de rendimiento. En un enfoque sistemático se dice que alguien es productivo con una cantidad de recursos o insumos, en un periodo de tiempo determinado si obtiene el máximo de resultados. La productividad del recurso humano o los trabajadores no está dada como parte de sus habilidades o conocimientos técnicos, sin embargo, se debe considerarse como factores que influyen.

Además de la relación de cantidad de servicios realizados por recursos utilizados, en la productividad se deben considerar otros aspectos muy importantes como:

La calidad que es la velocidad a la cual los servicios se proveen especialmente por unidad de mano de obra o trabajo.

Productividad = Salida/Entradas

Entradas: Materia prima, mano de obra, energía, capital, maquinaria y otros recursos,

Salidas: Productos.

Muchas empresas se enfocan a una gran variedad de servicios. Estas últimas suelen ser heterogéneas tanto en valor como en diversidad de servicios y su complejidad tecnológica puede presentar grandes diferencias. En estas empresas la productividad global se mide basándose en un número definido de “centros de utilidades” que representan en forma adecuada la actividad real de la empresa. La fórmula se convierte entonces en:

$$Productividad = \frac{SP \text{ equipo } A + SP \text{ equipo } B + \dots + SP \text{ equipo } n}{Insumos \text{ empleados}}$$

2.6.1. Importancia de la medición:

- La medición permite planificar con mayor certeza y confiabilidad.
- Permite discernir con mayor precisión las oportunidades de mejora de un proceso en específico.
- Permite analizar y explicar cómo ocurren los sucesos.

Muchas veces se interpreta que la medición solo, es útil para conocer las tendencias “promedios”, olvidando que estas son útiles dependiendo de cómo se presenten o se procesen y que cuando dirigimos procesos dentro de la empresa no bastan únicamente las tendencias “promedios” sino que se debe ir más allá, conociendo datos como la precisión la variabilidad en toda su gama y la interconexión de factores y causas en cada situación.

Sin medición no se puede evaluar, planificar, diseñar, prevenir, corregir y mantener, innovar sistemáticamente las actividades del proceso de mejoramiento.

La medición no solo se entiende como un proceso de toma de datos, sino que se debe insertar adecuadamente en el sistema para la toma de decisiones. Se pueden tener muchos datos sobre la causa de un efecto, pero si no se procede con su clasificación, se estudia su frecuencia, se aísla los principales y se establece sus relaciones, con la finalidad ya sea de poner bajo control el proceso o de mejorar su desempeño de poco servirán dichos datos y la medición.

Con el fin de garantizar la confiabilidad de los datos de un sistema de medición se necesita contar con un clima organizacional sano, donde los intereses comunes de la

organización prevalezcan sobre los intereses individuales de los departamentos y mucho más sobre aquellos intereses individuales incompatibles con el logro de los objetivos del sistema analizado.

Las mediciones deben ser transparentes y entendibles para quienes harán uso de ellas y, adicionalmente deberán reunir y tener una serie de atributos indispensables.

2.6.2. Pertinencia.

La pertinencia se refiere a que las mediciones que se realizan deben ser tomadas en cuenta y tener importancia en las decisiones que se toman en base de la misma.

El grado de pertinencia de una medición debe revisarse periódicamente, ya que algo que sea muy importante en un momento determinado, puede dejar de serlo al transcurrir un cierto periodo de tiempo. El grado de pertinencia de una medición, es relativa al conjunto de mediciones a realizar, dado a los recursos y capacidades de procesamiento y dirección que se tenga. A medida que se coloca un sistema bajo control, se puede manejar por excepción un conjunto de variables y ello ayuda a concentrarse en otras que requieren mayor atención.

2.6.3. Precisión.

Es el grado en que la medida obtenida refleje fielmente la magnitud que queremos analizar o corroborar, se interesa por conocer un proceso, tomar decisiones para tener resultados esperados. Debido a esto que se requiere conocer a fondo la precisión del dato que estamos obteniendo.

Para lograr la precisión de una medición, se deben realizar los siguientes pasos:

- Realizar una buena definición operativa, como la definición de la característica, de las unidades de escala de medición, número y selección de las muestras, cálculo de las estimaciones y errores permisibles o tolerancias de la medición.
- Elegir un instrumento de medición con el nivel de apreciación adecuado.
- Asegurar que el dato dado por el instrumento de medición, sea bien tomado por el operador, gerente, oficinista o inspector a cargo de hacerlo. Esto supone adiestrar al personal, además de que supone tener un buen clima organizacional donde todos estén interesados en la fidelidad cada lectura.

2.6.4. Oportunidad.

La medición es información para el logro de ese conocimiento profundo de los procesos, que permite toma de decisiones más adecuadas, bien sea para corregir estableciendo la estabilidad deseada del sistema, bien sea para prevenir y tomar decisiones antes de que se produzca la anomalía indeseada o más aún, para diseñar incorporando elementos que impiden que las características deseadas se salgan fuera de los límites de tolerancia.

2.6.5. Confiabilidad.

Se refiere fundamentalmente a que la medición en la empresa no es un acto que se haga una sola vez, por el contrario, es un acto repetitivo y de naturaleza realmente periódica. Si lo que se desea es estar seguros que lo que se mide sea la base

adecuada para las decisiones que serán tomadas, se debe revisar periódicamente todo sistema de medición.

2.6.6. Economía.

La justificación económica es sencilla y compleja a la vez. Sencilla, porque se refiere a la proporcionalidad que debe existir entre los costos incurridos entre la medición de una característica o hechos determinados y los beneficios y relevancia de la decisión que soportamos con los datos obtenidos. Aunque cuantificar esta proporcionalidad no es fácil en muchos casos, por lo complejo de cuantificar importancia y relevancia de decisiones.

2.7. Regresión Lineal Múltiple

Permite generar un modelo lineal en el que el valor de la variable dependiente o respuesta (Y) es determinada a partir de un conjunto de variables independientes o también llamadas predictores ($X_1, X_2, X_3, \dots, X_n$). Los modelos de regresión múltiple pueden ser utilizados para predecir el valor de la variable dependiente o para evaluar la influencia que tienen los predictores sobre la misma (esto se debe analizar con cautela para no malinterpretar la relación causa-efecto).

Los modelos lineales múltiples son expresados mediante siguiente ecuación:

$$Y_i = (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni}) + e_i$$

Donde:

- β_0 : es el valor de la variable dependiente Y cuando todos los predictores son cero también llamada ordenada en el origen.
- β_i : es el efecto promedio que tiene el incremento en una unidad de la variable predictora X_i sobre la variable dependiente y, se mantiene constante en el resto de variables. También conocido como coeficientes parciales de regresión.

- e_i : es la diferencia entre el valor observado y el estimado por el modelo, conocido también como residuo o error.

Se debe tener en cuenta que la magnitud de cada coeficiente parcial de regresión depende mucho de las unidades en las que se mida cada variable predictora a la que corresponde, por lo que su magnitud no se encuentra asociada con la importancia de cada predictor. Para poder determinar el impacto que tienen en el modelo cada una de las variables, se emplean coeficientes parciales estandarizados, estos se obtienen al estandarizar las variables predictoras previo ajuste del modelo, en el cual se debe sustraer la media y dividir entre la desviación estándar (Amat, 2020).

2.7.1. Condiciones para la Regresión Lineal Múltiple

Los modelos de regresión lineal múltiple requieren de las mismas condiciones que los modelos lineales simples además de otras adicionales, tales como multicolinealidad o no colinealidad, relación lineal entre los predictores numéricos y la variable respuesta (linealidad), distribución normal de sus residuos, variabilidad constante de los residuos (homocedasticidad) y no autocorrelación (Independencia).

2.7.1.1. No Colinealidad:

Los predictores deben ser independientes en los modelos lineales múltiples, en otros términos, no debe haber colinealidad entre ellos. Cuando un predictor está linealmente relacionado con uno o varios de los otros predictores del modelo o cuando es la combinación lineal de otros predictores entonces ocurre la colinealidad.

La colinealidad no se puede identificar de manera precisa, el efecto individual que tiene cada una de las variables colineales sobre la variable respuesta, lo cual causa un incremento de la varianza de los coeficientes de regresión estimados hasta un punto que resulta prácticamente imposible establecer su significancia estadística. Pequeños cambios en los datos pueden provocar grandes cambios en las estimaciones de los coeficientes.

Teóricamente la colinealidad propiamente dicha existe solo si el coeficiente de correlación simple o múltiple entre algunas de las variables independientes es 1, aunque, esto sucede muy raramente en la realidad. Sin embargo, frecuentemente se puede encontrar la llamada casi-colinealidad o multicolinealidad no perfecta.

No se aplica un método estadístico en concreto para determinar la existencia de colinealidad o multicolinealidad entre los predictores de un modelo de regresión, sin embargo, se han desarrollado reglas prácticas que ayudan a determinar en qué medida se afecta a la contraste y estimación de un modelo. Los pasos que se pueden recomendar a seguir son:

- Si el coeficiente de determinación R^2 es elevado, pero ninguno de los predictores resulta significativo, puede haber indicios de colinealidad.
- Estudiar la relación lineal entre cada par de predictores mediante el cálculo de una matriz de correlación. Se debe tener en cuenta de manera importante que, aunque no se obtenga ningún

coeficiente de correlación elevado, no se asegura que no exista multicolinealidad. Se puede dar algún caso de tener una relación lineal casi perfecta entre tres o más variables y que las correlaciones simples entre pares de estas mismas variables no sean mayores que 0.5.

- Elaborar un modelo de regresión lineal simple entre cada uno de los predictores. Si en alguno de los modelos el R^2 conocido también como coeficiente de determinación es elevado, podría estar señalando una posible colinealidad.
- Tolerancia (TOL) y Factor de Inflación de la Varianza (VIF). se trata de dos parámetros que cuantifican lo mismo (uno es el inverso del otro). El VIF de cada predictor se calcula según la siguiente fórmula:

$$VIF_{\hat{\beta}_j} = \frac{1}{1 - R^2}$$

$$Tolerancia_{\hat{\beta}_j} = \frac{1}{VIF_{\hat{\beta}_j}}$$

Donde R^2 es obtenida de la regresión del predictor X_j sobre los otros predictores. Esta opción es la más recomendada, a continuación, se observan los límites de referencia que se suelen emplear:

- $VIF = 1$: Ausencia total de colinealidad
- $1 < VIF < 5$: La regresión puede ser afectada por cierta colinealidad.
- $5 < VIF < 10$: Existe una alta probabilidad de que exista colinealidad entre las variables.
- El termino tolerancia es $1/VIF$ por lo que los límites recomendables están entre 1 y 0.1.

Si se encontrara colinealidad entre predictores, existen dos posibles soluciones.

- 1) Excluir uno de los predictores que dan problemas intentando conservar el que, bajo un buen criterio, se encuentra influyendo realmente en la variable respuesta. Esta solución no suele tener mucho impacto en el modelo en su capacidad predictiva debido a que, si llegara a existir colinealidad, la información que aporta uno de los predictores es redundante en presencia del otro.
- 2) Combinar las variables colineales en un único predictor, con esto, existe el riesgo de perder su interpretación.

Si se intenta establecer relaciones causa-efecto, la colinealidad puede llevar a conclusiones muy erróneas y, puede que se estime de manera errónea que una variable es la causa cuando en realidad es otra la que está influenciando sobre ese predictor (Amat, 2020).

2.7.1.2. Relación lineal entre los predictores numéricos y la variable respuesta

Todo predictor numérico debe estar linealmente relacionado con la variable respuesta (Y) mientras los demás predictores se mantienen constantes, de no ser así, no se puede ser introducido en el modelo. La manera más recomendable de comprobar esto es representar los residuos del modelo frente a cada uno de los predictores. Si esta relación es lineal, los residuos se distribuyen entorno a cero de forma aleatoria. Estos análisis únicamente son

aproximados, debido a que no hay forma de saber si realmente la relación es lineal cuando el resto de predictores se mantienen constantes (Amat, 2020).

2.7.1.3. Distribución normal de los residuos

Los residuos deben distribuirse de forma normal con media cero. Para comprobar esto, se recurre a los cuantiles normales, test de hipótesis o histogramas (Amat, 2020).

2.7.1.4. Variabilidad constante de los residuos (homocedasticidad):

La varianza de los residuos debe de ser constante en todo el rango de observaciones. Para comprobar esto, se representan los residuos. La varianza, de ser constante, se distribuye de forma aleatoria y mantiene una misma dispersión sin ningún patrón en específico. Una distribución cónica puede ser un identificador de falta de homocedasticidad. Otra forma de revisar este supuesto, es recurrir a contrastes de homocedasticidad como el test de Breusch-Pagan (Amat, 2020).

2.7.1.5. No autocorrelación (Independencia):

Los valores de cada una de las observaciones son independientes de los otros, cuando se trabaja con mediciones temporales es especialmente importante de comprobar. Es recomendado representar los residuos ordenados acorde al tiempo de registro de las observaciones, en caso de que exista cierto patrón existe indicios de autocorrelación. También se puede

utilizar el test de hipótesis de Durbin-Watson (Amat, 2020).

2.8. Simulación

Es un proceso de diseño y desarrollo de un modelo mediante computadoras de un proceso o sistema real y realizar experimentos mediante este modelo con el fin de entender el comportamiento del sistema estudiado o evaluar algunas estrategias con las que se puede manejar el sistema mediante un cambio en las variables o incluso predicciones acerca del comportamiento del sistema (Sotou, 2005).

Usualmente, los modelos que describen sistemas se desarrollan a través de modelos matemáticos que intentan encontrar soluciones de análisis de varios problemas que permiten predecir el comportamiento de un sistema de un conjunto de datos y condiciones iniciales.

La simulación se usa de manera frecuente como un accesorio o un sustituto de sistemas de modelado en los cuales las soluciones simples no son posibles.

Existen varios tipos de simulación por computadora, sin embargo, el punto en común que comparten es que intentan generar varios escenarios representativos de un modelo que bajo todos los escenarios propuestos quizás serían imposibles de visualizar.

2.8.1. Etapas de la Simulación

A continuación, se describen las etapas de una simulación:

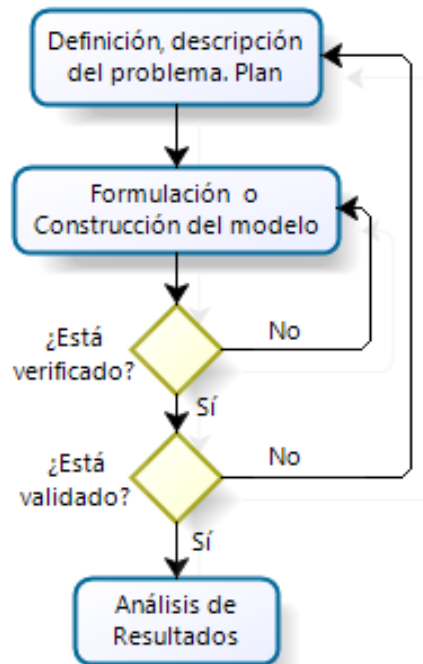


Figura 5. Diagrama de Flujo de una Simulación

2.8.1.1. Definición, descripción del problema. Plan.

En esta primera etapa se deben identificar los objetivos de la simulación, se especifican los índices de medición para determinar la efectividad del sistema a estudiar, se establecen objetivos específicos de los modelos y se definen los sistemas que van a ser modelados mediante la simulación.

2.8.1.2. Formulación o Construcción del modelo.

Una vez que se definen los resultados que se esperan del estudio, es necesario definir y construir un modelo con el cual se obtendrán los resultados deseados, se debe tener en cuenta que el propósito no es desarrollar el modelo en sí, sino que este debe resolver el problema para llegar al objetivo deseado. En esta etapa se deben definir las variables que forman parte del modelo, sus relaciones lógicas y los

diagramas de flujo que expliquen de manera completa al modelo.

Un modelo de simulación se refiere a un conjunto de hipótesis que describe el funcionamiento de un sistema que se expresa como relaciones matemáticas y/o lógicas entre varios elementos que conforman el sistema (Sotou, 2005).

2.8.1.3. Verificación del modelo.

En la verificación se necesita comprobar que el modelo simulado cumple con los requisitos de diseño para los cuales se construyó, es decir, se debe comprobar que el modelo de simulación sigue con las especificaciones del modelo conceptual.

2.8.1.4. Validación del modelo.

Durante esta etapa se deben valorar si existen diferencias entre la simulación y el sistema real que se está tratando de simular, comprueba que las hipótesis de trabajo sean las correctas. Algunas maneras de validar un modelo son:

- Opinión de expertos acerca de los resultados en una simulación.
- Predicción de datos históricos con exactitud.
- Predicción de datos futuros con exactitud.
- Comprobar fallas en el modelo mediante el uso de datos que hacen fallar al sistema real.
- Confianza y aceptación en el modelo por parte de la persona que utilizará los resultados que genere el experimento de simulación.
- Diseño de experimentos y plan de corridas.

Posterior a la validación del modelo se realiza una experimentación. En la experimentación se debe comprobar si los datos arrojados son los deseados y además se debe realizar un análisis de sensibilidad de los índices requeridos.

2.8.1.5. Análisis de resultados

En esta etapa se interpretan los resultados que genera la simulación y con estos se toman las decisiones. Los resultados obtenidos colaboran con el soporte de decisiones del tipo semi-estructurado.

Documentación de la simulación

Este documento sirve si en alguna ocasión alguien necesita volver a utilizar el modelo. La estructura de este documento podría ser: introducción, objetivos, hipótesis, descripción del sistema, descripción del modelo, análisis de los experimentos desarrollados y conclusiones.

Se pueden requerir de dos tipos de documentación para utilizar de mejor manera el modelo de simulación: documentación técnica y un manual de usuario para facilitar la interacción y la utilización del modelo construido.

2.8.2. Tipos de Simulación

2.8.3.1. Simulación continua:

Los estados del sistema cambian continuamente su valor con respecto al tiempo. Estas simulaciones se modelan generalmente con ecuaciones diferenciales. Algunos ejemplos de estos podrían ser:

- Sistemas físicos (mecánicos o eléctricos).
- Procesos físicos
- Dinámicas poblacionales.
- Ciertos modelos económicos.

2.8.3.2. Simulación por eventos discretos:

Es un modelo en el cual su comportamiento cambia en ciertos intervalos de tiempo dados. Los eventos del sistema o simulación son considerados como los momentos en los que se producen los cambios. Para aplicar esta técnica es necesario que las variables que definan al sistema no cambien su comportamiento durante el intervalo de tiempo simulado.

2.8.3.3. Simulación por autómatas celulares:

Es un modelo matemático que se aplica a casos complejos compuesto por un conjunto de subsistemas más pequeños conocidos como celdas o células que obtienen diferentes estados o valores. Estos pueden ser alterados de un momento a otro en unidades de tiempos discretos, por lo tanto, este conjunto de celdas o células consiguen evolucionar según una expresión matemática la cual posee sensibilidad a los estados de las células vecinas y se conoce con el nombre de regla de transición local.

2.8.3.4. Simulación de Monte Carlo:

Este tipo de simulación será descrito a detalle en el Capítulo 3.

2.9. Investigaciones Recientes

Se revisaron investigaciones recientes acerca de simulaciones de Monte Carlo para llegar a los resultados deseados.

2.9.1. Aplicación del Método Monte Carlo en Exploración y explotación de petróleo y gas natural

En esta investigación se realiza un análisis de la simulación de Monte Carlo como herramienta para resolver ecuaciones donde las variables necesarias para las mismas no se conocen con exactitud, sin embargo, se conocen los rangos de los datos reales de estas variables. Se da referencia al método, generación de números aleatorios e histogramas de frecuencia. Se trabajó mediante una hoja de cálculo para la simulación de Monte Carlo además de una simulación a la cual llamaron "MCP", el cual utiliza ecuaciones de gran importancia en el campo petrolífero.

La hoja de cálculo realizada tiene la capacidad de generar números aleatorios como si fueran los datos de las variables de entrada para posteriormente calcular los varios parámetros mediante las ecuaciones preestablecidas.

En resumen, este estudio usa hojas de cálculo para realizar una simulación de Monte Carlo, además de que se puede realizar cualquier ecuación que se necesite, además de que es una herramienta de uso común. El único problema que se genera es que se necesita desarrollar muchos pasos y chequeos de que las ecuaciones estén correctas antes de asegurarse de que la simulación sea exitosa (Santianello & Filippo, 2011).

2.9.2. Diseño e implementación de un sistema de inventarios, aplicando simulación Montecarlo, en una empresa de Servicios Petroleros

En este trabajo, se enfoca la simulación de Monte Carlo en los inventarios de una empresa, se definen las etapas de la simulación de Monte Carlo, en el cual se hace énfasis en la construcción de una distribución de probabilidades y la generación de números aleatorios. También se puede observar que se detallan las variables que influyen en el inventario.

Como consecuencia, se pudo determinar que la simulación de Monte Carlo no es precisa, pero tiene una aproximación muy cercana al sistema real. Mientras mayor cantidad de datos se generen, mayor será su precisión (Montenegro, 2011).

CAPÍTULO 3

3. METODOLOGÍA

3.1. Simulación de Monte Carlo

El análisis de riesgo es parte de cada decisión que se realiza. Constantemente las empresas y las personas se enfrentan con incertidumbres, ambigüedades y variabilidades, incluso si se tiene acceso a información sin precedentes, no existe forma de predecir el futuro con exactitud.

La Simulación de Monte Carlo es una técnica matemática computarizada cuya característica común es la utilización de números generados de forma aleatoria y el muestreo de distribuciones de probabilidad (Palisade, 2019). La simulación de Monte Carlo puede mostrar los extremos de las posibilidades, además de que, permite observar todos los escenarios posibles de las decisiones tomadas y evalúa el impacto del riesgo y, por lo tanto, se logra conseguir la mejor toma de decisiones bajo determinada incertidumbre incluso todas las posibles consecuencias de los posibles escenarios durante la simulación (antes de la toma de decisiones).

Esta técnica fue llamada Monte Carlo por los resorts de la ciudad de Mónaco renombrado por sus casinos y fue utilizada por primera vez por científicos que trabajaron en la bomba atómica. La simulación de Monte Carlo ha sido utilizada para modelar una variedad de sistemas físicos y conceptuales, desde su introducción en la Segunda Guerra Mundial (Palisade, 2019).

Estas simulaciones son utilizadas por profesionales en un sinnúmero de campos interdisciplinarios como finanzas, ingeniería, energía,

investigación y desarrollo, administración de proyectos, fabricación, servicios de seguros, gas y aceites, transportación y medio ambiente, entre otros.

La simulación de Monte Carlo realiza análisis de riesgos mediante construcción de modelos de posibles resultados utilizando un rango de valores (una distribución de probabilidad) aleatorios para cada factor que tiene una incertidumbre inherente.

Una secuencia de número aleatorios es aquella en la que es imposible predecir cuál será el siguiente número en una secuencia. Sin embargo, en computación las secuencias de números aleatorios que se usan en realidad son pseudoaleatorios, debido a que son generados por un algoritmo que se encarga de la secuencia sea lo suficientemente impredecible y que no se repita en ciclos muy cortos, de tal manera que la data disponible suficiente para ser considerada aleatoria. Estos algoritmos utilizan una semilla o número inicial como punto de partida para la generación de la secuencia. Además, esta secuencia de números aleatorios suele ser construida con distribución uniforme dentro del intervalo $(0,1)$, es decir, si se escoge una cantidad suficientemente grande de números de la secuencia se obtendrá la misma densidad de ellos en cada fracción de dicho intervalo.

De acuerdo con las recomendaciones formuladas por Cochran (1954), Roberts (2001), Flanagan (1993) y Morano (2003), la cantidad de números pseudoaleatorios (iteraciones) que pueden ser generados para la simulación deben ser determinados. De esta manera, la simulación se lleva a cabo por el tamaño de la muestra, las clases y la cantidad del conjunto de números pseudo aleatorios.

A continuación, se definen algunos conceptos de estadística que pueden ser de utilidad para entender estos métodos: La función de

distribución de probabilidad (FDP) de una variable x ($p(x)$) es una función que contiene la probabilidad de acierto para cada valor de x . Esta función no debe tener valores negativos como condición y debe estar normalizada (transformar a proporciones sin unidades de medida que permiten poder comparar elementos de distintas variables y distintas unidades de medida.) dentro de un intervalo $(x_{mín}, x_{máx})$.

$$p(x) \geq 0 \quad \equiv \quad \int_{x_{mín}}^{x_{máx}} p(x) dx = 1$$

La función de probabilidad acumulada (FPA) de una variable x es una función que contiene la probabilidad de acierto dentro de un intervalo $[x_{mín}, x]$. Por lo tanto, es una función monótona creciente con valor inicial $P(x_{mín})=0$ y $P(x_{máx})=1$. (López, 2016)

$$P(x) \equiv \int_{x_{mín}}^x p(x') dx'$$

La función generatriz a ser utilizada como función de distribución de probabilidad (FDP) debe presentar una forma semejante a la distribución de los elementos de la muestra para las clases a ser calculadas en condiciones predeterminadas.

Las distribuciones de probabilidad más comunes incluyen:

- a) Distribución Normal: Conocida también como campana de Gauss. Se puede definir la media o valor esperado y una desviación estándar para definir la variación de la media. Los valores en el medio cerca de la media son los más probables de ocurrencia. Esta distribución de probabilidad es simétrica y describe varios fenómenos naturales como por ejemplo altura de personas. Otros ejemplos podrían ser tasas de inflación y precios de energía.
- b) Distribución Logarítmica Normal: Cuando los valores están sesgados de forma positiva, no es simétrica como en la distribución

normal. Esta es utilizada para representar valores que no se encuentran bajo cero, pero tienen potenciales positivos ilimitados. Ejemplos de variables bajo esta distribución de probabilidad podrían incluir valores de propiedades del estado, precios de stock, reservas de aceite entre otras.

- c) Distribución Uniforme: Todos los valores tienen igual probabilidad de ocurrencia y se puede definir el máximo y el mínimo. Ejemplos de variables bajo esta distribución de probabilidad puede incluir costos de manufactura o futuras ventas de un nuevo producto.

- d) Distribución Triangular: Se define el mínimo y los valores máximos de mayor ocurrencia. Los valores alrededor del valor de mayor ocurrencia, tienen mayor probabilidad de ocurrir. Las variables que pueden ser descritas bajo esta distribución de probabilidad incluyen históricos de ventas por unidad y niveles de inventarios.

- e) Distribución PERT: Se define el mínimo y los valores máximos de mayor ocurrencia tal y como en la distribución triangular. Los valores alrededor del valor de mayor ocurrencia, tienen mayor probabilidad de ocurrir. Sin embargo, los valores entre el valor de mayor ocurrencia y sus extremos tienen mayor probabilidad de aparecer que en el triangular, esto es, los extremos no están enfatizados. Un ejemplo de distribución de probabilidad PERT podría ser la descripción de la duración de una tarea en un modelo de administración de proyectos (Palisade, 2019).

3.2. Pasos para Elaborar la Simulación de Monte Carlo

De acuerdo a Vélez (2003), Jorion (2006), Molina (2011) y Medina y Restrepo (2013) los pasos a seguir para la aplicación de la simulación Monte Carlo; adaptados para la construcción del modelo, son los siguientes:

- Identificar para cada una de las variables estocásticas las distribuciones de probabilidad con sus debidos parámetros. Esto son los volúmenes de servicio de análisis. Esto se encuentra explicado en la identificación de variables de entrada y de salida.
- Calcular las correlaciones entre las variables independientes, mediante matriz de correlación y otras técnicas. Incluir posibles dependencias entre las variables.
- Construir el modelo, comprobando todos sus supuestos.
- Generar números aleatorios o pseudoaleatorios, los cuales, como se indicó anteriormente, son comprendidos de 0 a 1.
- Alimentar los procesos estocásticos o las distribuciones de probabilidad inversas con los números aleatorios para determinar a qué clase pertenece cada probabilidad (número aleatorio).
- Calcular el número de simulaciones. Entre más grande sea la cantidad de simulaciones es mejor, esto basado en la ley de los grandes números. Cada corrida es un escenario posible y consistente basado en los supuestos establecidos en el modelo. Cada resultado debe ser guardado. En este estudio se realizaron los cálculos del número de simulaciones.

Cuando se realiza la corrida de simulación, los promedios de las variables del sistema tienen un tiempo de inestabilidad, posteriormente las variables tienden a estabilizarse y es cuando se considera que los valores de las variables son confiables. Existen algunas formas para estabilizar un modelo de simulación:

- Usar corridas suficientemente largas para que los datos de transición resulten insignificantes, este método es adecuado si la ejecución del modelo es muy breve.
- En caso de existan tiempos de transición más prolongados, se seleccionan las condiciones iniciales de arranque para que sean más representativas de la condición de estabilidad y por ende que se reduzca el periodo de transición. Este método tiene un problema

debido a que, si no se tiene una idea adecuada de la situación inicial, podría ocasionar una polarización de los resultados y producir en consecuencia un aumento en la varianza y por ende ocasiona tamaños de corrida más grandes.

- Otro método es determinar el momento en que los resultados obtenidos de las variables de salida lleguen a una estabilidad, esto se obtiene graficando los promedios de la variable de interés contra el tiempo de simulación y, cuando el promedio no cambia, se detiene la simulación.

El número de corridas de una simulación depende mucho de la distribución que se desea simular además de la bondad de los generadores de números aleatorios $U(0,1)$ que se utiliza y de las condiciones con que la simulación se inició. Por lo general para calcular el número de simulaciones se lo realiza mediante la fórmula:

$$n = \frac{\sigma^2(Z_{\alpha/2})^2}{\varepsilon^2}$$

En el cual:

Z → Estadístico normal para un determinado α

ε → Error estándar o desviación absoluta permitida sobre la muestra de la distribución a simular

σ^2 → Varianza de la distribución a simular

En caso de que la media y varianza de la distribución a simular se obtengan a partir de una población n_1 de 30 datos o menos, entonces este cálculo se modifica de tal forma que se utiliza la siguiente ecuación:

$$n = \frac{s^2(t_{n-1, \alpha/2})^2}{\varepsilon^2}$$

En el cual:

t → Estadístico de la distribución t-student

ε → Error estándar o desviación absoluta permitida sobre la muestra de la distribución a simular

s^2 → Estimador de la varianza de la distribución a simular

La ecuación anterior se emplea para calcular el número óptimo de corridas basada en una corrida de simulación de tamaño n_1 , esta corrida es conocida como prueba piloto. Se pueden utilizar ambas fórmulas únicamente cuando la información donde se obtienen los estimadores sigan estadísticamente una distribución normal (García, Sierra, & Guzmán, 2005).

Si los datos analizados siguen cualquier otra distribución, se debe utilizar el teorema de Tchebycheff de modo que el número de corridas se calcula mediante la ecuación:

$$n = \frac{m^2}{\alpha}$$

En el cual:

m^2 → Desviación estándar máximo permitido sobre la media de la distribución a simular

α → Probabilidad de error permitida

Para obtener el número de corridas óptimo de un modelo donde se tengan algunas variables de probabilidad, se realiza ejecutando el cálculo para cada una de ellas y se selecciona la mayor de todas ellas, entonces éste es el número de simulaciones del modelo (Rodríguez, 2016).

3.3. Descripción del Área de Trabajo

El presente trabajo se enfoca netamente en las actividades de análisis del laboratorio los cuales se basan en los procesos explicados en la Figura 6 a continuación:

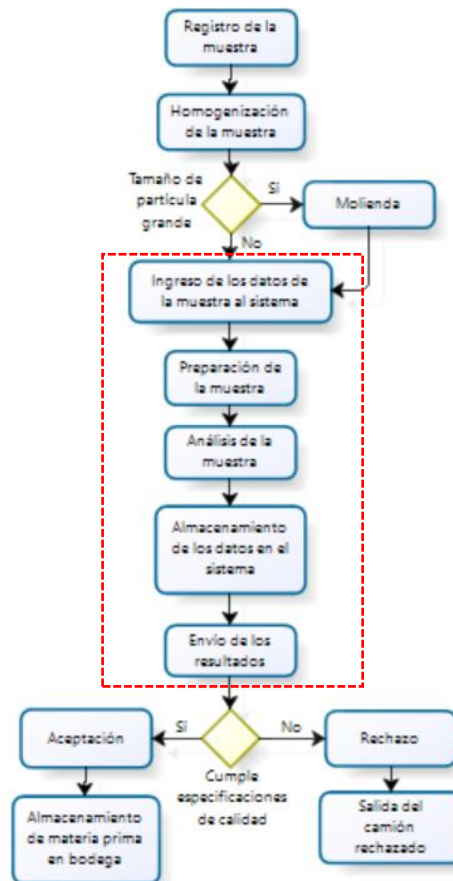


Figura 6. Diagrama de Proceso de los análisis de laboratorio

- Registro de la muestra: Toda muestra ingresada al laboratorio debe ser registrada en un formato para que quede constancia de que ha sido recibida y que debe ser analizada.
- Homogenización de la muestra: Es necesario para reducir de tamaño la muestra y asegurar que la mezcla resultante sea representativa de la cantidad de fundas, sacos o camiones al granel que fueron previamente muestreados. Generalmente la muestra que llega al laboratorio varía entre 1 a 3 kg y en el análisis únicamente se necesitan aproximadamente 10 g o menos.
- Molienda de la muestra: Existen materias primas que llegan en granos o demasiado gruesas, por este motivo, deben ser molidas para reducir su tamaño de partícula y proceder con su respectivo

análisis. Por el contrario, si las muestras son polvos finos, no es necesario realizar la molienda.

- Ingreso de la muestra al sistema: Es importante para ingresar toda la información relevante del producto al sistema respectivo tal como proveedor, lote, fechas de elaboración y caducidad, etc.
- Preparación de la muestra: Para preparar la muestra se necesitan tener los portamuestras del instrumento limpios y en buen estado, se necesita también colocar la cantidad correcta de muestra para que el análisis NIR sea preciso y no existan errores de predicción en los resultados.
- Análisis de la muestra: Se coloca el portamuestra al interior del instrumento NIR y se deja correr el instrumento para obtener los resultados.
- Almacenamiento de los datos en el sistema: Dado a que el sistema de estos análisis es automático, se espera que los resultados obtenidos en el análisis sean almacenados en la base de datos respectiva.
- Envío de los resultados: Mediante un correo electrónico, se deja saber a las áreas involucradas en este proceso si la materia prima cumple con las especificaciones y puede ser almacenada en las bodegas, o de lo contrario, la materia prima es rechazada y el camión que contiene el producto es enviado hacia afuera de la planta.

Este estudio fue considerado únicamente a partir del ingreso de la muestra al sistema (marcado en el cuadro rojo en la Figura 6), puesto

que, en los procesos anteriores a esta, no se pueden realizar cambios ya que son de vital importancia en el laboratorio.

Un día de trabajo, en el caso de los analistas de laboratorio empieza a las 08:00 y termina a las 17:00, con una hora de almuerzo que empieza a las 13:00 y finaliza a las 14:00.

3.4. Pasos Previos a la Simulación de Monte Carlo

Durante el estudio de este caso y previo a realizar la simulación, procedemos con los siguientes pasos:

3.4.1. Identificación de las Variables de Entrada y de Salida

A continuación, se identifican las variables de entrada y la variable de salida:

3.4.1.1. Variables de Entrada

Se identificaron cinco variables de entrada según el proceso explicado en la sección 3.3 de este capítulo:

- Tiempo de ingreso de los datos de la muestra al sistema: Es el tiempo que comprende el ingreso de toda la información de la materia prima tal como proveedor, fechas de elaboración y caducidad, etc. al sistema que es una base de datos cuya plataforma se encuentra en internet, esta es la etapa de inicio de los análisis hasta antes de iniciar la preparación de la muestra. Esta variable está identificada como T_{in} . A continuación, se muestra un histograma de distribución de estos datos:

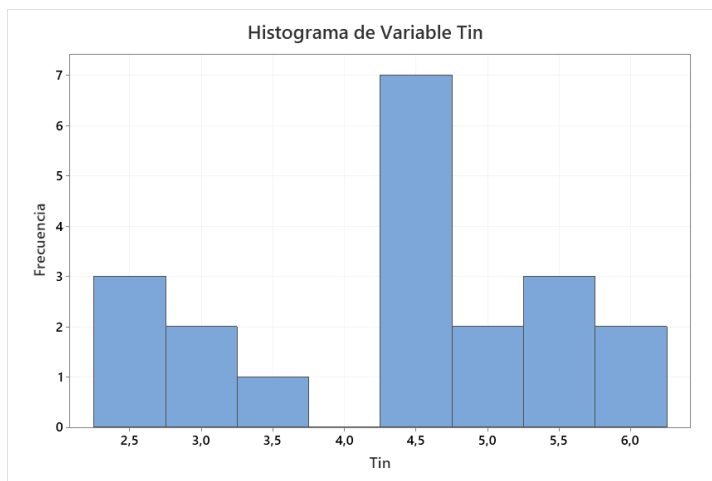


Figura 7. Histograma de la variable tiempo de ingreso de los datos al sistema T_{in}

Se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{in} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0.025
Exponencial	<0.003
Weibull	0.028
Gamma	<0.005

Tabla 1. Valor-p de la variable ingreso de los datos al sistema

De acuerdo con los valores-p mostrados en la Tabla 1, estos son <0.05, se rechaza la hipótesis nula. Por lo tanto, se puede inferir que ninguna de las distribuciones propuestas se ajusta al comportamiento de los datos de la variable T_{in} . Por consiguiente, se calcula como una Distribución Empírica:

$$f(x) \begin{cases} 0.15; 2.25 \leq x < 2.75 \\ 0.10; 2.75 \leq x < 3.25 \\ 0.05; 3.25 \leq x < 3.75 \\ 0 & ; 3.75 \leq x < 4.25 \\ 0.35; 4.25 \leq x < 4.75 \\ 0.10; 4.75 \leq x < 5.25 \\ 0.15; 5.25 \leq x < 5.75 \\ 0.10; 5.75 \leq x < 6.25 \end{cases}$$

- Tiempo de preparación de la muestra:
Corresponde al tiempo que toma el analista para colocar la muestra en las celdas portamuestras, esta etapa se realiza después del ingreso de los datos de la muestra y previo a realizar el análisis NIR. Esta variable se encuentra identificada como T_{pr} . A continuación, se muestra un histograma de distribución de estos datos:

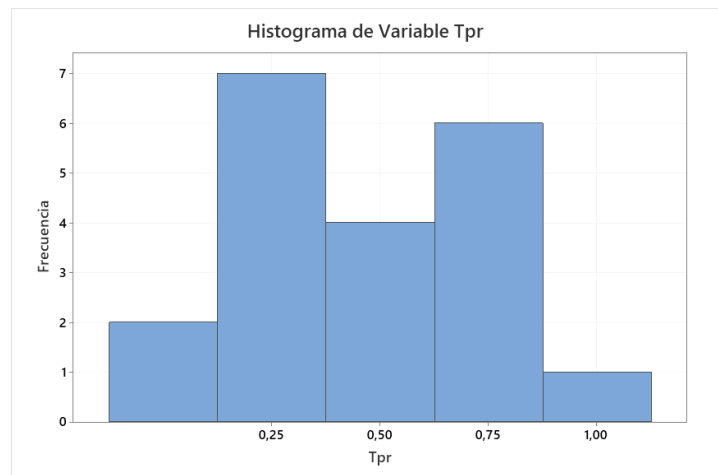


Figura 8. Histograma de los datos de la Variable T_{pr}

Se probaron 4 distribuciones para determinar cuál de ellas se ajusta al comportamiento de los datos de la variable T_{pr} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,074
Exponencial	0,067
Weibull	0,056
Gamma	0,045

Tabla 2. Valor-p de la variable preparación de la muestra

De acuerdo a los valores-p obtenidos en la Tabla 2, el comportamiento de los datos podría ajustarse a una Distribución Normal o Exponencial puesto a que tienen valores >0.05 , sin embargo, se toma el valor-p más alto, entonces, no se tiene evidencia estadística para rechazar la hipótesis nula y, por lo tanto, se considera que los datos de la variable T_{pr} siguen una Distribución Normal.

- Tiempo de análisis de la muestra: Corresponde al tiempo que toma el equipo en obtener los resultados de la muestra analizada, el análisis per se, se realiza después de la preparación de la muestra hasta obtener los resultados. Esta variable se encuentra identificada como T_{an} . A continuación, se muestra un histograma de distribución de estos datos:

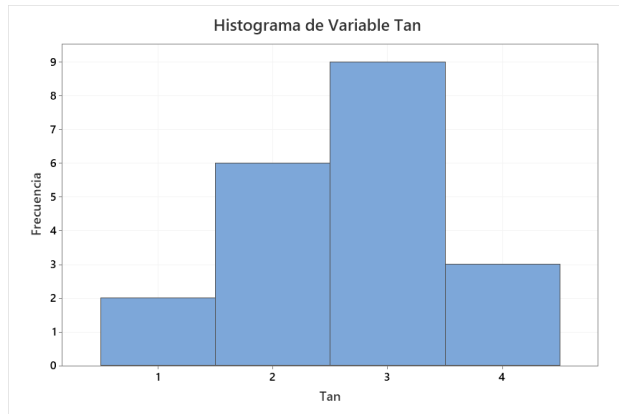


Figura 9. Histograma de los datos de la variable T_{an}

Se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{an} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,78
Exponencial	<0,003
Weibull	<0,003
Gamma	<0,005

Tabla 3. Valor-p de la variable análisis de la muestra

De acuerdo a los valores-p obtenidos en la Tabla 3, la Distribución Normal tiene un valor >0.05 , entonces, no se tiene evidencia estadística para rechazar la hipótesis nula y, por lo tanto, se considera que los datos de la variable T_{an} siguen una Distribución Normal.

- Tiempo de almacenamiento de los resultados en el sistema: Es el tiempo que toma al sistema en capturar los resultados obtenidos en el análisis,

este paso es automático y comprende desde que se obtienen los resultados hasta que estos son almacenados en la base de datos. Esta variable se encuentra identificada como T_{al} . A continuación, se muestra un histograma de distribución de estos datos:

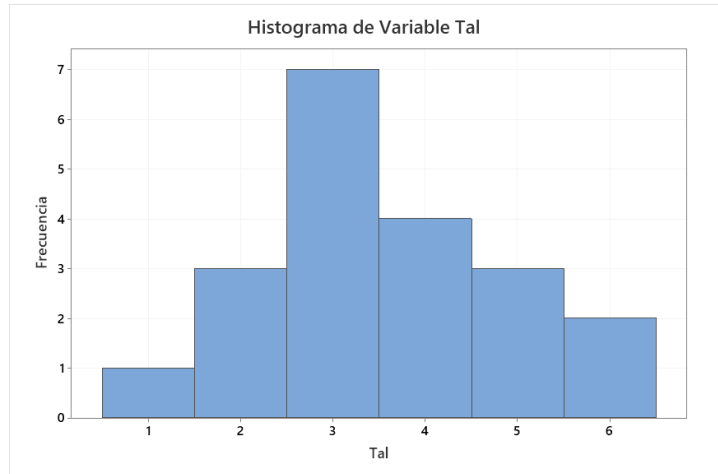


Figura 10. Histograma de los datos de la variable T_{al} .

Se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{al} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,689
Exponencial	<0,003
Weibull	<0,003
Gamma	<0,005

Tabla 4. Valor-p de la variable almacenamiento de los datos en el sistema

De acuerdo a los valores-p obtenidos en la Tabla 4, la Distribución Normal tiene un valor >0.05 , entonces, no se tiene evidencia estadística para rechazar la

hipótesis nula y, por lo tanto, se considera que los datos de la variable T_{al} siguen una Distribución Normal.

- Tiempo de envío de resultados: Corresponde al tiempo que toma el analista en elaborar un reporte de los resultados obtenidos mediante un correo electrónico, comprende desde que se almacenan los resultados en la base de datos hasta que el correo electrónico ha sido enviado a todas las partes interesadas en el resultado. Esta variable se encuentra identificada como T_{en} . A continuación, se muestra un histograma de distribución de estos datos:

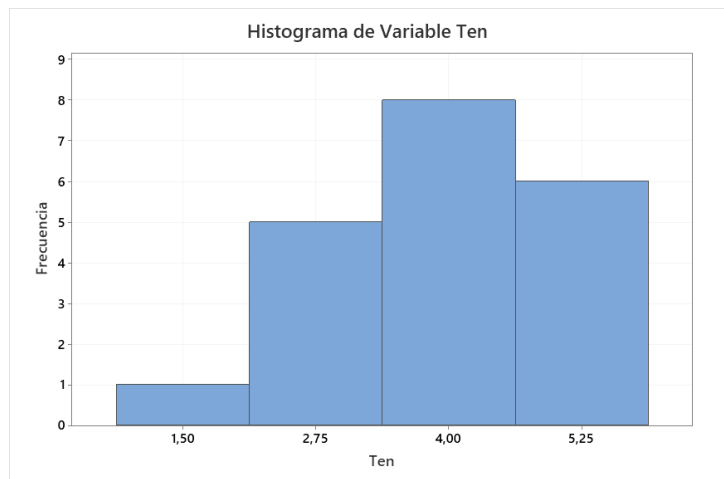


Figura 11. Histograma de los datos de la variable T_{en}

Se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{en} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,355
Exponencial	<0,003
Weibull	0,251
Gamma	0,205

Tabla 5. Valor-p de la variable tiempo de envío de resultados

Según los valores-p obtenidos en la Tabla 5, en el valor-p de la Distribución Normal, la distribución Weibull y la distribución gamma tienen valores $>0,05$ y en este caso se toma el valor-p más alto, entonces, no se tiene evidencia estadística para rechazar la hipótesis nula y, por lo tanto, se considera que los datos de la variable T_{en} siguen una Distribución Normal.

3.4.1.2. Variable de Salida

Los datos obtenidos de las mediciones entre los meses de Septiembre y Octubre del 2019 se encuentran presentadas como valores medios de los minutos que se toman en realizar un análisis (minutos/análisis). Esta variable se encuentra identificada como Mnts.

Posteriormente, esta variable de salida será transformada en términos de productividad (número de análisis/día de trabajo). Cabe indicar que el horario de trabajo va desde las 8:00 hasta las 17:00 con una hora de almuerzo que va desde las 13:00 hasta las 14:00, es decir, un día de trabajo consta de 8 horas laborables (480 minutos), entonces la transformación se la realiza mediante la siguiente fórmula:

$$\frac{\text{Análisis}}{\text{día}} = \frac{480 \text{ min/día}}{n \text{ min/análisis}} = \frac{480}{n} \text{ análisis/día}$$

A continuación, se muestra un histograma con la distribución de los datos obtenidos:

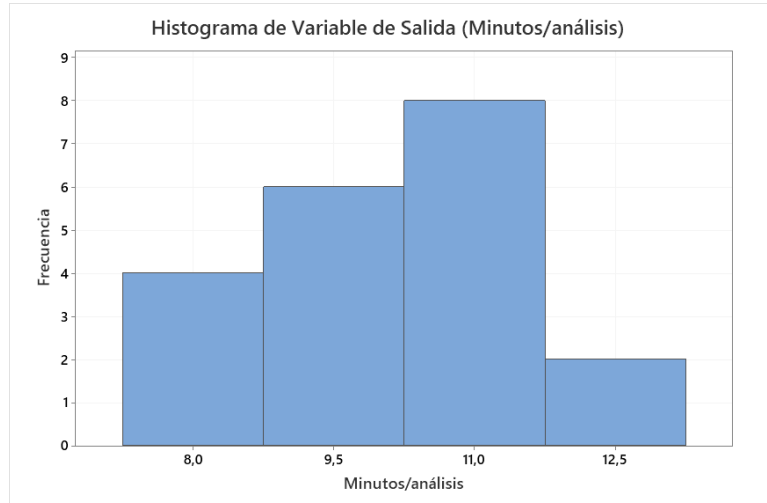


Figura 12. Histograma de los datos de la variable de salida Mnts (minutos/análisis)

Se probaron 4 distribuciones para determinar cuál de ellos se ajusta a la distribución de los datos de la variable Mnts mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,274
Exponencial	<0,003
Weibull	0,219
Gamma	0,251

Tabla 6. Valor-p de la variable de salida Mnts (minutos/análisis)

Según los valores-p obtenidos en la Tabla 6, en el valor-p de la distribución normal, la distribución Weibul y la distribución gamma tienen valores >0,05

y en este caso se toma el valor-p más alto, entonces, no se tiene evidencia estadística para rechazar la hipótesis nula y, por lo tanto, se considera que los datos de la variable de salida Mnts siguen una distribución normal.

3.4.2. Cálculo de las Correlaciones

Para este caso se utilizaron los datos tomados entre los meses de Septiembre a Octubre del 2019. A partir de estos datos tomados, se obtuvo un modelo el cual describe la situación inicial del laboratorio.

Con los datos obtenidos a través de las mediciones de los tiempos, procedemos con la construcción del modelo de la situación inicial en el laboratorio. Como primer paso elaboramos una matriz de correlación para determinar si existe algún problema de multicolinealidad entre las variables:

```
> cor(Datos)
      Tin      Tpr      Tan      Tal      Ten      Mnts
Tin  1.00000000 -0.037247358  0.34913250 -0.2400806  0.328537444  0.6137862
Tpr  -0.03724736  1.000000000  0.05282465  0.1378738  0.003952431  0.1804886
Tan  0.34913250  0.052824655  1.00000000  -0.1167116  0.377432795  0.6087820
Tal  -0.24008060  0.137873844 -0.11671159  1.0000000  -0.192122955  0.3934099
Ten  0.32853744  0.003952431  0.37743280 -0.1921230  1.000000000  0.5262257
Mnts 0.61378624  0.180488561  0.60878203  0.3934099  0.526225654  1.0000000
```

Figura 13. Matriz de Correlación de los datos de la situación inicial en el programa R

Dado a que los valores de las variables en la matriz de correlación no son muy altos, podemos indicar que no existen problemas de multicolinealidad. Dado a que no existen problemas de multicolinealidad, podemos construir el modelo correspondiente comprobando todos sus supuestos.

3.4.3. Construcción del modelo

Se procedió a realizar un modelo estadístico propuesto que posiblemente que pueda determinar la situación inicial del

laboratorio. En este caso se empleará un método de selección de variables mediante eliminación *stepwise* en el cual se inicia el modelo sin eliminar ninguna variable “Start AIC” y se eliminan o incluyen variables hasta obtener un modelo seleccionando las mejores variables predictoras con la medición *Akaike* (AIC) de menor valor.

```
> modelo1<-lm(Datos$Mnts~DTin+DTpr+DTan+DTal+DTen,data=Datos)
> step(object=modelo1,direction ="both",trace=1)
Start: AIC=-58.24
Datos$Mnts ~ DTin + DTpr + DTan + DTal + DTen

      Df Sum of Sq    RSS    AIC
<none> 1      0.5969  0.5969 -58.236
- DTpr  1      0.2615  0.8584 -52.968
- DTen  1      2.6361  3.2330 -26.446
- DTan  1      3.1722  3.7691 -23.378
- DTin  1      6.6637  7.2605 -10.266
- DTal  1     10.2050 10.8019  -2.320

Call:
lm(formula = Datos$Mnts ~ DTin + DTpr + DTan + DTal + DTen, data = Datos)

Coefficients:
(Intercept)      DTin      DTpr      DTan      DTal      DTen
  2.5475      0.5551      0.4200      0.5774      0.5520      0.4129
```

Figura 14. Medición de Akaike para determinar las variables significativas en el modelo de la situación inicial del laboratorio bromatológico

Según la Figura 14, en la medición de Akaike, el menor AIC (-58.236 por ser negativo) se encuentra con todas las variables iniciales (<none>), mientras que al eliminar cualquiera de las variables el valor de AIC se incrementa, por lo cual existe indicio de que la mejor propuesta de modelo que se ajusta a la realidad del sistema es aquella a la cual no se elimina ninguna variable. Por este motivo, podemos indicar preliminarmente que el modelo de regresión lineal propuesto para la situación inicial del laboratorio sería:

$$\text{Minutos/análisis} = \beta_0 + \beta_1 T_{in} + \beta_2 T_{pr} + \beta_3 T_{an} + \beta_4 T_{al} + \beta_5 T_{en}$$

Con este antecedente, procedemos a determinar los coeficientes β_0 , β_1 , β_2 , β_3 , β_4 y β_5 mediante la siguiente prueba de hipótesis con 0.05 de significancia:

$$H_0: \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$H_1: \beta_0 \neq \beta_1 \neq \beta_2 \neq \beta_3 \neq \beta_4 \neq \beta_5 \neq 0$

> summary(mode1o1)

```
Call:
lm(formula = Datos$Mnts ~ DTin + DTpr + DTan + DTa1 + DTen, data = Datos)

Residuals:
    Min       1Q   Median       3Q      Max
-0.22538 -0.13806 -0.01688  0.11909  0.33474

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.54752    0.30349   8.394 7.79e-07 ***
DTin         0.55505    0.04440  12.502 5.52e-09 ***
DTpr         0.42001    0.16958   2.477  0.0266 *
DTan         0.57742    0.06694   8.626 5.64e-07 ***
DTa1         0.55198    0.03568  15.471 3.38e-10 ***
DTen         0.41288    0.05251   7.863 1.67e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2065 on 14 degrees of freedom
Multiple R-squared:  0.9799,    Adjusted R-squared:  0.9728
F-statistic: 136.8 on 5 and 14 DF,  p-value: 2.26e-11
```

Figura 15. Modelo estadístico propuesto de la situación inicial del laboratorio bromatológico

Según el modelo de regresión lineal propuesto tiene un R^2 de 0.9799, por lo cual podríamos indicar que este modelo es capaz de explicar el 97.99% de la variable observada (minutos/análisis). El valor-p del modelo es significativo ($2.26e-11$) por lo que se rechaza la hipótesis nula y, por lo tanto, podemos indicar que al menos uno de los coeficientes parciales de regresión es distinto de 0. También se pueden observar los coeficientes para el modelo propuesto los cuales son $\beta_0=2.54752$, $\beta_1=0.55505$, $\beta_2=0.42001$, $\beta_3=0.57742$, $\beta_4=0.55198$ y $\beta_5=0.41288$.

Para poder determinar si este modelo de regresión lineal propuesto es adecuado, es necesario que se cumpla con todos los supuestos correspondientes:

a) Supuestos del modelo de la situación inicial

Linealidad

Esta condición se valida mediante diagramas de dispersión entre la variable dependiente y cada una de las variables o

con diagramas de dispersión entre cada una de las variables y los residuos del modelo (Amat, 2020). En este caso realizamos un diagrama de dispersión entre los residuos del modelo y cada una de las variables, como se muestra a continuación:

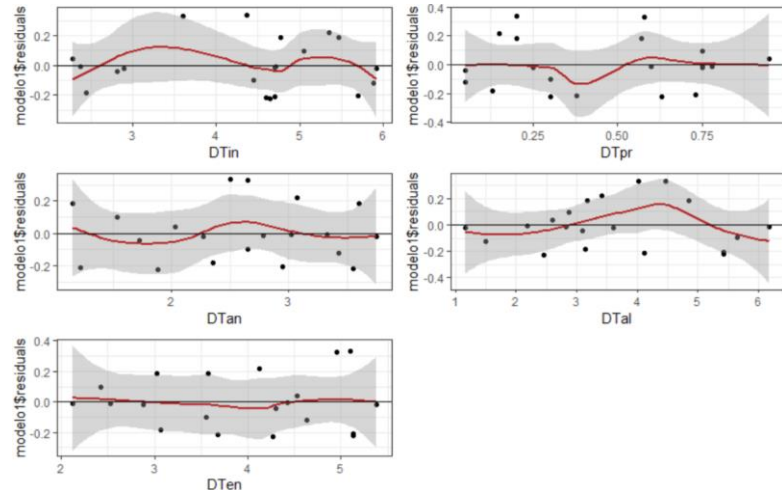


Figura 16. Diagrama de dispersión entre los residuos del modelo propuesto y las variables de la situación inicial

Tal y como se muestra en la Figura 16, dado a que la relación es lineal, los residuos se distribuyen en torno a 0 con variabilidad aproximadamente constante a lo largo del eje X. Entonces, se cumple la linealidad para cada una de las variables de entrada.

Independencia

Independencia entre los residuos mediante el estadístico de Durbin-Watson que toma valor 2 cuando los residuos son completamente independientes (entre 1.5 y 2.5 se considera que existe independencia). La prueba de hipótesis con 0.05 de significancia es la siguiente:

H_0 : Existe independencia

H_1 : No existe independencia

```

> dwtest(modelo1,alternative = "two.sided")

Durbin-Watson test

data: modelo1
DW = 2.7572, p-value = 0.08547
alternative hypothesis: true autocorrelation is not 0

```

Figura 17. Prueba de Durbin-Watson del modelo propuesto para la situación inicial

Tal y como se observa en la Figura 17, se observa que el valor-p de la prueba de hipótesis es mayor a 0.05, entonces no se tiene suficiente evidencia estadística para rechazar la hipótesis nula y, por lo tanto, podemos indicar que existe independencia entre las variables.

Homocedasticidad

Se procede a realizar la prueba de Breusch-Pagan para determinar si existe homocedasticidad, la prueba de hipótesis con 0.05 de significancia es la siguiente:

H₀: Existe homocedasticidad.

H₁: No existe homocedasticidad.

```

> bptest(modelo1)

studentized Breusch-Pagan test

data: modelo1
BP = 9.1809, df = 5, p-value = 0.1021

```

Figura 18. Prueba de homocedasticidad de Breusch-Pagan del modelo propuesto de la situación inicial

Dado a que los valores-p en ambos casos son mayores a 0.05, entonces, no existe suficiente evidencia estadística para rechazar la hipótesis nula y, por lo tanto, se puede interpretar entonces que, existe homocedasticidad.

Normalidad de los residuos

Luego a esto, se procede a realizar la prueba de normalidad de Shapiro-Wilk en los errores o residuos, la prueba de hipótesis con 0.05 de significancia es la siguiente:

H₀: Existe normalidad en los residuos.

H₁: No existe normalidad en los residuos.

```
> r1=residuals(modelo1)
> shapiro.test(r1)
```

Shapiro-Wilk normality test

```
data: r1
W = 0.92409, p-value = 0.1188
```

Figura 19. Prueba de normalidad de Shapiro-Wilk en los residuos del modelo propuesto en la situación inicial

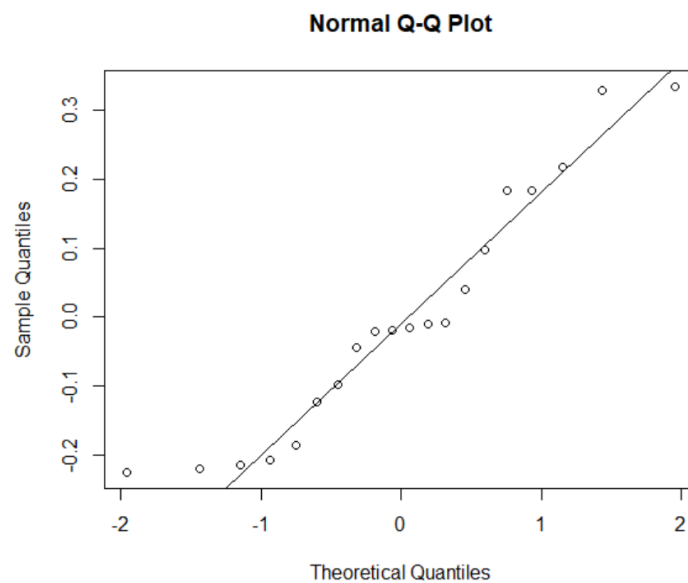


Figura 20. Gráfico Q-Q de normalidad de los residuos del modelo propuesto del laboratorio en la situación inicial

Dado a que el valor-p es mayor a 0.05, no se tiene suficiente evidencia estadística para rechazar la hipótesis nula y, por lo tanto, podemos confirmar la normalidad de los residuos. Además, podemos visualizar la linealidad de los residuos en la Figura 20.

No Colinealidad

Determinamos si en el modelo existen problemas de colinealidad mediante un análisis de inflación de varianza (VIF):

```
> vif(modelo1)
      DTin      DTpr      DTan      DTal      DTen
1.244223 1.025448 1.254815 1.098024 1.246929
```

Figura 21. Análisis de inflación de Varianza (VIF) del modelo propuesto en la situación inicial

En este caso no existen problemas de colinealidad dado a que los valores de los VIF son inferiores a 10 ($1 \leq \text{VIF} \leq 10$).

Finalmente, dado a que se cumplen los supuestos del modelo lineal, podemos indicar entonces que el modelo de la situación inicial del laboratorio es:

$$\text{Minutos/análisis} = 2.54752 + 0.55505T_{in} + 0.42001T_{pr} + 0.57742T_{an} + 0.55198T_{al} + 0.41288T_{en}$$

3.5. Verificación de la simulación

En esta etapa se realiza una prueba piloto de 20 simulaciones al igual que los 20 datos iniciales y se generan los valores de las variables aleatorias (V.A.) obtenidas a partir de los números aleatorios (N.A.) o también llamada probabilidad aleatoria que va de 0 a 1 y estas dependen de las funciones de probabilidad de cada una de las variables aleatorias. En las siguientes tablas se observarán los números aleatorios junto con los valores de las variables aleatorias obtenidas dependiendo de su función de probabilidad:

N.A. T_{in}	V.A. T_{in}
0,52	4,50
0,94	6,00
0,08	2,50
0,23	3,00
0,22	3,00
0,36	4,50
0,77	5,50

0,88	5,50
0,80	5,50
0,55	4,50
0,89	5,50
0,63	4,50
0,01	2,50
0,99	6,00
0,92	6,00
0,91	6,00
0,28	3,50
0,82	5,50
0,08	2,50
0,35	4,50

Tabla 7. Números aleatorios generados junto con los valores de la variable aleatoria T_{in}

N.A. T_{pr}	V.A. T_{pr}
0,33	0,33
0,21	0,23
0,43	0,40
0,94	0,90
0,85	0,74
0,27	0,28
0,65	0,56
0,19	0,21
0,43	0,40
0,39	0,38
0,23	0,24
0,06	0,02
0,89	0,80
1,00	1,38
0,12	0,13
0,87	0,77
0,27	0,28
0,94	0,90
0,15	0,16
0,33	0,33

Tabla 8. Números aleatorios generados junto con los valores de la variable aleatoria T_{pr}

N.A. T_{an}	V.A. T_{an}
0,84	3,36
0,85	3,38
0,46	2,49
0,93	3,73
1,00	4,79
0,27	2,08
0,12	1,63
0,35	2,26
0,39	2,35
0,41	2,39
0,13	1,67
0,43	2,43
0,10	1,55
0,01	0,69
0,02	0,97
0,77	3,15
0,52	2,61
0,70	2,98
0,18	1,83
0,67	2,93

Tabla 9. Números aleatorios generados junto con los valores de la variable aleatoria T_{an}

N.A. T_{al}	V.A. T_{al}
0,47	3,49
0,91	5,44
0,26	2,73
0,87	5,16
0,80	4,78
0,56	3,81
0,15	2,18
0,65	4,15
0,18	2,36
0,40	3,25
0,41	3,29
0,92	5,57
0,76	4,61
0,73	4,45
0,15	2,18
0,09	1,73
0,65	4,14
0,15	2,18
0,02	0,88
0,47	3,52

Tabla 10. Números aleatorios generados junto con los valores de la variable aleatoria T_{al}

N.A. T_{en}	V.A. T_{en}
0,49	3,93
0,12	2,75
0,29	3,40
0,14	2,83
0,11	2,71
0,50	3,93
0,56	4,08
0,70	4,46
0,94	5,49
0,87	5,09
0,53	4,03
0,97	5,81
0,46	3,85
0,81	4,82
0,09	2,58
0,92	5,39
0,03	2,00
0,77	4,68
0,11	2,69
0,49	3,90

Tabla 11. Números aleatorios generados junto con los valores de la variable aleatoria T_{en}

Mediante los valores de las variables aleatorias presentadas en las tablas anteriores se obtuvieron los siguientes resultados en la prueba piloto:

No. Simulación	Minutos/ análisis	No. Análisis/ día
1	10,67	45
2	12,07	40
3	8,45	57
4	10,76	45
5	11,05	43
6	10,09	48
7	9,66	50
8	11,12	43
9	10,70	45
10	10,48	46
11	10,14	47
12	11,93	40
13	9,30	52
14	11,30	42

15	8,76	55
16	11,20	43
17	9,23	52
18	10,83	44
19	6,66	72
20	10,43	46

Tabla 12. Minutos por análisis y número de análisis por día obtenidos de la simulación propuesta de la situación inicial del laboratorio

Para mejor comprensión, se realiza un comparativo gráfico mediante el cual se puede observar si los datos reales son similares a los datos obtenidos mediante la simulación piloto:

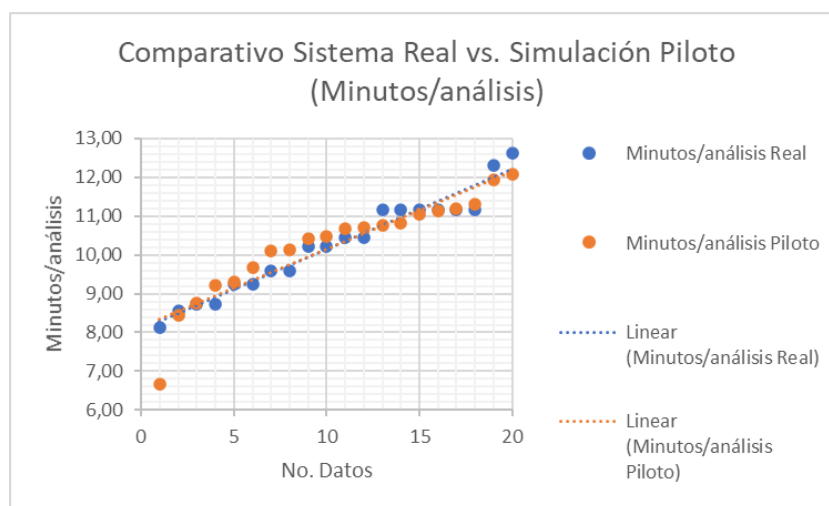


Figura 22. Comparativo del Sistema Real vs. Simulación Piloto (Minutos/análisis)

Tal como se observa en la Figura 22, si se analiza las pendientes del sistema real y de la simulación piloto de la variable de salida (Minutos/análisis) son muy parecidos, esto da un indicio de que no existe diferencia entre ambos.

3.6. Cálculo del número de simulaciones

Para calcular el número simulaciones, se debe determinar para cada una de las variables de entrada utilizadas en el modelo de simulación en la cual:

Para la variable T_{in} , dado a que los datos de esta variable siguen una distribución empírica, se utiliza la siguiente ecuación:

$$n = \frac{m^2}{\alpha}$$

Donde:

m: Desviación estándar = 1.29

α : Probabilidad de error permitida = 0.05

Entonces:

$$n = \frac{1.29^2}{0.05} = 33.282 \cong \mathbf{33}$$

Para la variable T_{pr} , dado a que los datos de esta variable siguen una distribución normal con una cantidad de datos menores a 30, se utiliza la siguiente ecuación:

$$n = \frac{s^2(t_{n-1,\alpha/2})^2}{\varepsilon^2}$$

Donde:

s^2 : Varianza = 0.34

ε : Error estándar = 0.076

$t_{19, 0.05}$: Estadístico de la distribución t-student = 2.093

Entonces:

$$n = \frac{0.34^2(2.093)^2}{0.076^2} = 87.674 \cong \mathbf{88}$$

Para la variable T_{an} , dado a que los datos de esta variable siguen una distribución normal con una cantidad de datos menores a 30, se utiliza la siguiente ecuación:

$$n = \frac{s^2(t_{n-1,\alpha/2})^2}{\varepsilon^2}$$

Donde:

s^2 : Varianza = 0.97

ε : Error estándar = 0.217

$t_{19, 0.05}$: Estadístico de la distribución t-student = 2.093

Entonces:

$$n = \frac{0.97^2(2.093)^2}{0.217^2} = 87.531 \cong \mathbf{88}$$

Para la variable T_{al} , dado a que los datos de esta variable siguen una distribución normal con una cantidad de datos menores a 30, se utiliza la siguiente ecuación:

$$n = \frac{s^2(t_{n-1,\alpha/2})^2}{\varepsilon^2}$$

Donde:

s^2 : Varianza = 1.31

ε : Error estándar = 0.293

$t_{19, 0.05}$: Estadístico de la distribución t-student = 2.093

Entonces:

$$n = \frac{1.31^2(2.093)^2}{0.293^2} = 87.568 \cong \mathbf{88}$$

Y, por último, para la variable T_{en} , dado a que los datos de esta variable siguen una distribución normal con una cantidad de datos menores a 30, se utiliza la siguiente ecuación:

$$n = \frac{s^2(t_{n-1,\alpha/2})^2}{\varepsilon^2}$$

Donde:

s^2 : Varianza = 1.09

ε : Error estándar = 0.244

$t_{19, 0.05}$: Estadístico de la distribución t-student = 2.093

Entonces:

$$n = \frac{1.09^2(2.093)^2}{0.244^2} = 87.420 \cong \mathbf{87}$$

Entonces para este estudio, necesitamos realizar un total de 88 simulaciones. La siguiente tabla describe los valores de minutos tomados por un análisis y su transformación a número de análisis por

día obtenidos a través de la simulación con el modelo de regresión propuesto anteriormente:

Sim.	Minutos/ análisis	No.Análisis/ día	Sim.	Minutos/ análisis	No.Análisis/ día
1	10,29	46,63	45	9,65	49,73
2	10,55	45,48	46	10,41	46,12
3	10,48	45,81	47	10,70	44,88
4	9,03	53,17	48	10,31	46,56
5	9,08	52,87	49	11,81	40,63
6	10,44	45,99	50	10,90	44,04
7	12,37	38,80	51	8,86	54,15
8	12,18	39,41	52	11,14	43,08
9	11,09	43,26	53	10,42	46,06
10	9,95	48,23	54	11,57	41,48
11	11,09	43,27	55	11,96	40,15
12	12,31	38,99	56	10,57	45,43
13	11,53	41,65	57	11,06	43,39
14	9,71	49,44	58	10,86	44,21
15	9,88	48,59	59	10,32	46,53
16	11,66	41,17	60	10,95	43,83
17	11,37	42,23	61	8,72	55,05
18	10,85	44,22	62	8,19	58,60
19	9,36	51,26	63	9,40	51,04
20	10,52	45,61	64	10,12	47,45
21	9,46	50,76	65	10,12	47,43
22	9,88	48,61	66	12,67	37,90
23	12,05	39,84	67	11,38	42,17
24	8,72	55,06	68	10,15	47,27
25	9,84	48,78	69	12,35	38,87
26	10,41	46,11	70	8,64	55,55
27	9,99	48,03	71	11,36	42,25
28	12,81	37,47	72	9,44	50,83
29	10,56	45,45	73	10,58	45,38
30	8,56	56,08	74	9,90	48,46
31	9,97	48,16	75	9,39	51,13
32	10,00	47,99	76	9,86	48,67
33	11,15	43,03	77	10,40	46,15
34	10,46	45,88	78	9,01	53,27
35	10,00	48,00	79	10,22	46,98
36	9,75	49,26	80	11,04	43,49
37	9,96	48,20	81	10,34	46,43

38	8,80	54,55	82	10,69	44,92
39	10,47	45,83	83	10,17	47,18
40	10,01	47,94	84	9,50	50,50
41	8,84	54,31	85	9,88	48,60
42	8,93	53,74	86	8,69	55,24
43	8,00	59,97	87	10,32	46,51
44	9,55	50,28	88	8,51	56,44

Tabla 13. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo

3.7. Validación del Modelo de situación inicial

3.7.1. Estadística Descriptiva de la Simulación

En la Tabla 16 se muestra el resumen de los estadísticos descriptivos de la variable de respuesta (minutos/análisis), así como la transformación en términos de productividad de los datos obtenidos de la simulación y de los datos obtenidos a partir de las mediciones reales:

Estadístico	Minutos/análisis		Análisis/día	
	Real	Simulación	Real	Simulación
μ	10,25	10,29	47,5	47,1
σ	1,25	1,05	5,9	4,9
Mín	8,14	8,00	38,0	37,5
Máx	12,63	12,37	59,0	55,1
Median	10,32	10,31	46,5	46,5
Q1	9,23	9,67	43,0	43,9
Q3	11,16	10,92	52,0	49,7

Tabla 14. Resumen estadístico de los minutos por análisis y del número de análisis diarios obtenidos a través de los datos reales y de la simulación

La variable de decisión en este estudio es el número de análisis diarios obtenidos a partir de la variable de respuesta minutos por análisis. En la tabla anterior, los valores de las medias varían muy poco, los valores de las medianas son casi idénticas mientras que podemos ver que la desviación estándar es algo menor en la simulación que en el sistema real.

A continuación, podremos observar si existe una diferencia significativa entre los datos reales y los obtenidos a través de la simulación:

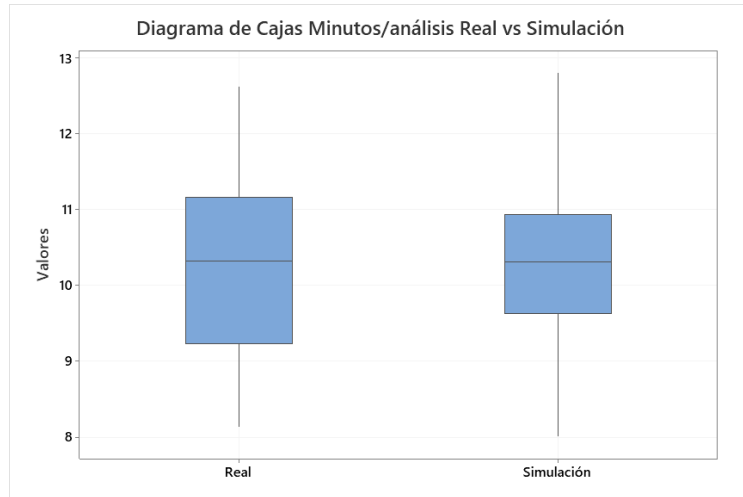


Figura 23. Diagrama de cajas del comparativo de la distribución de la variable de salida (minutos/análisis) de los datos reales vs los datos de la simulación

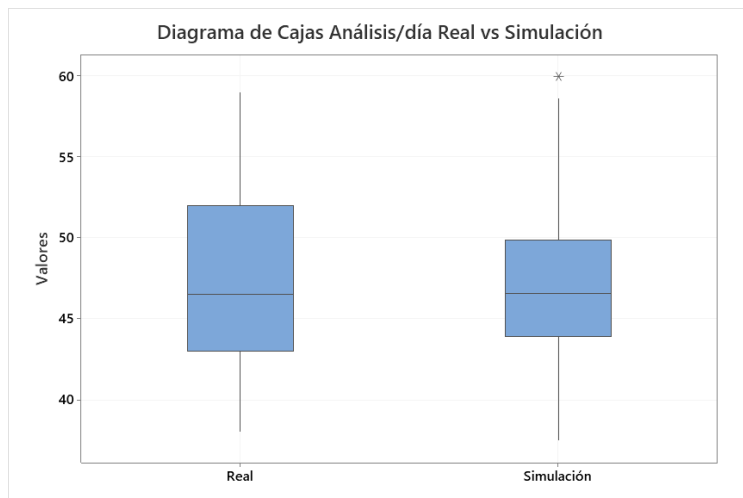


Figura 24. Diagrama de cajas del comparativo del número de análisis diarios de los datos reales vs los datos de la simulación

En el diagrama presentado en la Figura 24, se puede observar la diferencia entre la distribución de los datos de la variable de salida (minutos/análisis) en el sistema real y en la simulación. En este caso, los cuartiles representados de los datos reales son: $Q1=9.23$, $Q2=10.32$, $Q3=11.16$ y $Q4=12.63$, mientras que

en los datos obtenidos a través de la simulación son: $Q1=9.67$, $Q2=10.31$, $Q3=10.92$ y $Q4=12.37$. En esta representación, se puede observar que la dispersión de los datos reales y de la simulación es asimétrica negativa, es decir, que los datos se concentran mayormente por encima del valor de la mediana ($Q2$). Por último, si comparamos sus variabilidades, se observa que los datos obtenidos a través de la simulación se tiene una menor variabilidad (o rango intercuartil) de 1.31 minutos/análisis comparado con los datos reales (1.93 minutos/análisis). Por lo tanto, podemos concluir que el modelo propuesto de la situación inicial del laboratorio describe lo suficientemente bien el sistema real.

En el diagrama presentado en la Figura 25, se puede observar la diferencia entre la distribución de los datos de la transformación de la variable de salida (análisis/día) en el sistema real y en la simulación. Para el número de análisis diarios, los cuartiles representados de los datos reales son: $Q1=43.0$, $Q2=46.5$, $Q3=49.7$ y $Q4=55.1$, mientras que en los datos obtenidos a través de la simulación son: $Q1=43.9$, $Q2=46.5$, $Q3=52.0$ y $Q4=59$. En esta representación, se puede observar que la dispersión de los datos reales y de la simulación es asimétrica positiva, es decir, que los datos se concentran mayormente por debajo del valor de la mediana ($Q2$). Por último, si comparamos sus variabilidades, se observa que los datos obtenidos a través de la simulación se tiene una menor variabilidad (o rango intercuartil) de 9 análisis diarios comparado con los datos reales (5.97 análisis diarios).

3.7.2. Prueba Chi-cuadrado de Bondad de Ajuste

Finalmente, se procede a realizar una prueba Chi-cuadrado de bondad de ajuste para determinar si el modelo propuesto

describe lo suficientemente bien mediante una comparación entre las frecuencias esperadas y las frecuencias observadas en la simulación:

Categoría	Frecuencia Observada (Sim.)	Proporción Real	Frecuencia Esperada
8	3	0,05	4,4
9	20	0,25	22,0
10	31	0,30	26,4
11	24	0,30	26,4
12	7	0,05	4,4
13	3	0,05	4,4

Tabla 15. Tabla de Frecuencias Observadas en la Simulación comparados con las Proporciones del Sistema Real

La frecuencia esperada se obtuvo del producto del número total de corridas (88) y la proporción de los datos en el sistema real. La prueba Chi-cuadrado de bondad de ajuste se describe a continuación, además se genera la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : El modelo propuesto se ajusta al sistema real

H_1 : El modelo propuesto no se ajusta al sistema real

N	GL	Chi-cuad.	Valor p
88	5	3,62879	0,604

Tabla 16. Prueba Chi-cuadrado de bondad de ajuste

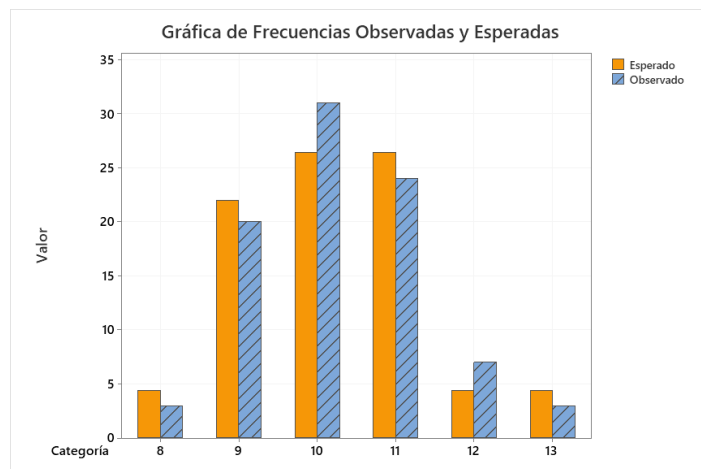


Figura 25. Gráfica de Frecuencias Observadas y Esperadas

Según los valores obtenidos en la Tabla 18, el valor-p es mayor a 0.05, no se tiene evidencia estadística para rechazar H_0 y, por lo tanto, podemos indicar que el modelo propuesto de la situación inicial del laboratorio se ajusta al sistema real.

CAPÍTULO 4

4. RESULTADOS

4.1. Análisis de Causa de los Retrasos

Dado al bajo número de análisis realizados durante varias horas del día, se tienen datos históricos de las causas de los retrasos en la entrega de los resultados, por lo cual se elaboró un diagrama de Ishikawa para encontrar las respectivas causas y posteriormente un diagrama de Pareto con las frecuencias de ocurrencia de los inconvenientes encontrados.

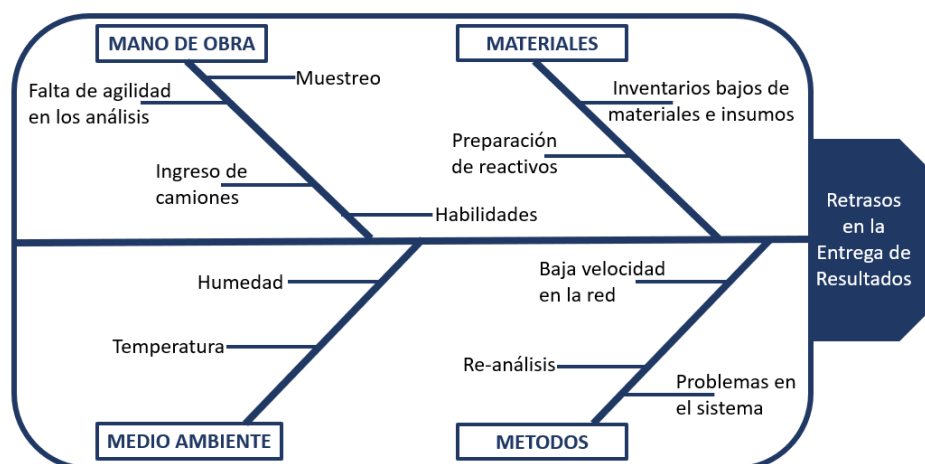


Figura 26. Diagrama de Causa-efecto en retraso de la entrega de resultados

No.	Causas	Frecuencia	Frec. Acum.	%	% Acum.
1	Baja velocidad en la red	106	106	75,71%	75,71%
3	Problemas en el sistema	13	119	9,29%	85,00%
4	Ingreso de camiones	8	127	5,71%	90,71%
5	Inventario de materiales	6	133	4,29%	95,00%
6	Preparación de reactivos	4	137	2,86%	97,86%
7	Re-análisis	2	139	1,43%	99,29%
8	Temperatura	1	140	0,71%	100,00%
Total		140		100,00%	

Tabla 17. Tabla de frecuencia de los problemas encontrados durante Septiembre-2019

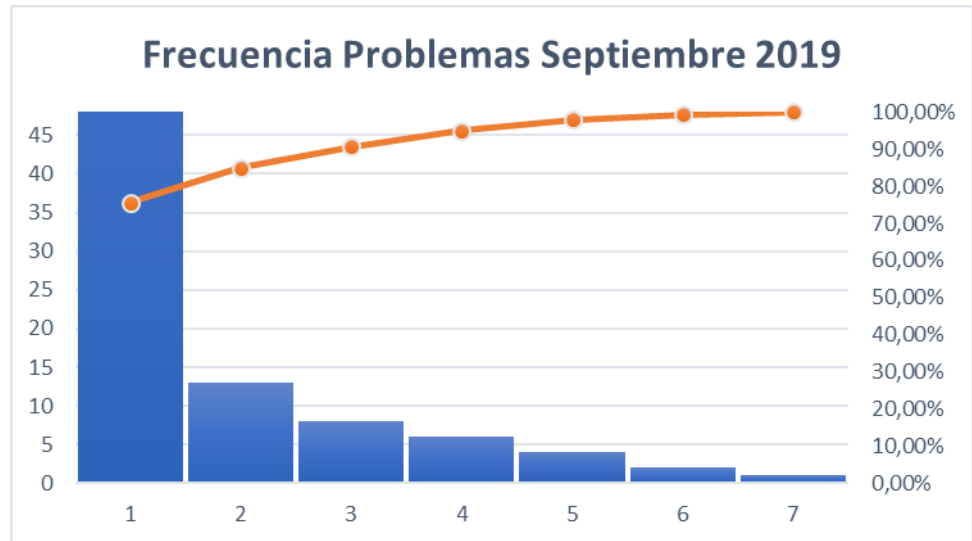


Figura 27. Diagrama de Pareto de las frecuencias de ocurrencia de inconvenientes Septiembre-2019

Dado al porcentaje de las frecuencias, tomamos en cuenta la regla de 80-20 o “pocos vitales y muchos triviales” por lo cual, identificamos la principal causa del problema que es el siguiente:

- Baja velocidad en la red: Esta causa ocasiona mayormente los retrasos durante el ingreso de los datos, análisis de la muestra, almacenamiento de los resultados en la base de datos y el envío de los resultados mediante un correo electrónico. Dado que la señal de la red es compartida por varios departamentos a la vez, ocasiona que esta sea lenta al momento de realizar los análisis.

Por este motivo, se procede a realizar la simulación en tres escenarios: probable, optimista y pesimista en las cuales, mediante datos históricos cambiamos los valores de las variables del tiempo de ingreso, tiempo de análisis, tiempo de almacenamiento en el sistema y tiempo de envío de correo (T_{in} , T_{an} , T_{al} y T_{en} respectivamente). Estos tiempos están directamente relacionados con el problema de la velocidad en la red utilizada en el laboratorio.

Los valores de la variable tiempo de preparación (T_{pr}) sigue con las mismas características dado a que este no puede ser cambiado bajo ningún concepto.

4.2. Aplicación del Método Monte Carlo en la resolución del Problema (Escenario Probable)

Mediante datos históricos, se observa que generalmente en el tiempo por análisis puede existir una disminución de hasta un 18% cuando la red trabaja a velocidad “normal” y no existe mucha demanda de usuarios utilizando la misma, por lo cual se podría incrementar el número de análisis diarios hasta un 23%.

Utilizando el modelo propuesto, se disminuyen los valores de las variables T_{in} , T_{an} , T_{al} y T_{en} en un 25% y se obtienen los siguientes resultados de la variable de salida:

Sim.	Minutos/ análisis	No.Análisis/ día	Sim.	Minutos/ análisis	No.Análisis/ día
1	9,53	50,34	45	8,2	58,52
2	7,77	61,78	46	8,36	57,44
3	7,93	60,56	47	8,61	55,74
4	8,13	59,05	48	8,04	59,7
5	8,71	55,11	49	8,26	58,13
6	7,06	68,03	50	9,38	51,18
7	9,18	52,29	51	9,24	51,92
8	7,96	60,33	52	7,97	60,26
9	8,75	54,88	53	8,75	54,88
10	8,83	54,35	54	8,06	59,54
11	7,29	65,85	55	8,93	53,73
12	7,06	68,02	56	9,94	48,31
13	8,07	59,5	57	8,98	53,46
14	9,29	51,69	58	7,79	61,63
15	7,44	64,49	59	8,92	53,82
16	9,24	51,93	60	6,91	69,46
17	7,61	63,07	61	8,02	59,88
18	8,62	55,65	62	6,68	71,87
19	7,45	64,43	63	7,15	67,11

20	8,02	59,83	64	7,94	60,43
21	8,91	53,89	65	8,43	56,91
22	8,2	58,53	66	8,74	54,92
23	8,86	54,16	67	8,08	59,41
24	7,81	61,49	68	7,43	64,64
25	7,5	63,97	69	7,67	62,58
26	6,81	70,49	70	7,83	61,32
27	8,05	59,6	71	7,25	66,19
28	7,48	64,18	72	9,84	48,76
29	6,69	71,74	73	8,66	55,45
30	7,33	65,45	74	9,29	51,68
31	8,17	58,73	75	8,68	55,30
32	9,20	52,19	76	7,58	63,33
33	8,03	59,75	77	7,00	68,58
34	7,02	68,37	78	8,04	59,71
35	8,94	53,69	79	8,66	55,46
36	9,12	52,62	80	10,81	44,39
37	9,32	51,5	81	8,78	54,66
38	7,27	66,05	82	8,87	54,14
39	7,23	66,37	83	7,97	60,22
40	8,95	53,63	84	9,15	52,48
41	9,32	51,51	85	7,99	60,07
42	7,33	65,51	86	6,92	69,37
43	7,81	61,45	87	7,68	62,48
44	8,94	53,7	88	7,90	60,75

Tabla 18. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo en el Escenario Probable

Procedemos entonces a comprobar si los datos simulados de las variables T_{in} , T_{an} , T_{al} y T_{en} siguen las mismas distribuciones que los datos observados al inicio:

Variable T_{in} : En los datos del sistema real, estos siguen una distribución empírica dado a que no se ajusta a ninguna distribución conocida. Se realiza un histograma de los datos obtenidos a través de la simulación:

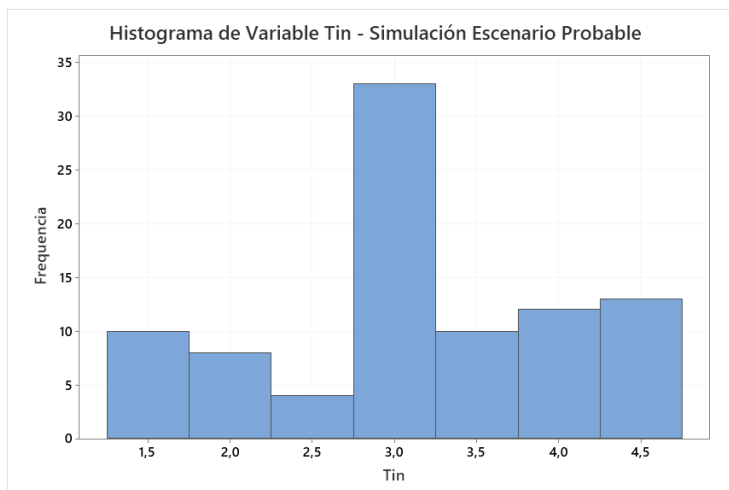


Figura 28. Histograma de la Variable T_{in} – Simulación escenario probable

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta a la distribución de los datos de la variable T_{in} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	<0,005
Exponencial	<0,003
Weibull	<0,010
Gamma	<0,005

Tabla 19. Valor-p de la variable tiempo de ingreso – Simulación escenario probable

Por los valores presentados en la Tabla 21, se puede observar que al igual que en los datos iniciales, los datos de la variable T_{in} obtenidos en la simulación en el escenario probable no se ajustan a ninguna distribución conocida debido a que el valor-p obtenido es <0.05 y, por lo tanto, se tiene evidencia estadística para rechazar H_0 . Entonces la distribución empírica de estos datos en la simulación en un escenario optimista es la siguiente:

$$f(x) = \begin{cases} 0.15; 0.0 \leq x < 1.5 \\ 0.10; 1.5 \leq x < 2.0 \\ 0.05; 2.0 \leq x < 2.5 \\ 0.35; 2.5 \leq x < 3.0 \\ 0.10; 3.0 \leq x < 3.5 \\ 0.15; 3.5 \leq x < 4.0 \\ 0.10; 4.0 \leq x < 4.5 \end{cases}$$

Variable T_{an} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

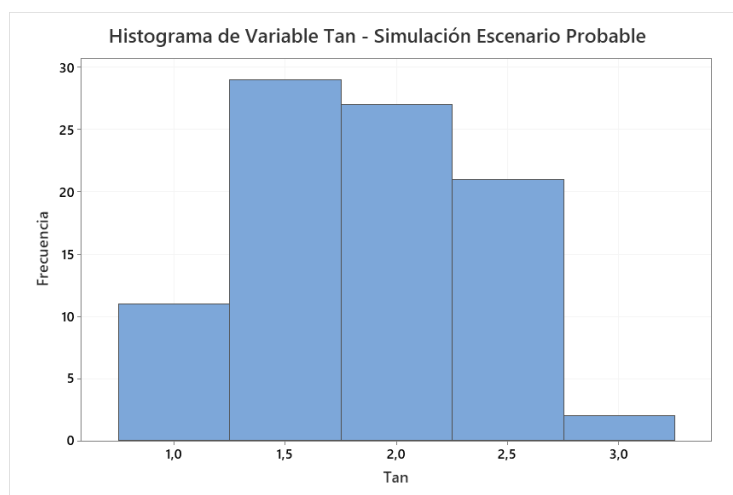


Figura 29. Histograma de la Variable T_{an} – Simulación escenario probable

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{an} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,171
Exponencial	<0,003
Weibull	0,012
Gamma	0,074

Tabla 20. Valor-p de la variable tiempo de análisis – Simulación escenario probable

Según los valores presentados en la Tabla 22, se puede observar que el valor-p obtenido en la Distribución Normal es el mayor y, por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{an} obtenidos en la simulación en el escenario probable se ajustan a una Distribución Normal, con una media de 1.9 y 0.5 de desviación estándar $N(1.9,0.5)$.

Variable T_{al} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

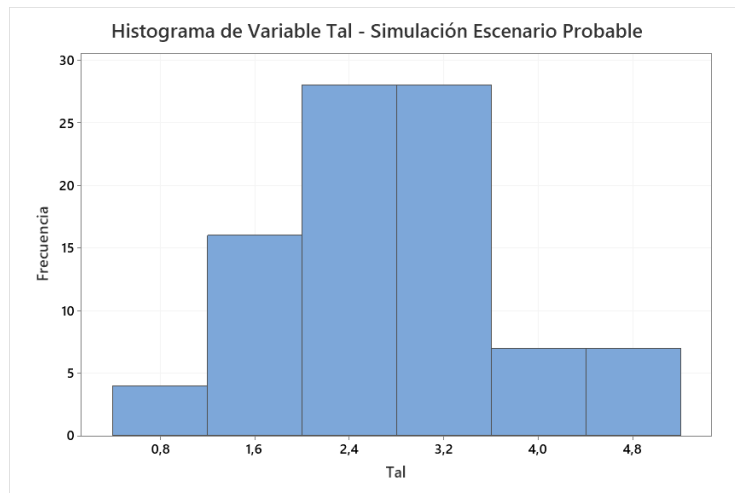


Figura 30. Histograma de la Variable T_{al} – Simulación escenario probable

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{al} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,460
Exponencial	<0,003
Weibull	0,251
Gamma	0,150

Tabla 21. Valor-p de la variable tiempo de almacenamiento de los datos en el sistema – Simulación escenario probable

Según los valores presentados en la Tabla 23, se puede observar que el valor-p obtenido en la Distribución Normal es el mayor y, por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{al} obtenidos en la simulación en el escenario probable se ajustan a una Distribución Normal, con una media de 2.7 y 0.9 de desviación estándar $N(2.7,0.9)$.

Variable T_{en} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

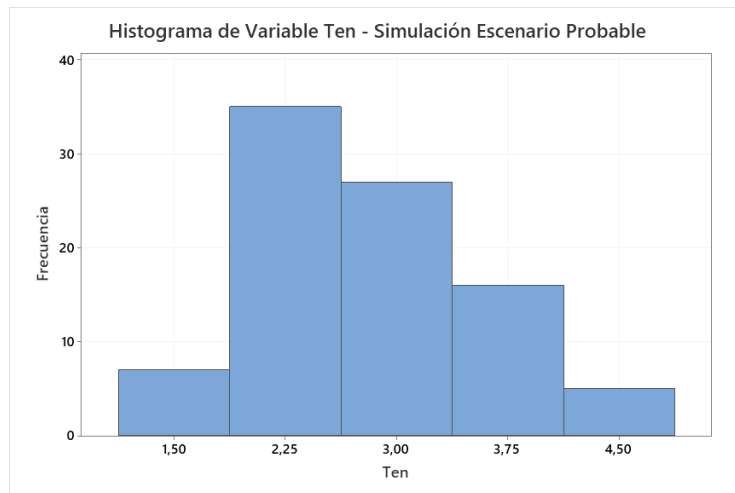


Figura 31. Histograma de la Variable T_{en} – Simulación escenario probable

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{en} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,054
Exponencial	<0,003
Weibull	0,019
Gamma	0,025

Tabla 22. Valor-p de la variable tiempo de envío – Simulación escenario probable

Según los valores presentados en la Tabla 24, se puede observar que el valor-p obtenido en la Distribución Normal es ≥ 0.05 , por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{en} obtenidos en la simulación en el escenario probable se ajustan a una Distribución Normal, con una media de 2.9 y 0.7 de desviación estándar $N(2.9,0.7)$.

4.3. Aplicación del Método Monte Carlo en la resolución del Problema (Escenario Optimista)

Mediante datos históricos, se observa que ciertos datos del tiempo por análisis pueden reducirse hasta un 36% cuando la red trabaja a velocidad “alta” y no existe demanda de otros usuarios utilizando la misma, por lo cual se podría incrementar el número de análisis diarios hasta un 56%.

Utilizando el modelo propuesto, se disminuyen los valores de las variables T_{in} , T_{an} , T_{al} y T_{en} en un 50% y se obtienen los siguientes resultados de la variable de salida:

Sim.	Minutos/ análisis	No.Análisis/ día	Sim.	Minutos/ análisis	No.Análisis/ día
1	5,90	81,37	45	6,77	70,94
2	6,95	69,09	46	6,47	74,16
3	7,31	65,69	47	5,47	87,71
4	6,33	75,84	48	7,17	66,92
5	6,16	77,93	49	6,43	74,63
6	5,66	84,78	50	6,32	75,90
7	6,58	72,92	51	6,91	69,49
8	6,56	73,17	52	6,68	71,88
9	6,46	74,27	53	6,09	78,81
10	6,68	71,86	54	6,64	72,31
11	5,58	86,07	55	5,89	81,51
12	6,61	72,64	56	7,23	66,41
13	7,25	66,18	57	5,97	80,41
14	7,30	65,77	58	6,16	77,86
15	7,01	68,44	59	5,83	82,33
16	7,03	68,27	60	6,80	70,56

17	6,38	75,27	61	5,77	83,12
18	6,18	77,64	62	6,89	69,66
19	6,64	72,27	63	6,84	70,19
20	6,17	77,79	64	6,34	75,70
21	5,75	83,45	65	7,16	67,08
22	5,66	84,75	66	7,00	68,60
23	6,48	74,05	67	7,30	65,73
24	5,80	82,77	68	5,96	80,49
25	6,75	71,15	69	6,82	70,36
26	6,87	69,88	70	5,73	83,77
27	7,75	61,93	71	5,63	85,24
28	5,87	81,77	72	6,90	69,52
29	5,24	91,59	73	7,61	63,09
30	5,91	81,25	74	6,23	77,04
31	6,55	73,33	75	6,46	74,35
32	7,40	64,84	76	6,14	78,21
33	5,89	81,56	77	6,24	76,97
34	6,16	77,96	78	6,99	68,66
35	5,67	84,61	79	6,64	72,24
36	6,62	72,51	80	6,23	77,00
37	6,30	76,24	81	6,74	71,24
38	6,09	78,78	82	6,87	69,87
39	7,30	65,75	83	6,06	79,21
40	6,35	75,60	84	6,18	77,62
41	6,41	74,85	85	6,05	79,31
42	6,34	75,69	86	6,40	74,98
43	5,31	90,35	87	7,54	63,65
44	7,68	62,46	88	6,69	71,70

Tabla 23. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo en el Escenario Optimista

Procedemos entonces a comprobar si los datos simulados de las variables T_{in} , T_{an} , T_{al} y T_{en} siguen las mismas distribuciones que los datos observados:

Variable T_{in} : En los datos del sistema real, estos siguen una distribución empírica dado a que no se ajusta a ninguna distribución conocida. Se realiza un histograma de los datos obtenidos a través de la simulación:

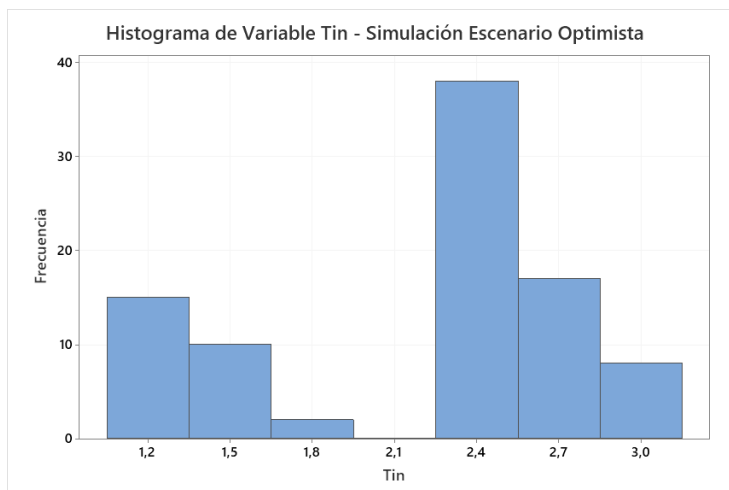


Figura 32. Histograma de la Variable T_{in} – Simulación escenario optimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{in} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	<0,005
Exponencial	<0,003
Weibull	<0,010
Gamma	<0,005

Tabla 24. Valor-p de la variable tiempo de ingreso – Simulación escenario optimista

Por los valores presentados en la Tabla 26, se puede observar que al igual que en los datos iniciales, los datos de la variable T_{in} obtenidos en la simulación en el escenario optimista no se ajustan a ninguna distribución conocida debido a que el valor-p obtenido es <0.05 y, por lo tanto, se tiene evidencia estadística para rechazar H_0 . Entonces la distribución empírica de estos datos en la simulación bajo un escenario optimista es la siguiente:

$$f(x) = \begin{cases} 0.17; & 0.00 \leq x < 1.35 \\ 0.11; & 1.35 \leq x < 1.65 \\ 0.03; & 1.65 \leq x < 1.95 \\ 0.00; & 1.95 \leq x < 2.25 \\ 0.41; & 2.25 \leq x < 2.55 \\ 0.19; & 2.55 \leq x < 2.85 \\ 0.09; & 2.85 \leq x < 3.15 \end{cases}$$

Variable T_{an} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

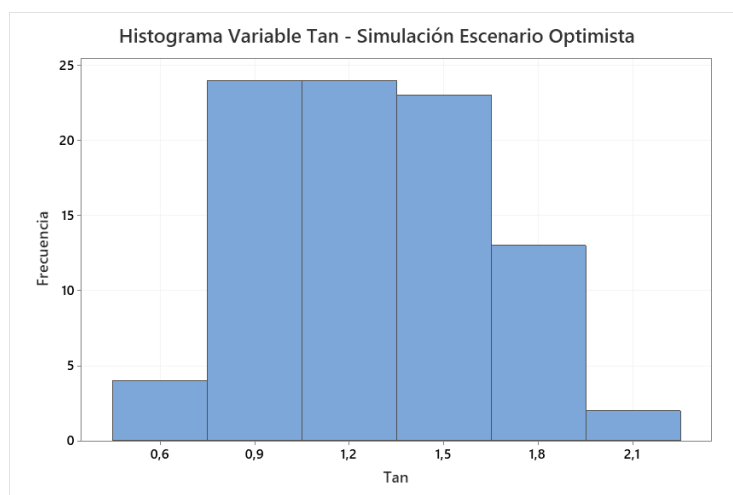


Figura 33. Histograma de la Variable T_{an} – Simulación escenario optimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{an} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,105
Exponencial	<0,003
Weibull	0,055
Gamma	0,010

Tabla 25. Valor-p de la variable tiempo de análisis – Simulación escenario optimista

Según los valores presentados en la Tabla 27, se puede observar que el valor-p obtenido en la Distribución Normal es el mayor, por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{an} obtenidos en la simulación en el escenario optimista se ajustan a una Distribución Normal, con una media de 1.3 y 0.4 de desviación estándar $N(1.3,0.4)$.

Variable T_{al} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

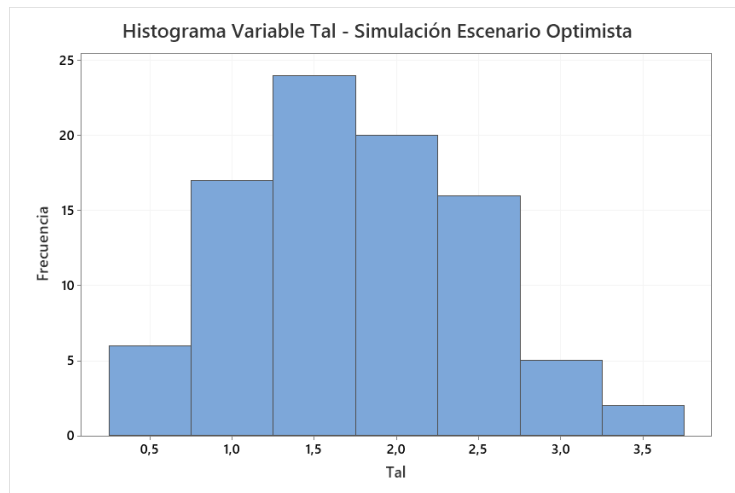


Figura 34. Histograma de la Variable T_{al} – Simulación escenario optimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{al} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,359
Exponencial	<0,003
Weibull	0,250
Gamma	0,180

Tabla 26. Valor-p de la variable tiempo de almacenamiento de los datos en el sistema – Simulación escenario optimista

Según los valores presentados en la Tabla 28, se puede observar que el valor-p obtenido en la Distribución Normal es el mayor, por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{al} obtenidos en la simulación en el escenario optimista se ajustan a una Distribución Normal, con una media de 1.8 y 0.7 de desviación estándar $N(1.8,0.7)$.

Variable T_{en} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

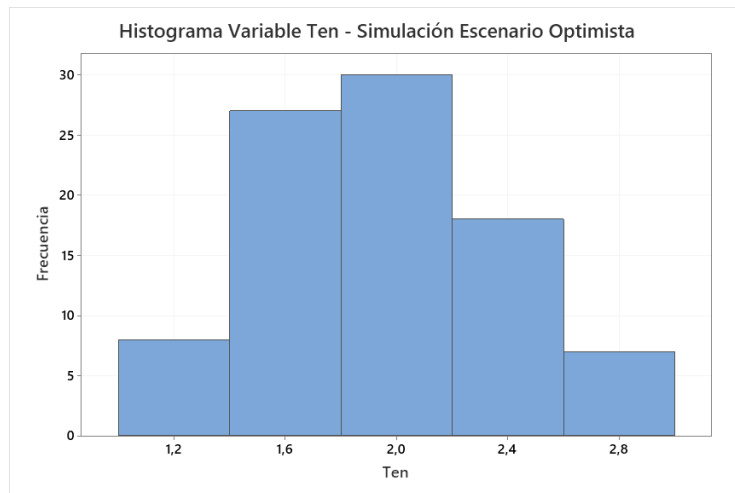


Figura 35. Histograma de la Variable T_{en} – Simulación escenario optimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{en} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,171
Exponencial	<0,003
Weibull	0,013
Gamma	0,025

Tabla 27. Valor-p de la variable tiempo de envío – Simulación escenario optimista

Según los valores presentados en la Tabla 29, se puede observar que el valor-p obtenido en la Distribución Normal es ≥ 0.05 , por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{en} obtenidos en la simulación en el escenario optimista se ajustan a una Distribución Normal, con una media de 1.9 y 0.4 de desviación estándar $N(1.9,0.4)$.

4.4. Aplicación del Método Monte Carlo en la resolución del Problema (Escenario Pesimista)

Mediante datos históricos, se observa que ciertos datos de la variable tiempo por análisis puede existir un incremento de hasta un 15% cuando la red trabaja a velocidad “lenta” y existe mucha demanda de usuarios utilizando la misma, por lo cual el número de análisis diarios puede disminuir hasta un 15%.

Utilizando el modelo propuesto, se incrementan los valores de las variables T_{in} , T_{an} , T_{al} y T_{en} en un 20% y se obtienen los siguientes resultados de la variable de salida:

Sim.	Minutos/ análisis	No.Análisis/ día	Sim.	Minutos/ análisis	No.Análisis/ día
1	14,15	33,93	45	11,91	40,29
2	12,94	37,08	46	13,03	36,85
3	13,82	34,72	47	11,39	42,15
4	12,20	39,33	48	11,01	43,61
5	7,77	61,80	49	10,84	44,28
6	12,48	38,47	50	11,97	40,08
7	12,48	38,46	51	9,78	49,06
8	9,20	52,16	52	12,75	37,66
9	13,17	36,44	53	11,44	41,96
10	12,58	38,15	54	11,12	43,15
11	9,68	49,57	55	11,04	43,50
12	12,77	37,60	56	12,34	38,91
13	13,41	35,79	57	11,74	40,88
14	13,12	36,60	58	10,29	46,65
15	13,28	36,16	59	13,61	35,27
16	12,56	38,21	60	12,36	38,84
17	10,32	46,50	61	12,59	38,14
18	12,98	36,98	62	10,26	46,80

19	12,34	38,89	63	12,37	38,82
20	10,63	45,14	64	12,24	39,20
21	10,76	44,61	65	13,44	35,71
22	12,58	38,16	66	10,97	43,74
23	11,91	40,32	67	10,38	46,23
24	13,83	34,71	68	11,73	40,91
25	13,42	35,76	69	10,74	44,69
26	10,36	46,32	70	11,80	40,69
27	13,58	35,34	71	10,90	44,04
28	13,11	36,62	72	11,89	40,37
29	11,41	42,07	73	11,29	42,53
30	13,72	34,99	74	11,46	41,87
31	11,05	43,44	75	13,34	35,98
32	13,73	34,96	76	9,82	48,88
33	11,83	40,58	77	13,60	35,29
34	13,89	34,55	78	11,25	42,68
35	13,08	36,71	79	10,70	44,85
36	10,73	44,72	80	11,56	41,53
37	11,88	40,40	81	11,75	40,84
38	8,39	57,18	82	10,96	43,80
39	12,36	38,85	83	12,74	37,68
40	9,50	50,55	84	9,82	48,88
41	11,47	41,84	85	9,43	50,88
42	11,58	41,45	86	11,06	43,39
43	10,46	45,90	87	11,26	42,64
44	11,46	41,90	88	13,10	36,65

Tabla 28. Valores de minutos por análisis y de análisis diarios obtenidos de la simulación de Monte Carlo en el Escenario Pesimista

Procedemos entonces a comprobar si los datos simulados de las variables T_{in} , T_{an} , T_{al} y T_{en} siguen las mismas distribuciones que los datos originales:

Variable T_{in} : En los datos del sistema real, estos siguen una distribución empírica dado a que no se ajusta a ninguna distribución conocida. Se realiza un histograma de los datos obtenidos a través de la simulación:

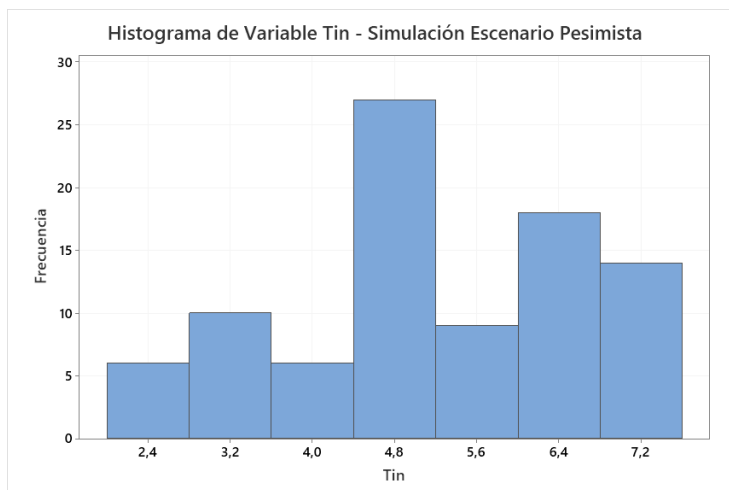


Figura 36. Histograma de la Variable T_{in} – Simulación escenario pesimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{in} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	<0,005
Exponencial	<0,003
Weibull	<0,010
Gamma	<0,005

Tabla 29. Valor-p de la variable tiempo de ingreso – Simulación escenario pesimista

Por los valores presentados en la Tabla 31, se puede observar que al igual que en los datos iniciales, los datos de la variable T_{in} obtenidos en la simulación en el escenario pesimista no se ajustan a ninguna distribución conocida debido a que el valor-p obtenido es <0.05 y por lo tanto, se tiene evidencia estadística para rechazar H_0 . Entonces la distribución empírica de los datos en la simulación bajo un escenario pesimista es la siguiente:

$$f(x) = \begin{cases} 0.07; 0.0 \leq x < 2.8 \\ 0.11; 2.8 \leq x < 3.6 \\ 0.07; 3.6 \leq x < 4.4 \\ 0.30; 4.4 \leq x < 5.2 \\ 0.10; 5.2 \leq x < 6.0 \\ 0.19; 6.0 \leq x < 6.8 \\ 0.16; 6.8 \leq x < 7.6 \end{cases}$$

Variable T_{an} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

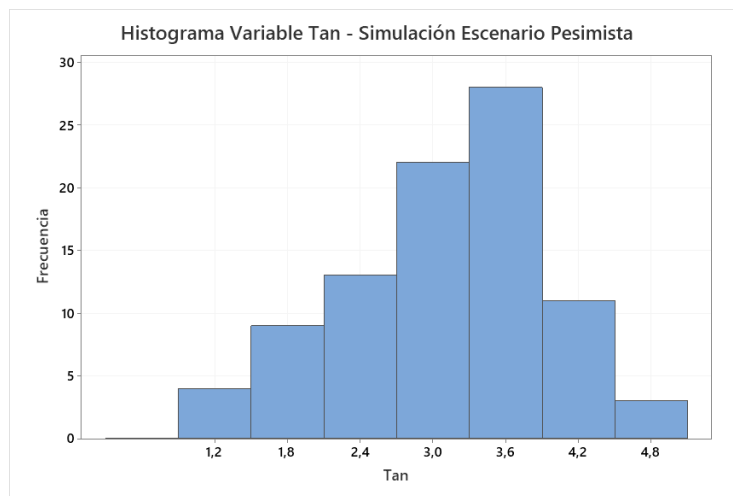


Figura 37. Histograma de la Variable T_{an} – Simulación escenario pesimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{an} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,055
Exponencial	<0,003
Weibull	<0,010
Gamma	<0,005

Tabla 30. Valor-p de la variable tiempo de análisis – Simulación escenario pesimista

Según los valores presentados en la Tabla 32, se puede observar que el valor-p obtenido en la Distribución Normal es ≥ 0.05 , por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{en} obtenidos en la simulación en el escenario pesimista se ajustan a una Distribución Normal, con una media de 3.1 y 0.9 de desviación estándar $N(3.1,0.9)$.

Variable T_{al} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

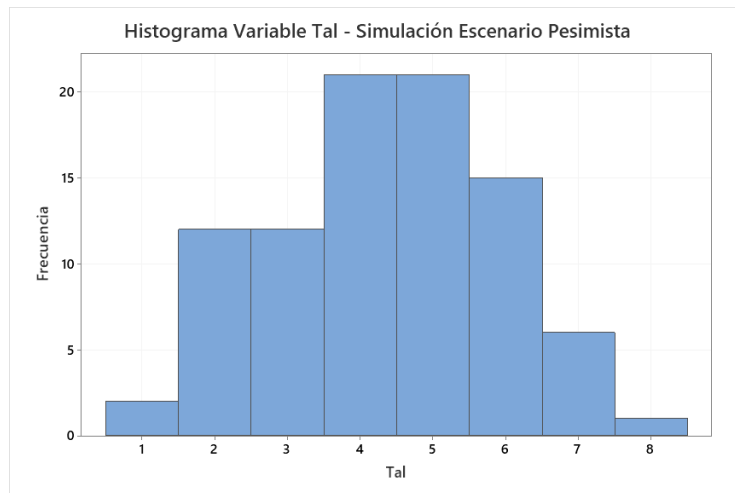


Figura 38. Histograma de la Variable T_{al} – Simulación escenario pesimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{al} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,183
Exponencial	<0,003
Weibull	0,137
Gamma	<0,005

Tabla 31. Valor-p de la variable tiempo de almacenamiento de los datos en el sistema – Simulación escenario pesimista

Según los valores presentados en la Tabla 33, se puede observar que el valor-p obtenido en la Distribución Normal es el mayor, por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{al} obtenidos en la simulación en el escenario pesimista se ajustan a una Distribución Normal, con una media de 4.4 y 1.5 de desviación estándar $N(4.4, 1.5)$.

Variable T_{en} : En los datos del sistema real, estos siguen una distribución normal. Se realiza un histograma de los datos obtenidos a través de la simulación:

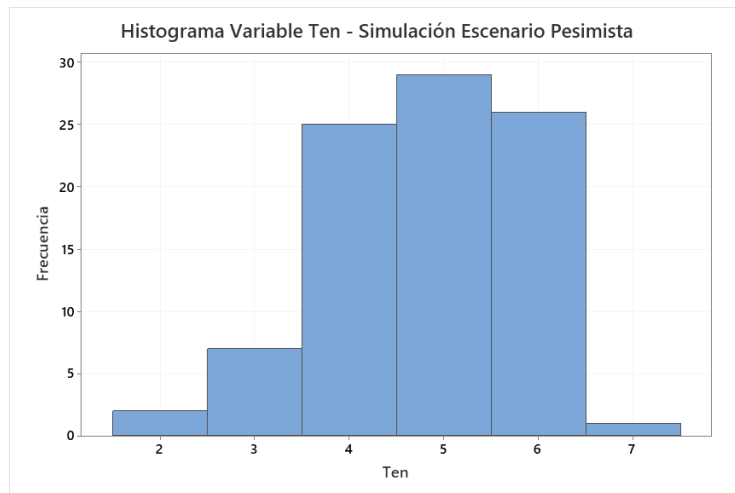


Figura 39. Histograma de la Variable T_{en} – Simulación escenario pesimista

Posterior a esto, se probaron 4 distribuciones para determinar cuál de ellos se ajusta al comportamiento de los datos de la variable T_{en} mediante la siguiente prueba de hipótesis con 0.05 de significancia:

H_0 : Los datos siguen una Distribución X.

H_1 : Los datos no siguen una Distribución X.

Distribución	Valor-p
Normal	0,183
Exponencial	<0,003
Weibull	0,137
Gamma	<0,005

Tabla 32. Valor-p de la variable tiempo de envío – Simulación escenario pesimista

Según los valores presentados en la Tabla 34, se puede observar que el valor-p obtenido en la Distribución Normal es el mayor, por lo tanto, no se tiene evidencia estadística para rechazar H_0 . Entonces los datos de la variable T_{en} obtenidos en la simulación en el escenario pesimista se ajustan a una Distribución Normal, con una media de 4.8 y 1.0 de desviación estándar $N(4.8,1.0)$.

4.5. Diferencia entre Escenarios de Simulación

En la siguiente Tabla se observarán los valores de los diferentes escenarios como resumen:

Estadístico	Escenario Probable		Escenario Optimista		Escenario Pesimista	
	Minutos/ análisis	Análisis/ día	Minutos/ análisis	Análisis/ día	Minutos/ análisis	Análisis/ día
Media	8,41	59	6,61	75	11,80	41
Incremento	-18%	23%	-36%	56%	15%	-15%

Tabla 33. Resumen de las Simulaciones en diferentes Escenarios

En la Tabla 35 se muestran las medias y cómo estas aumentan o disminuyen según su comportamiento durante la ejecución de las simulaciones:

- Escenario Probable: Los minutos por análisis disminuyen en un 18% mientras que los análisis por día se incrementan en un 23%.
- Escenario Optimista: Los minutos por análisis disminuyen en un 36% mientras que los análisis por día se incrementan en un 56%.
- Escenario Pesimista: Los minutos por análisis se incrementan en un 15% mientras que los análisis por día disminuyen en un 15%.

Según los valores obtenidos, el escenario optimista es el que tiene mayor incremento de la productividad puesto a que se incrementa hasta un 56% dado a que los tiempos se disminuyen en un 36%

CAPÍTULO 5

5. CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

- Si se considera solucionar la causa principal de los retrasos en la entrega de los resultados en el laboratorio (baja velocidad en la red), la productividad (en términos de análisis diarios) puede aumentar hasta en un 56%, con esto se evitan costos adicionales tales como horas extras pagadas a los analistas de laboratorio.
- Tanto los datos del sistema real, como en las simulaciones en cada escenario, están basados en datos históricos. En cada una de las variables de entrada y de salida se comprobó que sigan el mismo comportamiento de los datos del sistema real.
- En el presente estudio, se diseñó la simulación de Monte Carlo aplicando cada una de las etapas de la misma para que esta describa lo suficiente al sistema real de los datos tomados.
- La productividad (en términos de análisis diarios) se puede incrementar entre el 22 a 56%. Esto si consideramos el escenario probable o el escenario optimista, el cual es posible siempre y cuando no se comparta la red con otros departamentos para que la velocidad de la red no decaiga. Otra posibilidad puede ser tener una sola red para el laboratorio y que esta no sea compartida.

5.2. Recomendaciones

- Se podría definir si bajo el escenario optimista se puede llegar a incrementar aún más la productividad al considerar solucionar otras causas de retrasos en la entrega de resultados.
- Se podría utilizar el método de Monte Carlo para otros proyectos dentro del laboratorio, como por ejemplo disminuir la variabilidad de los resultados de los análisis, validación de métodos de laboratorio, entre otros.

6. REFERENCIAS

1. 18, S. d. (2020). *Formas de evaluar una distribución de datos*. Obtenido de Soporte de Minitab (R) 18: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/probability-distributions-and-random-data/supporting-topics/basics/ways-to-assess-a-data-distribution/>
2. 18, S. d. (2020). *Interpretar los resultados clave para la Identificación de distribución individual*. Obtenido de Soporte de Minitab (R) 18: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/quality-and-process-improvement/quality-tools/how-to/individual-distribution-identification/interpret-the-results/key-results/>
3. Amat, J. (2020). *Introducción a la Regresión Lineal Múltiple*. Obtenido de RPubS - Regresión Lineal Múltiple en R: https://rpubs.com/Joaquin_AR/226291
4. Ambientales, C. (2019). *Práctica 5: Determinación el Contenido Graso de Leche en Polvo: Extracción Soxhlet*. Obtenido de Técnicas Avanzadas en Química: https://www.upo.es/depa/webdex/quimfis/docencia/TAQ/curso0405/TAQP5_0405.pdf
5. Belda, C., & Grande, E. (2009). Los Modelos de Simulación: Una Herramienta Multidisciplinar de Investigación. *Encuentros Multidisciplinares*, 1-11.
6. Benítez-Pérez, H., & Cárdenas-Flores, F. y.-N. (2009). Red reconfigurable mediante el modelo de control predictivo para tres bandas transportadoras como caso de estudio. . *Información Tecnológica, Vol.20, No.1*, 111-127.

7. Calderón, E. (2010). Aplicaciones de la Simulación Monte Carlo en Dosimetría y Problemas de Física Médica. *ResearchGate*, 1-9.
8. Camo, A. (2019). *Chemometrics*. Obtenido de Understand chemical data: <https://www.camo.com/chemometrics/>
9. Canavos, G. (2003). Probabilidad y Estadística. Aplicaciones y Métodos. Ciudad de México: McGraw-Hill.
10. Cárdenas, N. (2010). Optimización del Servicio ofrecido por Multiservidores basado en Simulación Matemática para un almacén de Electrodomésticos en la Ciudad de Guayaquil. Obtenido de DSpace en Espol: <https://www.dspace.espol.edu.ec/browse?type=author&value=C%C3%A1rdenas+Escobar%2C+Nadia+Lorena>
11. Carollo, C. (2020). *Regresión Lineal Simple*. Obtenido de Universidad de Santiago de Compostella: http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat_50140116_Regr_%20simple_2011_12.pdf
12. CNA. (2019). *Cámara Nacional de Acuacultura - Estadísticas*. Obtenido de Cámara Nacional de Acuacultura: <http://www.cna-ecuador.com/estadisticas/>
13. Cochran, W. (1954). The Combination of Estimates from Different Experiment, *Biometrics. International Biometric Society*, 101-129.
14. Comparación de la estimación de la productividad del proceso de soldadura eléctrica por los métodos L. F. Martins, M. L. (2012). Comparación de la estimación de la productividad del proceso de soldadura eléctrica por los métodos de simulación de Monte Carlo e Hipercubo Latino. *Información Tecnológica Vol. 23(4)*, 21-32.

15. CRISP. (2019). *Electromagnetic Radiation*. Obtenido de Centre for Remote Imaging, Sensing & Processing: <https://crisp.nus.edu.sg/~research/tutorial/em.htm>
16. Devore, J. L. (2008). *Probabilidad y Estadística para Ingeniería y Ciencias*. Ciudad de México: International Thomson Editores.
17. Esbensen, K., & Geladi, P. (2019). *Wikipedia*. Obtenido de Chemometrics: <https://en.wikipedia.org/wiki/Chemometrics>
18. Flanagan, R., Norman, G., & Chapman, R. (2006). *Risk Management And Construction*. London: John Wiley & Sons Incorporated.
19. García, F., Sierra, J., & Guzmán, V. (2005). *Cálculo de Número Óptimo de Simulaciones*. Obtenido de Espacio de Formación Multimodal - Uaem.mx: http://metabase.uaem.mx/bitstream/handle/123456789/564/Calculo_de_numero_optimo_de_simulaciones.pdf?sequence=1&isAllowed=y
20. García, J., Reding, A., & López, J. (2013). Cálculo del tamaño de la muestra en investigación en educación médica. *ScienceDirect*, 217-224.
21. Kofman, E. (2020). *Simulación de Sistemas Continuos - Notas de Clase*. Obtenido de Facultad de Ciencias Exactas, Ingeniería y Agrimensura - Universidad Nacional de Rosario: https://www.fceia.unr.edu.ar/control/ssc/notas_ssc.pdf
22. Law, A. (2014). Simulation, Modeling and Analysis. En A. Law, *Simulation, Modeling and Analysis* (págs. 722-723). Tucson: McGraw-Hill International Editions.

23. López de Bermúdez, N. Y. (2016). Aspectos teóricos en la generación de variables aleatorias. *Facultad de Ciencias Naturales y Matemática.*, 28-34.
24. López, J. (2008). *Guía Básica para la Simulación de Monte Carlo*. Madrid: AENOR. Asociación Española de Normalización y Certificación.
25. López, M. (2020). *Pruebas de Normalidad: Prueba de Anderson-Darling*. Obtenido de es.slideshare.net: <https://es.slideshare.net/leugimxw/pruebas-de-normalidad-prueba-de-andersondarling>
26. McMurray, A., Pearson, T., & Casarim, F. (2017). Guía para aplicar el Enfoque de Monte Carlo al Análisis de Incertidumbre en la Contabilidad Forestal y de Gases de Efecto Invernadero. Obtenido de Winrock International: <https://www.winrock.org/wp-content/uploads/2018/03/UncertaintyReport-12.26.17-ES.pdf>
27. Montenegro, R. (2011). Diseño e implementación de un sistema de inventarios, aplicando simulación Montecarlo, en una empresa de Servicios Petroleros. Obtenido de BIBDIGITAL Escuela Politécnica Nacional: <https://bibdigital.epn.edu.ec/handle/15000/7871>
28. Montesinos, A. (2011). ESTUDIO DEL AIC Y BIC EN LA SELECCIÓN DE MODELOS DE VIDA CON DATOS CENSURADOS. *Centro de Investigación en Matemáticas A.C.*, 15.
29. Morano, C., & Ferreira, M. (2003). Aplicação do Método de Monte Carlo em Análise de Riscos em Projetos de Construção. *XXXV SBPO - Simpósio Brasileiro de Pesquisa Operacional* (pág. 1). Natal: XXXV SBPO - Simpósio Brasileiro de Pesquisa Operacional.

30. Nahum, A. (2013). *Monte Carlo Techniques in Radiation Therapy. Imaging in Medical Diagnosis and Therapy*. Boca Raton, FL: CRC Press.
31. Palisade, C. (2019). *Monte Carlo Simulation*. Obtenido de Palisade: http://www.palisade.com/risk/monte_carlo_simulation.asp
32. Peña, D. (2001). *Deducción de distribuciones: el método de Monte Carlo*. En D. (. Peña Sánchez de Rivera, *Fundamentos de Estadística*. Madrid: Alianza Editorial.
33. Puche, J. (2005). *Guía Práctica para la Simulación de Procesos Industriales*. En J. Puche, *Guía Práctica para la Simulación de Procesos Industriales* (págs. 12-40). Murcia: Centro Tecnológico del Mueble y la Madera de la Región de Murcia.
34. Quevedo, L. (2020). *Estudio sobre la Simulación de Eventos Discretos*. Obtenido de ResearchGate: https://www.researchgate.net/publication/259043035_Estudio_Sobre_la_Simulacion_de_Eventos_Discretos
35. Ramirez-Alan, O. (2020). *Chi-Cuadrado*. Obtenido de RPubs by RStudio: <https://rpubs.com/osoramirez/111403>
36. Rodríguez, I. (2016). *Simulación Análisis Salidas Simulación*. Obtenido de SlideShare: <https://es.slideshare.net/inesarodriguez58/informe-simulacion-analissalidassimulacion>
37. Salazar, E., & Alzate, W. (2018). *Aplicación de la simulación Monte Carlo en la proyección del estado de resultados. Un estudio de caso*. Caracas: Revista ESPACIOS.
38. Santianello, S., & Filippo, F. (2011). *Saber UCV: Aplicación del método de Monte Carlo en exploración y explotación de petróleo y gas natural*. Obtenido de Saber UCV: <http://saber.ucv.ve/handle/123456789/556>

39. Sotou, J. (2005). *Simulación Método Monte Carlo*. Obtenido de Wordpress: <https://jaimesotou.files.wordpress.com/2011/05/metodo-montecarlo-03.pdf>
40. UNAM. (05 de diciembre de 2019). *Fundamentos y Técnicas de Análisis de Alimentos. Laboratorio de Alimentos I Departamento de Alimentos y Biotecnología*. Obtenido de Facultad de Química: http://depa.fquim.unam.mx/amyd/archivero/FUNDAMENTOSYTECNICASDEANALISISDEALIMENTOS_12286.pdf
41. Wikipedia. (03 de diciembre de 2019). *Chemometrics*. Obtenido de Wikipedia: <https://en.wikipedia.org/wiki/Chemometrics>
42. Wikipedia. (03 de diciembre de 2019). *Espectro electromagnético*. Obtenido de Wikipedia: https://es.wikipedia.org/wiki/Espectro_electromagnético