



UNIVERSIDAD NACIONAL DE CHIMBORAZO
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO
DIRECCIÓN DE POSGRADO

“Análisis de la variabilidad y precisión en la estimación de intervalos de confianza en pruebas de hipótesis con desviaciones estándar variables: aplicación a datos de física de partículas del CERN y Zenodo”

Trabajo de Titulación para optar al título de Magíster en Matemática
Aplicada con mención en Matemática Computacional

AUTOR:

Salazar Andrade, Andrés Eduardo

TUTOR:

Haro Rivera Silvia Mariana, PhD.

Riobamba, Ecuador. 2025

Declaración y Cesión de Derechos de Autoría

Yo, **Andrés Eduardo Salazar Andrade**, con número único de identificación 180447111-6, declaro y acepto ser responsable de las ideas, doctrinas, resultados y lineamientos alternativos realizados en el presente trabajo de titulación denominado: **“Análisis de la variabilidad y precisión en la estimación de intervalos de confianza en pruebas de hipótesis con desviaciones estándar variables: aplicación a datos de física de partículas del CERN y Zenodo”** previo a la obtención del grado de Magíster en Matemática Aplicada con mención en Matemática Computacional.

- Declaro que mi trabajo investigativo pertenece al patrimonio de la Universidad Nacional de Chimborazo de conformidad con lo establecido en el artículo 20 literal j) de la Ley Orgánica de Educación Superior LOES.
- Autorizo a la Universidad Nacional de Chimborazo que pueda hacer uso del referido trabajo de titulación y a difundirlo como estime conveniente por cualquier medio conocido, y para que sea integrado en formato digital al Sistema de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor, dando cumplimiento de esta manera a lo estipulado en el artículo 144 de la Ley Orgánica de Educación Superior LOES.

Riobamba, 19 julio de 2025

Ing. Andrés Eduardo Salazar Andrade

N.U.I. 180447111-6



Dirección de
Posgrado
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO



ACTA DE CULMINACIÓN DE TRABAJO DE TITULACIÓN

En la ciudad de Riobamba, a los 21 días del mes de julio del año 2025, los miembros del Tribunal designado por la Comisión de Posgrado de la Universidad Nacional de Chimborazo, reunidos con el propósito de analizar y evaluar el Trabajo de Titulación bajo la modalidad Proyecto de titulación con componente investigación aplicada y/o desarrollo, CERTIFICAMOS lo siguiente:

Que, una vez revisado el trabajo titulado: "**Análisis de la variabilidad y precisión en la estimación de intervalos de confianza en pruebas de hipótesis con desviaciones estándar variables: aplicación a datos de física de partículas del CERN y Zenodo**", perteneciente a la línea de investigación: Ingeniería Informática, presentado por el maestrante **Salazar Andrade Andrés Eduardo**, portador de la cédula de ciudadanía No. **1804471116**, estudiante del programa de **Maestría en Matemática Aplicada con mención en Matemática Computacional**, se ha verificado que dicho trabajo cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo cuanto podemos certificar, en honor a la verdad y para los fines pertinentes.

Atentamente,



PhD. Silvia Mariana
Haro Rivera

TUTOR



Mgs. Henry Mauricio
Villa Yanez

**MIEMBRO DEL
TRIBUNAL 1**



PhD. Manuel Antonio
Meneses Freire

**MIEMBRO DEL
TRIBUNAL 2**



Campus La Dolorosa
Av. Eloy Alfaro y 10 de Agosto
Teléfono (593-3) 373-0880, ext. 2002
Riobamba - Ecuador.

Unach.edu.ec
en movimiento



Dirección de
Posgrado
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO



Riobamba, 31 de julio de 2025

CERTIFICADO

De mi consideración:

Yo, Silvia Mariana Haro Rivera, certifico que Andrés Eduardo Salazar Andrade con cédula de identidad No. 1804471116 estudiante del programa de Maestría en Matemática Aplicada con mención en Matemática Computacional, tercera cohorte (2023-2024), presentó su trabajo de titulación bajo la modalidad de Proyecto de titulación con componente de investigación aplicada/desarrollo denominado: Análisis de la variabilidad y precisión en la estimación de intervalos de confianza en pruebas de hipótesis con desviaciones estándar variables: aplicación a datos de física de partículas del CERN y Zenodo, el mismo que fue sometido al sistema de verificación de similitud de contenido COMPILATION identificando el **0%** de similitud en el texto y el **7%** de similitud en inteligencia artificial.

Es todo en cuanto puedo certificar en honor a la verdad.

Atentamente,



SILVIA MARIANA HARO
RIVERA

Dra. Silvia Mariana Haro Rivera

CI: 0603082041

Adj.-

- Resultado del análisis de similitud (Compilation)



sQudable

Av. Eloy Alfaro y 10 de Agosto
Teléfono (593-3) 373-0880, ext. 2100 - 29
Riobamba - Ecuador
Unach.edu.ec
en innovación

Dedicatoria

Dedico este trabajo a mi madre, pilar fundamental de mi vida, cuya constancia, esfuerzo y amor incondicional me han inspirado a no rendirme ante los desafíos, por su sabiduría y fortaleza silenciosa. También dedico esta tesis a mi hermana y sobrinas, por creer en mí incluso en los momentos más inciertos. A todos ellos, gracias por ser mi sostén emocional y mi guía espiritual.

Agradecimiento

Agradezco profundamente a Dios, por brindarme la vida, la fortaleza y la sabiduría necesarias para culminar esta etapa académica con perseverancia y dedicación.

Mi más sincero agradecimiento a la Universidad Nacional de Chimborazo por acogerme en sus aulas y desarrollar profesionales en el país.

A los docentes del programa de posgrado, por transmitir su pasión por la ciencia y por impulsar en mí el pensamiento crítico y analítico, fundamentales en la construcción de este trabajo. A mis compañeros y compañeras de maestría, por los diálogos compartidos, el apoyo mutuo y la colaboración académica que hicieron más llevadero este recorrido.

Índice General

Declaración y Cesión de Derechos de Autoría.....	ii
Acta de Culminación de Trabajo de Titulación.....	iii
Certificado de Contenido De Similitud.....	iv
Dedicatoria.....	v
Agradecimiento.....	vi
Índice General.....	vii
Índice de Tablas.....	x
Índice de Figuras.....	xi
Resumen.....	xii
Abstract.....	xiii
Introducción.....	1
Capítulo 1 Generalidades.....	4
1.1 Planteamiento del problema.....	4
1.2 Justificación de la investigación.....	5
1.3 Objetivos.....	7
1.3.1 Objetivo General.....	7
1.3.2 Objetivos Específicos.....	7
Capítulo 2 Estado del Arte y la Práctica.....	9
2.1 Antecedentes investigativos.....	9
2.2 Fundamentación legal.....	11
2.2.1 Legislación internacional.....	11
2.2.2 Legislación nacional (Ecuador).....	12
2.3 Fundamentos de la inferencia estadística.....	13
2.3.1 Estimación de parámetros.....	13

2.3.2 Teorema del límite central.....	13
2.3.3 Propiedades de los intervalos de confianza.....	14
2.3.4 Bases matemáticas del intervalo clásico.....	14
2.4 Intervalos de confianza para la desviación estándar.....	15
2.4.1 Construcción clásica bajo normalidad.....	15
2.4.2 Limitaciones en presencia de heterocedasticidad.....	15
2.5 Estadística no paramétrica y submuestreo.....	16
2.5.1 Métodos comparativos para estimar intervalos.....	17
2.5.2 Indicadores de evaluación: precisión y eficiencia de los intervalos bajo heterocedasticidad.....	21
2.6 Aplicaciones en física de partículas: variable Jet _{pt} y jets del CMS.....	23
Capítulo 3 Diseño Metodológico.....	26
3.1 Enfoque de la investigación.....	26
3.2 Diseño de la investigación.....	27
3.3 Tipo de investigación.....	28
3.4 Nivel de investigación.....	30
3.5 Técnicas e instrumentos de recolección de datos.....	31
3.6 Técnicas para el procesamiento e interpretación de datos.....	33
3.7 Población y muestra.....	36
3.7.1 Población.....	36
3.7.2 Tamaño de la Muestra.....	37
3.8 Formulación de hipótesis estadísticas.....	39
3.9 Tratamiento estadístico.....	40
Capítulo 4 Análisis y Discusión de los Resultados.....	44
4.1 Análisis descriptivo de los resultados.....	44
4.2 Análisis comparativo de los métodos de estimación.....	47

4.2.1 Ancho del IC según la desviación estándar estimada - Método de transformación logarítmica ($\log - x^2$).....	47
4.2.2 Ancho del IC según la desviación estándar estimada - Método de cuantiles empíricos.....	49
4.2.3 Ancho del IC según la desviación estándar estimada - Método Bootstrap.....	50
4.3 Discusión de los resultados.....	58
Capítulo 5 Marco Propositivo.....	61
5.1 Planificación de la actividad preventiva.....	61
Conclusiones.....	63
Recomendaciones.....	65
Referencias Bibliográficas.....	66
Anexos.....	71

Índice de Tablas

Tabla 1 Paquetes y librerías para el análisis	36
Tabla 2 Resumen de resultados de intervalos de confianza (IC) para la desviación estándar muestral usando los métodos estudiados	56
Tabla 3 Matriz comparativa final entre los tres métodos con base en los indicadores de eficiencia en las submuestras.....	56

Índice de Figuras

Figura 1 Normalidad de los datos analizados por el detector CMS	44
Figura 2 Desviación estándar para cada submuestra.	46
Figura 3. Ancho del IC según la desviación estándar estimada - Método de transformación logarítmica	48
Figura 4 Ancho del IC según la desviación estándar estimada - Método de cuantiles empíricos	50
Figura 5 Ancho del IC según la desviación estándar estimada - Método Bootstrap	50
Figura 6 Relación entre desviación estándar estimada y ancho del IC	52
Figura 7 Diagrama de cajas del ancho del IC por método	55

Resumen

El presente estudio, titulado “Análisis de la variabilidad y precisión en la estimación de intervalos de confianza en pruebas de hipótesis con desviaciones estándar variables: aplicación a datos de jets de física de partículas del CERN y Zenodo”, tiene como propósito evaluar comparativamente la eficacia de tres métodos estadísticos no paramétricos en la construcción de intervalos de confianza para la desviación estándar muestral, en presencia de heterocedasticidad.

La investigación toma como base empírica la variable física momento transversal del jet (Jet_{pt}), proveniente de datos experimentales reales del detector CMS. El objetivo general fue analizar la precisión y eficiencia de los intervalos generados mediante tres enfoques: la transformación logarítmica del intervalo chi-cuadrado, el método de cuantiles empíricos y el bootstrap percentil. Se empleó un diseño cuantitativo de tipo comparativo con submuestreo aleatorio y procesamiento en lenguaje R, aplicando 113 submuestras por escenario.

Los resultados revelan que el método logarítmico presenta la menor longitud promedio del intervalo, mientras que el bootstrap mantiene un adecuado balance entre cobertura empírica y amplitud. El enfoque de cuantiles, aunque robusto en distribuciones sesgadas, mostró mayor dispersión bajo condiciones de alta variabilidad.

Se concluye que el método logarítmico ofrece la mejor precisión inferencial bajo heterocedasticidad, y que el bootstrap constituye una herramienta flexible frente a distribuciones no normales, validando así la importancia del enfoque no paramétrico en escenarios reales de alta complejidad estadística.

Palabras claves: intervalos de confianza, desviación estándar, heterocedasticidad, métodos no paramétricos, submuestreo.

ABSTRACT

The current research work 'Analysis of variability and precision in the estimation of confidence intervals in hypothesis testing with variable standard deviations: application to particle physics jet data from CERN and Zenodo', aims to comparatively evaluate the effectiveness of three non-parametric statistical methods in constructing confidence intervals for the sample standard deviation in the presence of heteroscedasticity.

The research is based on the empirical variable transverse momentum of the jet (Jet_pt), derived from real experimental data from the CMS detector. The overall objective was to analyze the precision and efficiency of the intervals generated using three approaches: the logarithmic transformation of the chi-square interval, the empirical quantile method, and the percentile bootstrap. A comparative quantitative design with random subsampling and processing in R language was used, applying 113 subsamples per scenario.

The results reveal that the logarithmic method has the shortest average interval length, while the bootstrap maintains an adequate balance between empirical coverage and width. The quantile approach, although robust in skewed distributions, showed greater dispersion under conditions of high variability.

It is concluded that the logarithmic method offers the best inferential precision under heteroscedasticity, and that bootstrap is a flexible tool for non-normal distributions, thus validating the importance of the non-parametric approach in real scenarios of high statistical complexity.

Keywords: confidence intervals, standard deviation, heteroscedasticity, non-parametric methods, subsampling.

Reviewed by:



MISHELL GABRIELA
SALAO ESPINOZA

Mgtr. Mishell Salao Espinoza
ENGLISH PROFESSOR
C.C. 0650151566

Introducción

En contextos experimentales reales, es habitual observar fluctuaciones en la dispersión de los datos, lo cual compromete la fiabilidad de las estimaciones realizadas a través de intervalos de confianza. Esta situación, asociada a la presencia de desviaciones estándar no constantes, plantea un reto importante para las técnicas inferenciales clásicas que asumen homocedasticidad, particularmente cuando la distribución subyacente no es normal. Ante ello, se vuelve necesario analizar rigurosamente el rendimiento de métodos alternativos capaces de ajustarse a escenarios con variabilidad en la varianza, como el log-chi-cuadrado o el bootstrap.

El presente trabajo está direccionado al análisis técnico de estimación de intervalos de confianza robustos a través de desviaciones estándar muestrales en datos no paramétricos. La variable escogida por su variabilidad de datos es el momento transversal de jets (Jet_{pt}) en colisiones de partículas.

Esta variable es ampliamente empleada en el ámbito de la física de altas energías, exhibe una notable variabilidad y asimetría en su distribución convirtiéndose en un modelo apropiado para la investigación en el comportamiento de estimadores bajo escenarios de heterocedasticidad de datos.

Con el fin de evaluar el desempeño de diferentes enfoques inferenciales, se lleva a cabo una comparación entre tres métodos estadísticos no paramétricos por su variación de datos, los cuales son: transformación logarítmica aplicada a intervalos clásicos basado en la distribución chi-cuadrado, el método de cuantiles empíricos a partir de percentiles directos en submuestras, y el remuestreo bootstrap percentil. Estos métodos permiten estimar

intervalos sin suponer una distribución subyacente específica los cuales son idóneos para datos reales no normales y con estructuras complejas.

El uso del lenguaje R, es ideal para análisis estadísticos sin asumir distribuciones específicas mostrando gráficamente hallazgos. A través de simulaciones computacionales, se crean pequeños grupos de datos para probar y evaluar la eficacia de distintos enfoques de estimación, y así poder contrastar cómo funcionaban. En ello se utilizaron métricas de presión y confianza como es el error estándar y el ancho de amplitud.

Rodríguez Medina et al. (2021) explican que la prueba de Wald clásica se fundamenta en estimaciones derivadas de la muestra completa de los datos y no en el análisis de submuestras, siendo una técnica ampliamente utilizada para contrastar hipótesis sobre parámetros poblacionales. No obstante, también se señala que los estimadores por máxima verosimilitud tienden a aproximarse más rápidamente a la normalidad, lo que sugiere la necesidad de una elección cuidadosa del método en función de la estructura y distribución de los datos. Por su parte, (Rakhimov, 2024) argumenta que los intervalos de confianza generados mediante técnicas bootstrap por sesgos o percentiles pueden ofrecer ventajas significativas cuando se presenta heterocedasticidad en los datos, ya que no requieren supuestos estrictos sobre la homogeneidad de la varianza. Aunque el método jackknife no se aplica en el presente estudio, se reconoce que investigaciones anteriores, como la de (Severiano A. et al., 2011), su capacidad para generar estimaciones más precisas que el bootstrap en contextos donde se evalúan medidas de concordancia entre pares, lo que reafirma la utilidad de los enfoques no paramétricos en escenarios complejos.

El aporte de esta investigación contribuye desde una perspectiva metodológica permitiendo contrastar la eficacia de distintas técnicas no paramétricas para la construcción

de intervalos de confianza de dispersión, desde el ámbito científico, promueve la incorporación de métodos estadísticos adaptativos no paramétricos en el análisis de datos experimentales con alta variabilidad; y finalmente, en el plano académico ayuda al fortalecimiento de habilidades analíticas mediante la aplicación de herramientas estadísticas avanzadas que permiten una mejor interpretación de fenómenos empíricos con datos reales y de gran impacto.

Este documento se organiza en cinco capítulos. El Capítulo 1 presenta el planteamiento del problema, los objetivos y la justificación. El Capítulo 2 desarrolla el marco teórico y los fundamentos estadísticos del estudio a través de los diferentes parámetros a utilizar en la presente investigación. El Capítulo 3 describe la metodología aplicada, incluyendo la selección de datos, el diseño experimental, hipótesis, métodos de empleo y los procedimientos de estimación. En el Capítulo 4 se exponen los resultados y su análisis comparativo entre los métodos a trabajar. Finalmente, el Capítulo 5 presenta las conclusiones, limitaciones y recomendaciones para futuras investigaciones.

Capítulo 1

Generalidades

1.1 Planteamiento del problema

La estimación de parámetros poblacionales a partir de muestras aleatorias constituye una de las tareas centrales de la inferencia estadística. Uno de los enfoques más extendidos para esta finalidad es la construcción de intervalos de confianza, los cuales permiten acotar el valor probable de un parámetro mediante una medida de incertidumbre asociada. No obstante, la formulación clásica de estos intervalos descansa sobre supuestos que no siempre se cumplen en contextos experimentales reales. Entre ellos, destaca la homocedasticidad, es decir, la suposición de varianza constante a lo largo de las observaciones. Cuando este supuesto se viola, como sucede en entornos donde la dispersión de los datos varía según las condiciones o subconjuntos muestrales, la validez de los intervalos puede verse afectada.

La complejidad de esta problemática se observa claramente al analizar los datos provenientes de experimentos en física de partículas por su variabilidad, especialmente aquellos registrados por el detector CMS (Compact Muon Solenoid) del CERN (Organización Europea para la Investigación Nuclear). Los conjuntos de datos generados en colisiones protón-protón de alta energía contienen información sobre jets de partículas, dentro de los cuales la variable del momento transversal del jet (Jet_{pt}) presenta distribuciones con marcada asimetría y elevada dispersión entre eventos. Esta naturaleza heterogénea dificulta la aplicación de técnicas inferenciales convencionales que suponen normalidad en los datos y homogeneidad en la varianza, lo que evidencia la necesidad de métodos estadísticos más flexibles.

Frente a esta situación, se plantea como objetivo central analizar el comportamiento de tres métodos de estimación de intervalos de confianza para la desviación estándar muestral en presencia de heterocedasticidad y ausencia de normalidad. Específicamente, se compararán el método de transformación logarítmica aplicado al intervalo chi-cuadrado, el método de cuantiles empíricos (percentiles directos de submuestras), y el enfoque bootstrap basado en remuestreo con reemplazo. Estos métodos se aplicarán sobre submuestras extraídas de los datos reales del CMS, lo cual permitirá observar cómo varía la precisión de los estimadores a medida que cambian el tamaño y la dispersión de las muestras.

La importancia de esta investigación radica en que la estimación adecuada de la desviación estándar y la construcción precisa de intervalos de confianza son fundamentales para la interpretación estadística de fenómenos físicos. Una metodología robusta no solo contribuye a mejorar la inferencia sobre parámetros de dispersión, sino que también proporciona lineamientos metodológicos replicables para otros campos donde los datos presentan estructuras no paramétricas y dispersión variable.

1.2 Justificación de la investigación

Esta investigación surge de la necesidad metodológica y aplicada de evaluar la precisión de los intervalos de confianza bajo condiciones de heterocedasticidad, es decir, cuando la varianza de los datos no se mantiene constante. Este fenómeno representa un desafío significativo para los enfoques clásicos de inferencia estadística, que suelen fundamentarse en el supuesto de homocedasticidad y normalidad para construir intervalos alrededor de parámetros poblacionales (Efron & Hastie, 2016; Rakhimov, 2024).

Desde el punto de vista matemático, el problema se enfoca en la construcción de intervalos de confianza para la desviación estándar muestral, particularmente cuando el

estimador presenta alta variabilidad entre subconjuntos (Navidi, 2021). En condiciones ideales, la estimación clásica basada en distribuciones normales o chi-cuadrado puede expresarse como:

$$IC = \left[\bar{x} \pm z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right] \quad (1)$$

donde \bar{x} es la media muestral, $z_{\frac{\alpha}{2}}$ es el valor crítico de la distribución normal estándar y s es la desviación estándar muestral. Sin embargo, en presencia de varianza no constante, este tipo de expresiones pierde validez, lo cual impacta la cobertura nominal y la eficiencia del intervalo (Yang & Xie, 2020).

Este reto se acentúa en contextos de alta complejidad experimental, como en la física de partículas. En particular, el experimento CMS del CERN registra jets de partículas producto de colisiones protón-protón, donde el momento transversal del jet (Jet_pt) presenta una distribución sesgada y dispersa. Esta variable se considera adecuada para estudiar empíricamente los efectos de la heterocedasticidad sobre la confiabilidad de los intervalos de confianza.

En consecuencia, se comparan tres enfoques robustos para la estimación de intervalos de confianza: la transformación logarítmica del intervalo chi-cuadrado, el método de cuantiles empíricos mediante percentiles directos de submuestras, y el procedimiento bootstrap (Efron & Hastie, 2016), que genera intervalos mediante remuestreo con reemplazo:

$$IC^* = \left[\hat{\theta}_{inf}^*, \hat{\theta}_{sup}^* \right] \quad (2)$$

donde los valores $\hat{\theta}_{inf}^*$, $\hat{\theta}_{sup}^*$ y $\hat{\theta}_{med}^*$ derivan de la distribución bootstrap. La comparación se basa en indicadores como la amplitud del intervalo, el error estándar

estimado y la tasa de cobertura empírica, aplicados sobre submuestras obtenidas de datos reales con distintos niveles de dispersión (Ciucan-Rusu et al., 2022).

La metodología empleada en la investigación resulta relevante para el análisis robusto a través de paquetes estadísticos como readxl y ggplot2. Este tipo de aplicación además de sistematizar el análisis estadístico permite su utilización en otras áreas investigativas como la bioestadística, la economía, ciencias experimentales que manifiesten datos robustos.

La relevancia del presente estudio radica en la significativa contribución al análisis estadístico robusto. Bajo condiciones no ideas propone criterios de estimación aplicables para datos reales en condiciones múltiples como en este caso en colisiones de partículas protón-protón.

1.3 Objetivos

1.3.1 Objetivo General

- Analizar la variabilidad y precisión en la estimación de intervalos de confianza en pruebas de hipótesis con desviaciones estándar variables, mediante la aplicación y comparación de métodos estadísticos sobre datos de jets de física de partículas del CERN y Zenodo.

1.3.2 Objetivos Específicos

- Analizar el impacto de la variabilidad de la desviación estándar en la construcción de intervalos de confianza para la variable física momento transversal (Jet_{pt}), utilizando datos provenientes de experimentos reales del CMS del CERN.

- Aplicar métodos de estimación estadística no paramétricos, entre ellos la transformación logarítmica del intervalo chi-cuadrado, el método de cuantiles empíricos y el procedimiento de remuestreo bootstrap, en conjuntos de datos caracterizados por distintos niveles de dispersión.
- Evaluar la precisión y la eficiencia de los intervalos generados por cada método bajo condiciones de heterocedasticidad empleando indicadores la amplitud del intervalo.

Capítulo 2

Estado del Arte y la Práctica

2.1 Antecedentes investigativos

La evaluación de la precisión en la estimación de intervalos de confianza bajo condiciones de heterocedasticidad ha motivado diversos estudios en estadística aplicada y computacional. A continuación, se presentan algunas investigaciones recientes que abordan enfoques no paramétricos y métodos robustos para estimar parámetros de dispersión, y que aportan elementos teóricos y metodológicos relevantes para el desarrollo de esta tesis.

En primer lugar, Rakhimov (2024) desarrolló un estudio de simulación para analizar el rendimiento del método bootstrap en modelos lineales afectados por heterocedasticidad. Frente a los métodos tradicionales el autor realiza una comparación del desempeño de este enfoque basados en la normalidad, determinando mediante bootstrap una mejor estabilidad de los datos cuando la desviación estándar no era constante. A través de esto se sustenta que el uso del método de bootstrap se utiliza para datos robustos donde los supuestos clásicos no se cumplen en desarrollo de datos no paramétricos y con submuestras.

Por otro lado, Zitzmann et al. (2023) propusieron una variante computacionalmente optimizada del método jackknife, orientada a mejorar la estimación de errores estándar en modelos multinivel bajo heterocedasticidad. A través de un enfoque de simulación Monte Carlo, demostraron que su técnica superó en eficiencia y consistencia a los estimadores clásicos. Aunque la presente tesis no emplea jackknife, este estudio confirma que los métodos alternativos al enfoque clásico, como los utilizados aquí: bootstrap, cuantiles y log chi-cuadrado pertinentes para escenarios de varianza no constante y datos complejos.

En un análisis más amplio, (Ciucan-Rusu et al., 2022) realizaron una comparación de varios métodos bootstrap aplicados a distribuciones no normales. Entre los enfoques evaluados estuvieron el bootstrap normal, el percentil y el bootstrap studentizado. Los autores concluyeron que los métodos basados en percentiles directos presentaron mayor precisión y una tasa de cobertura más estable, especialmente en muestras pequeñas y distribuciones asimétricas. Este hallazgo es particularmente relevante para esta investigación, ya que uno de los métodos comparados será precisamente el de cuantiles empíricos mediante percentiles de submuestras.

Zhang et al. (2023) investigaron modelos de funciones lineales a través de bootstrap en errores heterocedásticos. Se aplicó el remuestreo emparejado evaluado por pruebas empíricas en datos masivos. En su conclusión respaldan el mejoramiento de la cobertura de ancho de amplitud y estabilidad de los estimadores en datos heterocedasticidad respaldando a lo propuesto en esta investigación en dispersión de datos que no presenta homogeneidad en su desviación estándar.

Por otro lado, Fernández-Casal et al. (2024) dieron a conocer un método de análisis de bootstrap percentiles para datos no paramétricos en datos reales con heterocedasticidad. Confirmaron que su metodología determinaba una estimación precisa en presencia de datos no normales. Los aportes de esta investigación están en concordancia con la lógica metodológica empleada. A través de lo cual se respalda la aplicabilidad de procedimientos robustos en estructuras complejas.

Estos antecedentes permiten identificar un consenso en la literatura contemporánea sobre la necesidad de abandonar los métodos clásicos en presencia de heterocedasticidad y adoptar estrategias estadísticas no paramétricas y basadas en remuestreo. La presente tesis se diferencia de los estudios revisados al aplicar estos métodos no sólo desde una perspectiva

simulada, sino sobre datos experimentales reales en física de partículas, específicamente sobre el momento transversal de jets (Jet_{pt}), lo que representa una contribución original tanto en el plano estadístico como en su aplicación interdisciplinaria.

2.2 Fundamentación legal

La presente investigación se apoya en datos científicos obtenidos del experimento CMS (Compact Muon Solenoid) del CERN y de repositorios científicos de acceso abierto, por lo cual requiere fundamentarse jurídicamente tanto en la normativa internacional sobre uso de datos experimentales, como en las leyes nacionales que regulan la producción y divulgación de conocimientos en el ámbito académico. La utilización de datos públicos debe garantizar el respeto a principios de integridad científica, transparencia y responsabilidad ética. En este sentido, se establece a continuación el marco legal relevante.

2.2.1 Legislación internacional

El Convenio de Berna para la Protección de las Obras Literarias y Artísticas, adoptado por la Organización Mundial de la Propiedad Intelectual (OMPI), establece principios fundamentales sobre el respeto a los derechos de autor en la utilización y reproducción de obras intelectuales. En el contexto académico, este tratado garantiza que el uso de datos, incluso de carácter científico, debe realizarse con la debida atribución a los autores o instituciones que los generan (Convenio de Berna Para La Protección de Las Obras Literarias y Artísticas, 1971). En esta tesis, al emplearse datos abiertos del CERN y Zenodo, se cumple con dicho principio al respetar las licencias de uso y la citación correspondiente.

Asimismo, la Declaración de Berlín sobre el Acceso Abierto al Conocimiento en Ciencias y Humanidades, firmada en 2003, promueve el libre acceso a los resultados de investigaciones científicas, siempre que se asegure el reconocimiento de la autoría y se

permita su reutilización en favor del avance científico. Esta declaración constituye uno de los pilares del movimiento de ciencia abierta, al que el CERN se adhiere mediante su portal oficial de datos abiertos (CERN, 2021; Max Planck Society, 2003).

Además, los datos utilizados en esta investigación están amparados bajo la Licencia Creative Commons Attribution 4.0 (CC BY 4.0), la cual autoriza la utilización, modificación y distribución del contenido siempre que se atribuya adecuadamente a la fuente. Esta licencia es ampliamente empleada por plataformas científicas como Zenodo, lo que permite el uso académico de los datos experimentales en condiciones claras y legalmente aceptadas (Creative Commons, 2024).

2.2.2 Legislación nacional (Ecuador)

En el ámbito jurídico ecuatoriano, la Ley Orgánica de Educación Superior (LOES) establece, en su artículo 7, que uno de los fines de la educación superior es fomentar la generación y difusión del conocimiento científico. Así mismo, el artículo 88 promueve principios de ética académica, integridad y reconocimiento de autoría, elementos esenciales para el desarrollo de investigaciones con responsabilidad (Asamblea Nacional del Ecuador, 2018).

Complementariamente, el Código Orgánico de la Economía Social de los Conocimientos, Creatividad e Innovación (Código Ingenios) plantea en sus artículos 4 y 27 la democratización del acceso al conocimiento, fomentando el uso de repositorios de libre acceso, así como la protección de los derechos morales y patrimoniales de los creadores. Este código legitima la investigación basada en ciencia abierta y promueve la reutilización de datos científicos para fines educativos y académicos (Código Orgánico de La Economía Social de Los Conocimientos, Creatividad e Innovación, 2016)

Por último, la Ley Orgánica de Transparencia y Acceso a la Información Pública (LOTAIP) reconoce el derecho de las personas a acceder libremente a la información generada por las instituciones públicas, incluyendo bases de datos científicas o educativas. Este principio respalda el uso legítimo de información pública para fines investigativos, como es el caso de los datos del CMS del CERN, disponibles en repositorios abiertos (Asamblea Nacional del Ecuador, 2004).

2.3 Fundamentos de la inferencia estadística

La inferencia estadística es el proceso mediante el cual se extraen conclusiones sobre poblaciones a partir de información limitada contenida en muestras, reconociendo la presencia de incertidumbre o error muestral. Este procedimiento se asienta en las teorías de probabilidad y en las propiedades de los estimadores como la bondad de ajuste y la eficiencia (Hirschauer et al., 2021). Su cometido principal se divide en dos tareas esenciales: la estimación de parámetros (puntual y por intervalo) y el contraste de hipótesis.

2.3.1 Estimación de parámetros

La estimación puntual consiste en determinar un valor único (como la media muestral) que represente el parámetro poblacional. Sin embargo, como tal estimación carece de información sobre su precisión, se recurre a la estimación por intervalo, que proporciona un rango dentro del cual se espera encontrar el parámetro con cierto nivel de confianza.

2.3.2 Teorema del límite central

El teorema del límite central (TLC) fundamenta la construcción de intervalos de confianza, ya que establece que, para muestras suficientemente grandes, la distribución de la media muestral se aproxima a una distribución normal, independientemente de la

distribución original de los datos. Este efecto permite usar técnicas como el IC clásico, incluso sin normalidad completa de la población (Hirschauer et al., 2021).

2.3.3 *Propiedades de los intervalos de confianza*

Un intervalo de confianza debe cumplir con los siguientes criterios:

- Cobertura nominal: la proporción de intervalos construidos que contienen el parámetro real debe aproximarse al nivel de confianza seleccionado (por ejemplo, 95 %).
- Longitud mínima: menor amplitud implica mayor precisión de estimación.
- Consistencia: al aumentar el tamaño de muestra, el intervalo debe reducir su amplitud y la estimación mejorar su exactitud.

2.3.4 *Bases matemáticas del intervalo clásico*

En el contexto frecuentista, para una variable x con media poblacional μ , el IC μ para cuando la desviación es desconocida se define como:

$$IC_{1-\alpha} = \left[\bar{x} \pm z_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right] \quad (3)$$

donde \bar{x} es la media muestral, s la desviación estándar muestral y z el cuantil normal asociado al nivel de confianza. Este intervalo posee una base probabilística sólida gracias al TLC, aunque su efectividad disminuye si n es pequeño o si existen problemas como heterocedasticidad en las varias observaciones, lo cual afecta directamente la cobertura y precisión de la estimación (Hirschauer et al., 2021).

2.4 Intervalos de confianza para la desviación estándar

En el campo de la inferencia estadística, la estimación de parámetros de dispersión como la desviación estándar poblacional (σ) constituye una herramienta clave en el análisis de datos cuantitativos. A partir de una muestra aleatoria X_1, X_2, \dots, X_n la desviación estándar muestral s es un estimador común de σ . Para evaluar la precisión de esta estimación, se construyen intervalos de confianza (IC), cuyo desempeño depende críticamente de los supuestos estadísticos subyacentes.

2.4.1 Construcción clásica bajo normalidad.

Cuando los datos provienen de una población normal $N(\mu, \sigma^2)$ la estadística.

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1} \quad (4)$$

permite construir un intervalo de confianza para σ^2 con nivel de confianza $1 - \alpha$.

El IC correspondiente es:

$$IC_{1-\alpha}(s^2) = \left[\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}, \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}, n-1}} \right] \quad (5)$$

Aplicando la raíz cuadrada se obtiene el IC para σ . Este procedimiento es ampliamente utilizado cuando se supone normalidad y varianza constante, siendo una técnica robusta bajo estos escenarios (Navidi, 2021).

2.4.2 Limitaciones en presencia de heterocedasticidad

No obstante, cuando los datos presentan heterocedasticidad, es decir la varianza no constante a lo largo de las observaciones, el método clásico puede volverse ineficiente. Estudios como el de Lumley (2012) muestran que la suposición de normalidad, aunque aceptable en muestras grandes, puede inducir errores significativos si se ignora la

variabilidad en la dispersión. En particular, los intervalos tienden a subestimar la verdadera incertidumbre y a perder cobertura nominal.

De acuerdo con Hayes & Cai (2007) , en modelos donde se aplican regresiones u otras técnicas inferenciales bajo condiciones de heterocedasticidad, los errores estándar estimados mediante métodos tradicionales pueden perder su consistencia, lo que compromete la validez de los intervalos de confianza construidos a partir de ellos. Esta limitación refuerza la necesidad de emplear metodologías y estrategias alternativas, particularmente aquellas que se basan en procedimientos robustos o para datos no paramétricos, los cuales permiten obtener estimaciones con mayor grado de confiabilidad sin depender del supuesto de varianza constante.

2.5 Estadística no paramétrica y submuestreo

La estadística no paramétrica resulta especialmente adecuada para datos que no satisfacen los supuestos clásicos de normalidad o varianza homogénea. En tales circunstancias, se recurre a métodos de remuestreo como el bootstrap percentil o a enfoques basados en cuantiles, los cuales permiten desarrollar procedimientos inferenciales sin depender de modelos paramétricos específicos (Fernández-Casal et al., 2024). Una técnica complementaria es el submuestreo, que consiste en extraer múltiples subconjuntos del conjunto original de datos con el propósito de estimar la variabilidad de los estadísticos, lo cual resulta ventajoso en presencia de heterocedasticidad o distribuciones no estándar.

En particular, el submuestreo aleatorio sin reemplazo constituye una estrategia eficaz para aproximar la distribución de una estadística sin necesidad de asumir una función para la población. Además, variantes como el bootstrap percentil aplicado sobre submuestras han sido desarrolladas con el objetivo de optimizar el rendimiento computacional en contextos

de datos masivos como las variables de protón-protón, ofreciendo estimaciones robustas sin requerir normalidad en la población (Ma et al., 2023).

En la estimación de intervalos de confianza con datos robustos los métodos no paramétricos dividido en submuestras forman parte esencial sobre todo cuando se emplea en datos empíricos como es el caso de la presente investigación a través de datos experimentales del LHC (Gran Colisionador de Hadrones) del CERN, con varianzas variables.

2.5.1 Métodos comparativos para estimar intervalos

Para el análisis de datos con supuestos no paramétricos y que presentan heterocedasticidad se requiere buscar alternativas que sean robustas y mantengan un análisis de separación de submuestras a la variabilidad de la varianza y que no son parte de una distribución poblacional, lo cual se vuelve crucial al momento de determinar la precisión y confiabilidad de los intervalos.

- **Método de transformación logarítmica ($\log - x^2$)**

El método TL (Transformación Logarítmica) en intervalos de confianza es una opción robusta para datos no paramétricos con desviación estándar no constante, especialmente para datos con heterocedasticidad. A veces se utiliza para corregir simetría de distribución chi-cuadrado. En conclusión, este método estabiliza la varianza y los intervalos de confianza determinando precisión.

Fundamento estadístico:

El procedimiento comienza aplicando una transformación logarítmica al estimador de la desviación estándar muestral \hat{s} :

$$\theta = \log(\hat{s}) \quad (6)$$

Bajo supuestos de normalidad, la varianza del estimador transformado puede aproximarse como:

$$EE_{\log(\hat{s})} \approx \frac{1}{\sqrt{2(S_k - 1)}} \quad (7)$$

lo cual permite construir un intervalo de confianza simétrico en la escala logarítmica:

$$\log(\hat{s}) \pm z_{\alpha/2} \cdot \frac{1}{\sqrt{2(S_k - 1)}} \quad (8)$$

donde $z_{\alpha/2}$ es el cuantil superior de la distribución normal estándar.

Finalmente, el intervalo se reconvierte a la escala original mediante la exponencial inversa, resultando en el intervalo final:

$$IC_{1-\alpha}(s) = \left[\hat{s} \cdot e^{-z_{\alpha/2}/\sqrt{2(n-1)}}, \hat{s} \cdot e^{+z_{\alpha/2}/\sqrt{2(n-1)}} \right] \quad (9)$$

$$\text{Ancho del } IC_{TL} = \hat{s} \cdot e^{\frac{z_{\alpha/2}}{\sqrt{2(S_k - 1)}}} - \hat{s} \cdot e^{-\frac{z_{\alpha/2}}{\sqrt{2(S_k - 1)}}} \quad (10)$$

Esta metodología produce intervalos contruidos de manera proporcional respecto al estimador central, lo que favorece una mayor fidelidad en la cobertura observada, especialmente en contextos donde los datos muestran asimetrías marcadas o niveles elevados de dispersión. Compite favorablemente con IC tradicionales, aunque en situaciones de valores atípicos extremos el (BCa) Bootstrap es más robusto (Jung et al., 2019).

- **Método de cuantiles empíricos (Percentil directo de submuestras)**

El método de cuartiles empíricos construye intervalos de confianza mediante distribución empírica a través de submuestras (remuestreos), también es llamada intervalo de percentiles directos. Esta distribución no es necesaria partir de la teoría específica y es utilizada especialmente en situaciones donde se muestra asimetría, es decir datos no paramétricos, varianzas no constantes o aceleración de sesgos.

Sea $X = \{x_1, x_2, \dots, x_n\}$ un conjunto de observaciones reales extraídas de un proceso experimental (por ejemplo, mediciones de momento transversal Jet_{pt} en colisiones protón-protón del CMS). Se generan S_k submuestras de tamaño ($m \leq n$), denotadas como:

$$S_{k_1}, S_{k_2}, \dots, S_{k_n} \quad (11)$$

Para cada submuestra S_k , se calcula un estimador \hat{s} , siendo en este caso la desviación estándar muestral. A partir de estos S_k valores se obtiene una distribución empírica del estimador \hat{s} :

$$\hat{s}_1, \hat{s}_2, \dots, \hat{s}_n \quad (12)$$

Para un nivel de confianza $1 - \alpha$ (por ejemplo, 95%), se ordenan los valores estimados y se seleccionan los percentiles correspondientes:

$$IC_{1-\alpha} = \left[Q_{\alpha/2}, Q_{1-\alpha/2} \right] \quad (13)$$

donde:

- $Q_{\alpha/2}$ es el percentil inferior (2,5%)
- $Q_{1-\alpha/2}$ es el percentil superior (97,5%)

Así, el intervalo de confianza final basado en cuantiles empíricos es:

$$IC_{1-\alpha}(\hat{s}) = \left[\hat{s}_{\frac{\alpha}{2}}, \hat{s}_{1-\frac{\alpha}{2}} \right] \quad (14)$$

$$\text{Ancho del } IC_{CE} = \left[Q_{1-\frac{\alpha}{2}}(\hat{s}) - Q_{\frac{\alpha}{2}}(\hat{s}) \right] \quad (15)$$

Este método no requiere suposiciones de normalidad ni homocedasticidad, y su efectividad aumenta con un número elevado de remuestreos (S_k), típicamente mayor a 1000. Además, es especialmente robusto en presencia de sesgo, colas largas o estructuras de datos no paramétricos (Justus et al., 2024).

Según (Ciucan-Rusu et al., 2022), en comparaciones empíricas con otros métodos bootstrap (como BCa o normal), el intervalo por percentiles directos ofreció una cobertura cercana al 95 % y mostró un mejor ajuste a la forma real de la distribución en muestras pequeñas y asimétricas, lo que lo convierte en una opción eficiente cuando se analizan variables físicas dispersas como Jet_pt.

- **Método Bootstrap percentil**

El método de percentiles bootstrap es una herramienta de remuestreo muy útil cuando las reglas tradicionales no se cumplen (como la normalidad en los datos o que la varianza sea parecida entre grupos). Básicamente, lo que hace es crear un montón de copias de la muestra original, sacando datos al azar con reemplazo. Dando una perspectiva de cómo se puede comportar realmente in intervalo. Para adaptarlos en la construcción de los percentiles de la distribución de los datos, sin tener que proyectar la forma de la población original (Wilcox, 2020).

Este enfoque no requiere suposiciones de normalidad ni manejo directo de ajustes por sesgo o aceleración, a diferencia del método Bootstrap por sesgo. Su implementación consiste en ordenar las réplicas bootstrap del estadístico y seleccionar directamente los percentiles correspondientes al nivel de confianza deseado (intervalos de confianza del 95 % se toman el percentil 2.5 % y el percentil 97.5 % de la distribución bootstrap simulada).

La fórmula del IC bajo el método es:

$$IC_{1-\alpha}^{percentil} = \left[\hat{s}_{\alpha/2}, \hat{s}_{(1-\frac{\alpha}{2})} \right] \quad (16)$$

Donde:

- \hat{s}_p representa el percentil p de la distribución Bootstrap del estimador \hat{s} .
- α es el nivel de significación (por ejemplo, 0.05 para un 95 % de confianza), y no se aplican correcciones por sesgo ni aceleración.

Este método es elegido por su acople computacional, su transparencia y su capacidad para adaptarse a contextos con datos que no cumplen con los supuestos tradicionales de la inferencia paramétrica (Ciucan-Rusu et al., 2022).

2.5.2 Indicadores de evaluación: precisión y eficiencia de los intervalos bajo heterocedasticidad.

La evaluación del desempeño de los intervalos de confianza construidos mediante métodos estadísticos no paramétricos; requiere indicadores cuantificables que permitan comparar su rendimiento bajo condiciones de heterocedasticidad. Se consideran tres indicadores fundamentales: la amplitud del intervalo, el error estándar del estimador y la tasa de cobertura empírica. Estos indicadores permiten medir la precisión, robustez y

confiabilidad de cada método de estimación aplicado sobre datos reales con alta dispersión, como los provenientes de la variable física Jet_pt.

Dado un conjunto de datos $X = \{x_1, x_2, \dots, x_n\}$, y considerando la estimación de la desviación estándar poblacional σ mediante submuestras, se evalúan los métodos en función de tres indicadores clave:

a) Precisión: Amplitud media del intervalo de confianza

La amplitud de un intervalo de confianza es una medida directa de la precisión de la estimación del parámetro.

La amplitud del intervalo de confianza $IC_{1-\alpha}^{(s_k)} = [L^{(s_k)}, U^{(s_k)}]$ en la submuestra s_k es:

$$\text{Ancho del IC} = \theta_U - \theta_L \tag{17}$$

Cuanto menor sea esta distancia, mayor será la precisión del estimador, siempre que se mantenga la tasa de cobertura deseada. La precisión se evalúa empíricamente como la media de las amplitudes obtenidas sobre múltiples repeticiones:

$$\bar{A} = \frac{1}{B} \sum_{s_k=1}^{s_k=n} (\theta_U^{(s_k)} - \theta_L^{(s_k)}) \tag{18}$$

Este criterio se ha aplicado con éxito en evaluaciones de métodos percentiles bootstrap y transformaciones logarítmicas, especialmente en condiciones de no normalidad

b) Error estándar del estimador

El error estándar (EE) mide la variabilidad del estimador entre submuestras:

$$EE(\hat{S}) = \sqrt{\frac{1}{s_{kn}-1} \sum_{s_k=1}^{s_{kn}} (\hat{S}^{(s_k)} - \bar{S})^2} \tag{19}$$

Este indicador refleja la estabilidad del procedimiento, siendo fundamental en métodos sensibles a remuestreo como el bootstrap (Wilcox, 2020).

Bajo escenarios donde la varianza de los datos no es constante, se ha evidenciado que la capacidad de los métodos bootstrap, particularmente aquellos basados en percentiles para mantener la cobertura nominal se ve comprometida cuando la muestra presenta sesgo o asimetría. Esta limitación resalta la necesidad de contrastar su rendimiento frente a alternativas metodológicas más robustas que consideren dichas irregularidades en la distribución de los datos.

2.6 Aplicaciones en física de partículas: variable Jet_{pt} y jets del CMS

En el campo de la física de partículas de alta energía, el análisis estadístico de los datos provenientes de colisionadores como el Large Hadron Collider (LHC) constituye una herramienta esencial para la reconstrucción e interpretación de fenómenos subatómicos. Uno de los experimentos más relevantes del LHC es el Compact Muon Solenoid (CMS), cuya misión científica incluye la detección y caracterización de jets de partículas originados en colisiones protón-protón a muy alta energía (Chatrchyan, 2017).

Entre las variables observables más significativas en estos eventos se encuentra el momento transversal del jet (Jet_{pt}), que representa la componente del momento del jet perpendicular al eje del haz de colisión. Esta magnitud resulta fundamental para inferir la dinámica de los procesos partónicos subyacentes en colisiones de partículas. Durante el año 2011, el acelerador operó colisionando protones a una energía de 7 TeV, lo que permitió registrar millones de eventos físicos en el experimento.

Desde el punto de vista estadístico, el Jet_{pt} presenta una distribución altamente asimétrica, con colas largas y una varianza no constante que varía según la energía de la

colisión, la configuración angular del detector y el algoritmo de reconstrucción empleado (Sirunyan et al., 2020). Esta heterogeneidad justifica el uso de técnicas no paramétricas como los métodos bootstrap o de cuantiles empíricos para la estimación de intervalos de confianza, ya que los supuestos clásicos de normalidad y homocedasticidad se ven frecuentemente vulnerados.

En esta investigación, el conjunto de datos empleado corresponde al Run2011A Jet AOD (campana específica de adquisición de datos en el LHC durante el año 2011) del experimento CMS, disponible en acceso abierto a través del CERN Open Data Portal. Estos datos permiten realizar un análisis detallado del comportamiento del Jet_{pt} en distintas condiciones experimentales, segmentando las observaciones en submuestras para capturar la variabilidad de la desviación estándar muestral. A través de esta estructura de análisis, se evalúa la precisión de diferentes métodos de estimación de intervalos en condiciones reales de dispersión no uniforme.

Estudios recientes respaldan este enfoque. Por ejemplo, Tripathee et al. (2019) muestran que la resolución de los jets reconstruidos depende críticamente de la fluctuación estadística en el Jet_{pt} , afectando la identificación de eventos raros como los bosones Z' o señales supersimétricas. Asimismo, Duarte (2021) propone modelos basados en aprendizaje automático para predecir la distribución del Jet_{pt} , pero subraya que incluso dichos enfoques requieren validación estadística robusta, especialmente bajo presencia de valores atípicos y heterocedasticidad.

En términos generales, el estudio de la variable Jet_{pt} representa un escenario idóneo para la evaluación de metodologías robustas en inferencia estadística, debido a su origen empírico, la complejidad inherente a su estructura de datos y su alta sensibilidad frente a la variabilidad derivada del proceso experimental.

En la física de partículas de alta energía, el momento lineal de una partícula se expresa usualmente en unidades de GeV/c, donde GeV equivale a gigaelectronvoltios ($1 \text{ GeV} = 10^9 \text{ eV}$) y c representa la velocidad de la luz en el vacío ($3 \times 10^8 \text{ m} \cdot \text{s}^{-1}$). Esta unidad proviene del formalismo relativista empleado en aceleradores como el LHC, donde las energías involucradas alcanzan escalas subatómicas extremas por las colisiones protón-protón. El uso de GeV/c permite una representación coherente entre las dimensiones de energía y momento, ya que dicha unidad corresponde al producto de masa por velocidad. En particular, esta notación se emplea para describir componentes específicos del momento, como el momento transversal del jet (Jet_{pt}).

Capítulo 3

Diseño Metodológico

3.1 Enfoque de la investigación

El presente estudio adopta un enfoque cuantitativo no paramétrico, fundamentado en el análisis estadístico de la variabilidad y precisión de métodos de estimación de intervalos de confianza, aplicados a datos experimentales provenientes del ámbito de la física de partículas. Este estudio se sitúa dentro del paradigma positivista, al orientarse hacia la descripción, modelización y comparación objetiva del comportamiento de diversos procedimientos inferenciales bajo condiciones de heterocedasticidad, empleando para ello herramientas estadísticas formales y validación empírica a través de métodos computacionales.

Metodológicamente, la investigación adopta un enfoque comparativo y aplicado, centrado en el contraste de tres técnicas para la estimación de intervalos de confianza de la desviación estándar muestral (transformación logarítmica del intervalo clásico, los percentiles empíricos y el método bootstrap percentil). Estas técnicas son evaluadas sobre submuestras obtenidas a partir de un conjunto de datos reales proporcionados por el detector CMS del CERN de carácter no paramétrico y la variabilidad estructural de dichos datos. Esta elección metodológica se respalda en trabajos recientes como el de Justus et al. (2024), quienes destacan la eficacia de los enfoques de remuestreo en escenarios caracterizados por distribuciones complejas y no estándar con datos no paramétricos.

Asimismo, la investigación se adscribe a un enfoque computacional aplicado, dado que el diseño experimental y la validación de los métodos se realiza a través de simulaciones en lenguaje R, con paquetes estadísticos especializados. Esta orientación permite, además,

replicar los resultados en otros contextos con estructura de datos similar, fortaleciendo la validez externa del análisis (Wilcox, 2020).

Este camino es crucial para entender cómo varían, en términos estadísticos, las estimaciones cuando la dispersión de los datos cambia. Esto es algo común en los datos que obtenemos en experimentos de física de alta energía (Sirunyan et al., 2020). Por lo tanto, el objetivo del estudio es doble: por un lado, validar si las técnicas que elegimos son realmente acertadas; y por otro, demostrar con ejemplos reales que tienen sentido en situaciones reales donde los datos no son siempre iguales, sino que cambian.

3.2 Diseño de la investigación

El diseño metodológico adoptado en esta investigación es de tipo cuantitativo, no experimental, transversal y comparativo, orientado al análisis estadístico de submuestras provenientes de datos experimentales con estructura heterocedástica. Al no existir manipulación de variables independientes ni intervención directa sobre los datos, el estudio se enmarca dentro de un diseño observacional con enfoque descriptivo-comparativo, caracterizado por la recolección y análisis de datos existentes en fuentes científicas abiertas, como los registros del experimento CMS del CERN.

El enfoque transversal de esta investigación se manifiesta en el análisis de los datos en un único momento temporal, sin contemplar un análisis longitudinal de las submuestras.

El análisis se apoya en datos específicos del conjunto Run2011AJetAOD del CERN, de la variable de momento transversal creando con ellos varias submuestras separadas. Esto nos deja ver cómo se comporta la desviación estándar muestral y sus rangos de confianza, teniendo en cuenta la diferente dispersión que hay en los datos reales.

El diseño comparativo se manifiesta en la evaluación sistemática de tres métodos estadísticos para estimar intervalos de confianza: el método de transformación logarítmica, el método de cuantiles empíricos y el bootstrap. Estos enfoques son aplicados sobre conjuntos de submuestras generadas aleatoriamente a partir de los datos reales. En cada escenario experimental, se mide y compara la eficiencia estadística de los métodos en términos de amplitud del intervalo, error estándar y cobertura empírica. Esta estrategia ha sido empleada exitosamente en investigaciones recientes que buscan validar procedimientos estadísticos en condiciones de varianza no constante (Ciucan-Rusu et al., 2022).

El diseño metodológico a través de la utilización del lenguaje de programación planteado interpreta los resultados en instrumentos de computación valorando el método.

Esta fase se apoya en los principios de la ciencia abierta, lo que significa que se busca que todo el proceso de investigación sea transparente y fácil de entender. Como señalan Efron & Hastie (2016), al usar métodos estadísticos que se pueden replicar, la investigación tendrá mayor fiabilidad, y se abre puerta a que más investigadores examinen a fondo desde diferentes ángulos.

3.3 Tipo de investigación

La presente investigación se encuentra bajo el enfoque aplicativo y comparativo con un análisis de datos cuantitativo y no paramétrico, analizados a través de métodos estadístico en estimación de intervalos. La precisión de estos intervalos en los parámetros estadísticos desarrolla metodologías robustas donde no se cumple los supuestos de normalidad tradicional.

Desde el punto de vista del propósito, se clasifica como investigación aplicada, ya que busca ofrecer una solución técnica a un problema específico: la estimación confiable de

la desviación estándar en conjuntos de datos reales y complejos, como los generados en experimentos de física de partículas. De acuerdo con Hernández et al. (2021), este tipo de estudios se orienta a la producción de conocimiento con valor práctico a respuestas efectivas frente a problemáticas concretas del entorno científico o técnico.

El análisis cuantitativo se determina bajo la recopilación de datos, los cuales son procesados en un sistema computacional, cuyos datos son obtenidos del CMS del CERN. Los métodos son cuantificados a medida del ancho del intervalo y el error estándar del sistema.

Además, se trata de una investigación comparativa, ya que confronta el rendimiento de tres enfoques distintos para la construcción de intervalos de confianza: el método de transformación logarítmica, los cuantiles empíricos y el bootstrap. Este tipo de estudios es especialmente útil cuando se pretende determinar cuál técnica ofrece mejor precisión bajo ciertas condiciones distribucionales, como lo demuestran (Ciucan-Rusu et al., 2022) en su análisis de métodos de remuestreo en escenarios no normales.

Finalmente, se clasifica como una investigación no paramétrica, debido a que los datos analizados (momento transversal de los jets de partículas (Jet_{pt}) no siguen una distribución normal y presentan evidencia de varianza no constante. Por ello, se recurre a métodos inferenciales que no dependen de supuestos paramétricos rígidos, lo cual es coherente con los lineamientos metodológicos para el tratamiento de datos experimentales altamente dispersos (Justus et al., 2024).

3.4 Nivel de investigación

El nivel de investigación adoptado en este estudio es de tipo explicativo-comparativo, dado que se busca identificar relaciones estadísticas entre los métodos de

estimación de intervalos de confianza y su comportamiento en presencia de heterocedasticidad, utilizando datos experimentales del campo de la física de partículas. Este nivel permite ir más allá de la simple descripción o correlación, ya que profundiza en las causas que explican la variación de la precisión y eficiencia entre los distintos métodos aplicados sobre la variable física (Jet_{pt}), que representa el momento transversal de los jets de partículas.

El nivel explicativo de esta investigación se fundamenta en que no solo se limita a la medición de la dispersión muestral y a la construcción de intervalos de confianza, sino que profundiza en el análisis del impacto que dicha variabilidad tiene sobre el desempeño de cada método considerado examinando el error estándar y la amplitud de los intervalos generados. Esta aproximación responde a lo planteado por Efron & Hastie (2016), quienes argumentan que los modelos estadísticos explicativos deben buscar en los mecanismos que originan la variabilidad de los estimadores, especialmente en escenarios donde los supuestos clásicos de homocedasticidad y normalidad no se cumplen.

Asimismo, el estudio tiene un componente comparativo, al establecer un marco riguroso de análisis entre tres métodos: transformación logarítmica ($\log - X^2$), percentil empírico y bootstrap. Esta comparación permite generar evidencia cuantitativa sobre cuál de ellos proporciona resultados más robustos bajo distintos escenarios de dispersión, replicando procedimientos similares a los utilizados en estudios de simulación estadística moderna (Justus et al., 2024)

Por su naturaleza matemática, computacional y aplicada, este tipo de investigación exige un nivel alto de formalismo estadístico. En ese sentido, el diseño metodológico también considera como parte del nivel explicativo la evaluación de comportamiento de los estimadores bajo condiciones reales de heterogeneidad, basándose en simulaciones

controladas y submuestreo estratificado, técnica ampliamente reconocida en inferencia moderna para evaluar precisión de métodos no paramétrico (Ciucan-Rusu et al., 2022).

3.5 Técnicas e instrumentos de recolección de datos

La presente investigación se fundamenta en el análisis cuantitativo de datos secundarios, obtenidos del repositorio abierto del experimento CMS (Compact Muon Solenoid) del CERN, específicamente del conjunto de colisiones protón-protón correspondientes al período Run2011A Jet AOD. Este conjunto de datos, estructurado bajo el formato ROOT, contiene eventos experimentales reales recolectados a 7 TeV de energía en el Gran Colisionador de Hadrones (LHC). La selección se ha orientado a la variable física denominada momento transversal del Jet p_t , la cual presenta una alta variabilidad y es estadísticamente relevante en el estudio de la dispersión y la heterocedasticidad (Chatrchyan, 2017).

En recolección de los datos se realizó a través de la descarga de la base del CERN (página principal) en formato ROOT, siendo una arquitectura de datos parte de Python. La data extraída sobre la variable analizar (momento transversal) fue convertida en formato .csv de forma eficiente para su análisis estadístico en RStudio. Es importante recalcar que se escoge este tipo de lengua por su eficiencia en el trabajo con grandes volúmenes de datos.

La técnica de recolección implementada es de carácter documental computacional y se apoya en un diseño de submuestreo aleatorio con reemplazo, con tamaños variables de submuestras ($n = 1000, 2000, 3000, 4000$) hasta completar n . Este enfoque permite evaluar la estabilidad de las estimaciones de la desviación estándar bajo distintas condiciones de dispersión muestral. Cada submuestra sirve como base para calcular el estimador s (desviación estándar muestral), sobre el cual se construyen intervalos de

confianza utilizando tres enfoques metodológicos distintos: el intervalo transformado logarítmicamente basado en chi-cuadrado, el intervalo de cuantiles empíricos y el intervalo tipo bootstrap (Justus et al., 2024; Wilcox, 2020).

Los instrumentos principales son:

- Scripts automatizados en R para generar submuestras, aplicar los métodos estadísticos, calcular indicadores de evaluación (precisión, amplitud, error estándar, cobertura), y visualizar resultados mediante ggplot2.
- Protocolos de validación empírica, que incluyen pruebas gráficas, comparación de la longitud de los intervalos, simulaciones repetidas ($B = 1000$ repeticiones) y análisis de cobertura empírica.
- Bitácoras digitales de procesamiento, que documentan cada etapa del flujo de datos y garantizan la trazabilidad metodológica conforme a estándares de reproducibilidad científica (Efron & Hastie, 2016).

La combinación de los diferentes métodos utilizados valida el desarrollo del estudio al controlar las restricciones analíticas y disminución de sesgos. Además, la transformación de los datos a través de su complejidad, facilitan su análisis (réplicas) en la construcción de escenarios computacionales (Ciucan-Rusu et al., 2022).

- **Herramientas computacionales utilizadas**

Para el procesamiento estadístico, generación de submuestras, estimación de intervalos de confianza y visualización gráfica, se empleó el lenguaje de programación R, reconocido por su versatilidad en análisis estadístico avanzado y su amplia disponibilidad de bibliotecas especializadas. En particular, se utilizaron los paquetes boot, ggplot2, dplyr, purrr y tibble, los cuales permitieron implementar los métodos de estimación no

paramétricos (transformación logarítmica, cuantiles empíricos y bootstrap) con control preciso sobre las condiciones de remuestreo, generación de percentiles y cálculo de errores estándar.

Esta infraestructura computacional permitió automatizar la generación de más de cien submuestras con tamaños crecientes, realizar los cálculos inferenciales de forma repetida para cada método, y producir salidas gráficas como trayectorias de desviación estándar, diagramas de cajas, gráficos de amplitud y curvas de estabilidad de los intervalos, lo que garantizó solidez en la comparación de métodos bajo heterocedasticidad.

3.6 Técnicas para el procesamiento e interpretación de datos

El tratamiento de los datos en este estudio se apoya en una estrategia computacional sólida en torno al lenguaje de programación R, que permite la implementación de métodos estadísticos no paramétricos, técnicas de remuestreo y simulaciones mediante submuestras. En este caso correspondiente al momento transversal de los jets generados en colisiones de partículas, fue seleccionada debido a su elevada dispersión.

El flujo de procesamiento parte de la importación de los datos obtenidos del CERN Open Data Portal, previamente estructurados en formato .csv. A partir de esta base, se aplican funciones del paquete boot para la implementación de los métodos bootstrap percentil y BCa, mientras que los métodos basados en cuantiles empíricos y en transformación logarítmica ($\log-\chi^2$) son programados de forma personalizada mediante rutinas construidas en R, siguiendo la teoría estadística formal (Justus et al., 2024; Wilcox, 2020).

El proceso analítico incluye las siguientes etapas:

- Como toda base de datos se deben eliminar datos no atípicos o que presenten anomalías en el caso, estableciendo filtros en la variable analizar y exista una tendencia en la misma.
- Construcción de submuestras: Mediante la agrupación de datos aleatorios (submuestras) con tamaños diferentes ($n = 1000, 2000, 3000, 4000$) desde una base de datos originales hasta completar las 113 réplicas para evaluar el comportamiento de la desviación estándar en datos no paramétricos.
- Cálculo de intervalos de confianza:
 - Para el método $(\log - x^2)$, se aplica la transformación logarítmica al estimador de la desviación estándar muestral s^2 , obteniendo $\theta = \log(s^2)$. Luego, asumiendo aproximación normal de esta transformación, se construye el intervalo de confianza utilizando la fórmula (9).
 - Para el método de percentil empírico (percentil directo), Se generan s_k submuestras con reemplazo desde los datos originales. Para cada submuestra se calcula el estimador \hat{s} , y se ordena el conjunto de estos valores simulados. El intervalo de confianza se define entre los percentiles $\alpha/2$ y $1 - \alpha/2$. Aplicando finalmente la expresión matemática (14) antes mencionada. Este método captura la variabilidad de la muestra sin requerir supuestos de normalidad.
 - Para el bootstrap percentil, el intervalo de confianza se obtiene directamente a partir de los percentiles de la distribución bootstrap simulada, sin necesidad de aplicar factores de corrección por sesgo ni

aceleración. Este enfoque asume que la distribución del estimador, generada mediante remuestreos con reemplazo, es representativa de la variabilidad real de los datos. Tal como se describe en la literatura especializada (Efron & Hastie, 2016). Se calcula a través de la expresión matemática (16). Este método permite construir intervalos de confianza tomando simplemente los valores al percentil, ofreciendo una solución práctica y robusta en situaciones donde no se cumplen los supuestos tradicionales de normalidad o varianza constante.

- Indicadores de evaluación: Para comparar el rendimiento de los métodos de estimación de intervalos de confianza bajo condiciones de heterocedasticidad, se calcularon dos métricas estadísticas fundamentales.:
 - Amplitud del intervalo: La amplitud \bar{A} se define como la diferencia entre los extremos superior e inferior del intervalo de confianza generado en cada submuestra. Dada una submuestra, el ancho del intervalo se calcula a través de la formula (17). Este indicador refleja la precisión del método (menor amplitud sugiere una estimación más eficiente).
 - Error estándar promedio: Este indicado cuantifica la variabilidad de las estimaciones y complementa la interpretación de la amplitud del intervalo, siendo útil para determinar la robustez del método ante fluctuaciones en los datos, es medida mediante la expresión antes mencionada (18).

Esto permite interpretar visualmente la estabilidad, el sesgo y la precisión de cada método comparado, al analizar cómo se comportan sus intervalos de confianza en términos de variabilidad. En particular, la amplitud del intervalo refleja la precisión del método,

mientras que la dispersión de los valores estimados señala su estabilidad. Además, el error estándar promedio calculado internamente en la programación mediante código en R, constituye un insumo clave para determinar la amplitud de los intervalos generados, especialmente en métodos como el de transformación logarítmica, el bootstrap y los cuantiles empíricos.

Tabla 1

Paquetes y librerías para el análisis

Paquetes	Librería	Descripción
<code>install.packages("readxl")</code>	<code>library(readxl)</code>	Carga el archivo con extensión .xlsx
<code>install.packages("ggplot2")</code>	<code>library(ggplot2)</code>	Crea gráficos estadísticos.
<code>install.packages("systemfonts")</code>	<code>library(systemfonts)</code>	Utiliza un tipo de letra específico.
<code>install.packages("nortest")</code>	<code>library(nortest)</code>	Incorpora la prueba M.KS que no viene por defecto para la normalidad.

Nota. Los paquetes y librerías indicados corresponden al entorno estadístico **R**, utilizado en este estudio para el procesamiento, análisis gráfico y evaluación estadística de datos experimentales.

3.7 Población y muestra

En investigaciones con enfoque cuantitativo de alta precisión, como la estimación estadística bajo condiciones de heterocedasticidad, la definición rigurosa de la población y la muestra constituye un pilar fundamental.

3.7.1 Población

En esta investigación, la población se compone de eventos físicos registrados en el detector CMS (Compact Muon Solenoid) del Gran Colisionador de Hadrones (LHC), durante las colisiones protón-protón de 7 TeV de energía, específicamente en el periodo denominado Run2011A Jet AOD, publicado a través del repositorio CERN Open Data Portal.

Estos eventos constituyen un universo de datos reales de alta resolución provenientes de experimentación científica en física de partículas, y representan una fuente confiable y accesible para estudios estadísticos sobre magnitudes físicas. Cada evento contiene múltiples variables físicas reconstruidas mediante algoritmos sofisticados, dentro de las cuales se selecciona la variable momento transversal del jet (Jet_{pt}), una medida fundamental en análisis hadrónico, ya que refleja la componente del momento perpendicular al eje del haz en colisiones (Chatrchyan, 2017).

La población total está conformada por millones de registros individuales, organizados en estructuras tipo árbol en formato ROOT, los cuales se encuentran disponibles bajo licencias de uso libre para fines académicos y científicos (CERN Open Data Portal, 2021). Esta población presenta heterogeneidad inherente, tanto en la energía de los jets como en su dispersión estadística, lo cual permite explorar la eficacia de métodos inferenciales robustos bajo condiciones reales de heterocedasticidad.

Dada la magnitud y complejidad del conjunto poblacional, la recolección completa de la población no es operacionalmente viable. Por ello, se recurre al uso de técnicas de muestreo computacional controlado para extraer subconjuntos representativos que capturen la variabilidad y la estructura de dispersión de la población original (Justus et al., 2024).

3.7.2 Tamaño de la Muestra

La muestra utilizada en esta investigación consistió en un conjunto consolidado de 113000 observaciones extraídas del total de datos del CMS (Compact Muon Solenoid), seleccionadas tras un riguroso proceso de filtrado computacional que garantizó la calidad y validez estadística del subconjunto. Para ello, se eliminaron registros con valores faltantes,

duplicados o considerados atípicos extremos, aplicando criterios robustos como el rango intercuartílico, la inspección de densidades y el análisis gráfico.

El enfoque muestral adoptado responde a un diseño cuantitativo de tipo no probabilístico por conveniencia, pero con aleatorización interna mediante técnicas de remuestreo y submuestreo con reemplazo. Este diseño híbrido permite simular distintos escenarios poblacionales bajo condiciones de heterocedasticidad y variabilidad estructural, lo cual resulta apropiado en contextos donde la distribución real de los datos se aleja de la normalidad. De acuerdo con Wilcox (2020), este tipo de diseño es adecuado para comparar la cobertura y precisión de intervalos de confianza bajo condiciones no paramétricas, especialmente cuando el supuesto de homocedasticidad no se cumple.

Adicionalmente, la muestra fue estratificada en subconjuntos de tamaños controlados ($n = 1000, 2000, 3000, 4000$), a fin de observar empíricamente cómo varía la precisión de los intervalos construidos en función del tamaño muestral. El valor de 113000 observaciones se justifica metodológicamente como un umbral eficiente para estudios de simulación con enfoque no paramétrico, de acuerdo con recomendaciones recientes en investigación computacional intensiva (Ciucan-Rusu et al., 2022; Justus et al., 2024) quienes emplean muestras entre 30.000 y 70.000 observaciones al evaluar cobertura y eficiencia de intervalos bajo técnicas bootstrap y de cuantiles.

El carácter no paramétrico de la variable `Jet_pt`, validado previamente mediante la prueba de Kolmogorov–Smirnov con corrección de Lilliefors, justifica el uso de técnicas como el percentil directo, el bootstrap clásico y el método log-chi-cuadrado transformado para la comparación de la eficiencia y cobertura de los intervalos. Estas técnicas requieren

múltiples submuestras generadas a partir de la muestra base, las cuales fueron creadas en el entorno estadístico R, utilizando funciones específicas como `sample()`.

De acuerdo con el uso de remuestreo en submuestras para evaluar el comportamiento de los estimadores en presencia de heterocedasticidad es adecuado, y proporciona resultados robustos cuando la estructura de la población es compleja y de difícil modelación analítica. Esta estrategia también ha sido empleada en simulaciones con datos experimentales similares en el área de física computacional, validando la pertinencia de la metodología permitiendo un análisis riguroso y replicable bajo condiciones de variabilidad no constante.

3.8 Formulación de hipótesis estadísticas

Dado el carácter comparativo y experimental del estudio, orientado a evaluar la precisión y eficiencia de diferentes métodos estadísticos para la construcción de intervalos de confianza bajo condiciones de heterocedasticidad, se plantea formalmente un conjunto de hipótesis que permitirán validar si existen diferencias significativas entre los tres métodos analizados: la transformación logarítmica ($\log-\chi^2$), los cuantiles empíricos y el método bootstrap.

La formulación de estas hipótesis busca establecer si alguno de los métodos supera a los demás en términos de cobertura empírica, precisión (amplitud del intervalo) y estabilidad (error estándar) al aplicarse sobre submuestras del momento transversal (`Jet_pt`) obtenidas de datos experimentales del detector CMS.

Hipótesis general:

- H_0 : No existen diferencias estadísticamente significativas entre los métodos de estimación ($\log-\chi^2$, cuantiles empíricos, bootstrap percentil) respecto a los indicadores de cobertura, amplitud del intervalo y error estándar.

- H_1 : Existen diferencias estadísticamente significativas entre los métodos de estimación respecto a los indicadores mencionados.

3.9 Tratamiento estadístico

El análisis estadístico de la presente investigación requirió la necesidad de establecer criterios acordes a los métodos establecidos aplicados a la variable *Jet_pt*. Uno de los estadísticos considerados fue la desviación estándar muestral (s), utilizada para estimar la desviación estándar submuestral (\hat{S}) en base a la variación de la muestra (n), es decir mediante la aplicación de submuestras (S_k). La idea principal consistió en verificar el comportamiento de los intervalos de confianza (*IC*) construidos para cada valor estimado de \hat{S} , en función de los métodos empleados para el proceso de estimación.

Al trabajar con una variable continua, fue indispensable explorar los datos con el fin de verificar uno de los supuestos fundamentales de la estadística paramétrica: la normalidad de los datos, es decir, comprobar si los datos están distribuidos normalmente. Como primer paso se utilizó un gráfico de cuantiles (*Q-Q plot*) para comparar de forma visual la distribución del conjunto de datos muestrales con una distribución normal teórica. A continuación, se utilizó la prueba de Kolmogorov–Smirnov con corrección de Lilliefors (M.KS) para evaluar la normalidad de la distribución muestral de la variable *Jet_pt*. Esta prueba es adecuada cuando los parámetros poblacionales (media y desviación estándar) se estiman a partir de la misma muestra, y no se conocen a priori. El contraste se aplicó bajo un nivel de significancia de $\alpha = 0.05$, es decir, con un nivel de confianza del 95 %. Un valor de $p < 0.05$ permitió rechazar la hipótesis nula de normalidad, confirmando la necesidad de métodos no paramétricos para el análisis inferencial posterior, recomendada cuando los parámetros poblacionales son desconocidos y se estiman a partir de la muestra (Sayili &

Gunver, 2025). Esta prueba se basa en el estadístico D' , se aplica cuando $n > 50$, se tiene la media aritmética muestral (\bar{x}) y s se estiman a partir de los datos:

$$D' = \sup_x \left| F_n(x) - \Phi \left(\frac{x-\bar{x}}{s} \right) \right| \quad (20)$$

donde $F_n(x)$ es la función de distribución empírica, Φ es la función de distribución acumulativa normal estándar, \bar{x} es la media muestral y s es la desviación estándar muestral corregida. Esta prueba permite validar si el conjunto de datos puede considerarse normalmente distribuido; en este estudio, los resultados indicaron que los datos no siguen una distribución normal, lo que justifica el uso de métodos no paramétricos.

La prueba M.KS (denotado como D' , estadístico de prueba) consiste en una versión modificada de la prueba KS, la diferencia hace énfasis en que no asume que media y desviación estándar poblacional sean conocidas, más bien son estimados a partir de un conjunto de datos, es decir de una muestra. En aspectos de sensibilidad la prueba M.KS es más precisa para la normalidad de datos con parámetros desconocidos.

$$M.KS = \max(D'^+, D'^-) \quad (21)$$

$$D'^+ = \max \left| F_n(x) - \Phi \left(\frac{x-\mu}{s} \right) \right| \quad (22)$$

$$D'^- = \max \left| \Phi \left(\frac{x-\mu}{s} \right) - F_n(x) \right| \quad (23)$$

Donde $F_n(x)$ es la función de distribución empírica; Φ es la función de distribución acumulativa de la distribución normal estándar; $\mu = \bar{x}$ es la media muestral y s es la desviación estándar muestral corregida con $n-1$. Además, $s = \sqrt{\sum (x_i - \bar{x})^2 / (n-1)}$ siendo n el tamaño de la muestra. Los datos de la muestra deben estar ordenados de manera

creciente, además, esta prueba genera mayor potencia estadística al utilizar muestras de gran tamaño.

Al analizar previamente la normalidad de los datos, se definieron los métodos estadísticos para realizar las estimaciones. Los métodos empleados fueron la transformación logarítmica, los cuantiles empíricos y bootstrap, que son apropiados para el análisis de datos no paramétricos.

Para el análisis del comportamiento de s mediante la variación de la muestra, se definió el conjunto muestral $X = \{x_1, x_2, x_3, \dots, x_n\}$ donde x representa la observación individual del espacio muestral. A partir del conjunto completo de $n = 113000$ observaciones validas, generaron submuestras crecientes de tamaño s_k , incrementadas de 1000 en 1000 observaciones, es decir $s_k = \{x_1, x_2, x_3, \dots, x_{k>1000}\}$, para $k = 1, 2, 3, \dots, 113$. Este esquema permite analizar cómo evoluciona la estimación de la desviación estándar muestral a medida que se incrementa el tamaño de la muestra, simulando condiciones de variabilidad creciente. Donde s_k representa el tamaño de la submuestra. La forma matricial para s_k se detalla en la expresión:

$$S_k = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_{1000} \\ x_1 & x_2 & x_3 & \cdots & x_{2000} \\ x_1 & x_2 & x_3 & \cdots & x_{3000} \\ \vdots & \vdots & \vdots & & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n \end{pmatrix} \quad (24)$$

Para cada submuestra se obtuvo $\hat{s} = s = \{s_1, s_2, s_3, \dots, s_n\}$ los mismos que fueron comparados con s de la muestra total.

Para la ejecución de los métodos cuantiles empíricos y bootstrap, se utilizaron submuestras internas de hasta 1000 observaciones. Este procedimiento fue definido

previamente con la finalidad de minimizar el costo computacional que ocasiona el remuestreo, además de reducir el tiempo de ejecución y sobre todo evitar la sobrecarga de memoria.

Los parámetros analizar en la precisión de los intervalos de confianza se tomó como métrica el IC (ancho del intervalo) y el error estándar determinados por los métodos empleados.

Mediante la aplicación de cada método para estimar el *IC* se obtuvo el *Ancho del IC* = $\{a_1, a_2, a_3, \dots, a_{s_k}\}$ donde *a* representa el ancho del *IC* para cada S_k hasta completar la muestra total, es decir las observaciones de muestra ($n = k > 1000$).

Capítulo 4

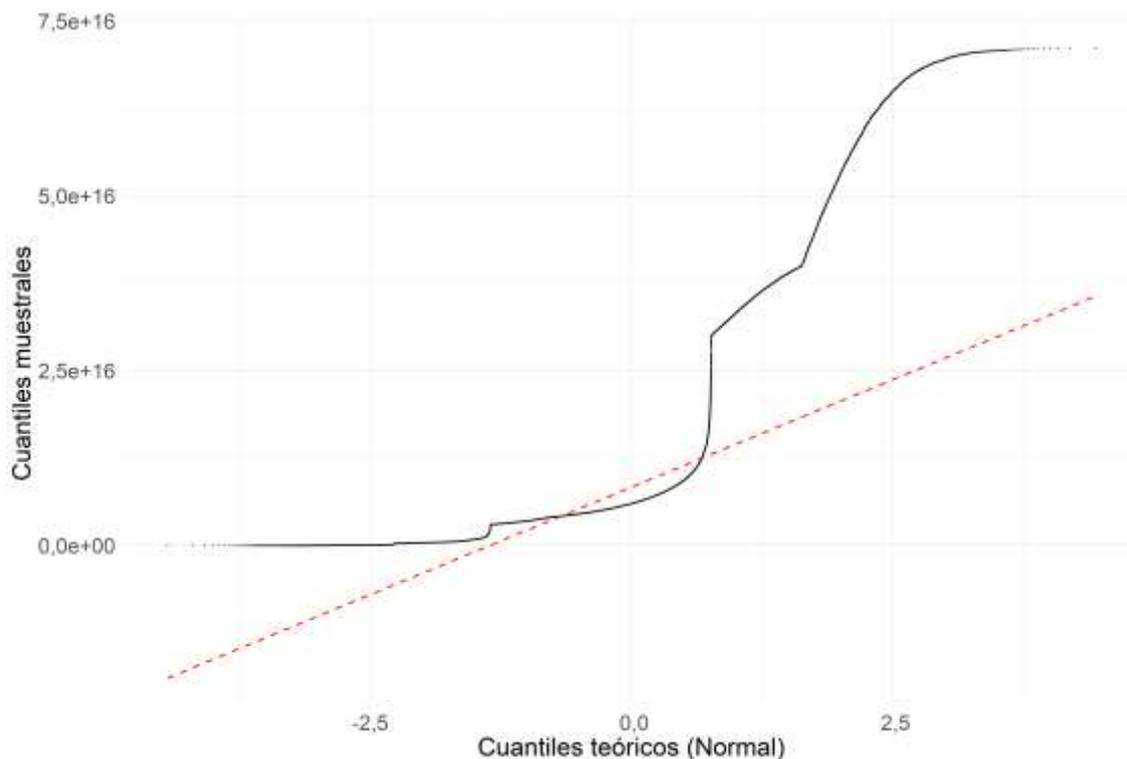
Análisis y Discusión de los resultados

4.1 Análisis descriptivo de los resultados

La variable principal de esta investigación es el momento transversal del jet (Jet_{pt}), correspondiente a eventos registrados por el detector CMS en colisiones protón-protón. Para evaluar su comportamiento estadístico inicial, se realizó un análisis exploratorio con énfasis en la verificación de supuestos de normalidad, dado que estos condicionan la elección de los métodos inferenciales a emplear en los capítulos siguientes.

Figura 1

Normalidad de los datos analizados por el detector CMS



Nota. La Figura 1 muestra un gráfico Q-Q del momento transversal del jet (Jet_{pt}) en unidades de GeV/c, obtenido de los datos experimentales del CMS (CERN, 2021). La fuerte desviación respecto

a la línea de referencia (en rojo) indica un claro alejamiento de la normalidad, lo que justifica el uso de métodos no paramétricos para la estimación estadística.

En la Figura 1 se observa el análisis de relación entre los cuantiles muestrales y los cuantiles teóricos basado a una distribución normal. Se puede determinar una desviación pronunciada de los datos analizados en comparación a la línea de referencia (línea roja discontinua), especialmente en los extremos. Demostrando así una asimetría en los datos (no se distribuyen normalmente y se encuentran altamente sesgados).

El resultado fue confirmado con la prueba de Kolmogorov–Smirnov modificada por Lilliefors, cuyo estadístico fue $D' = 0,29127$ y el valor de significancia reportado fue de $p < 2,2 \times 10^{-16}$, lo que lleva a rechazar contundentemente la hipótesis nula de normalidad al nivel de significancia de $\alpha = 0,05$.

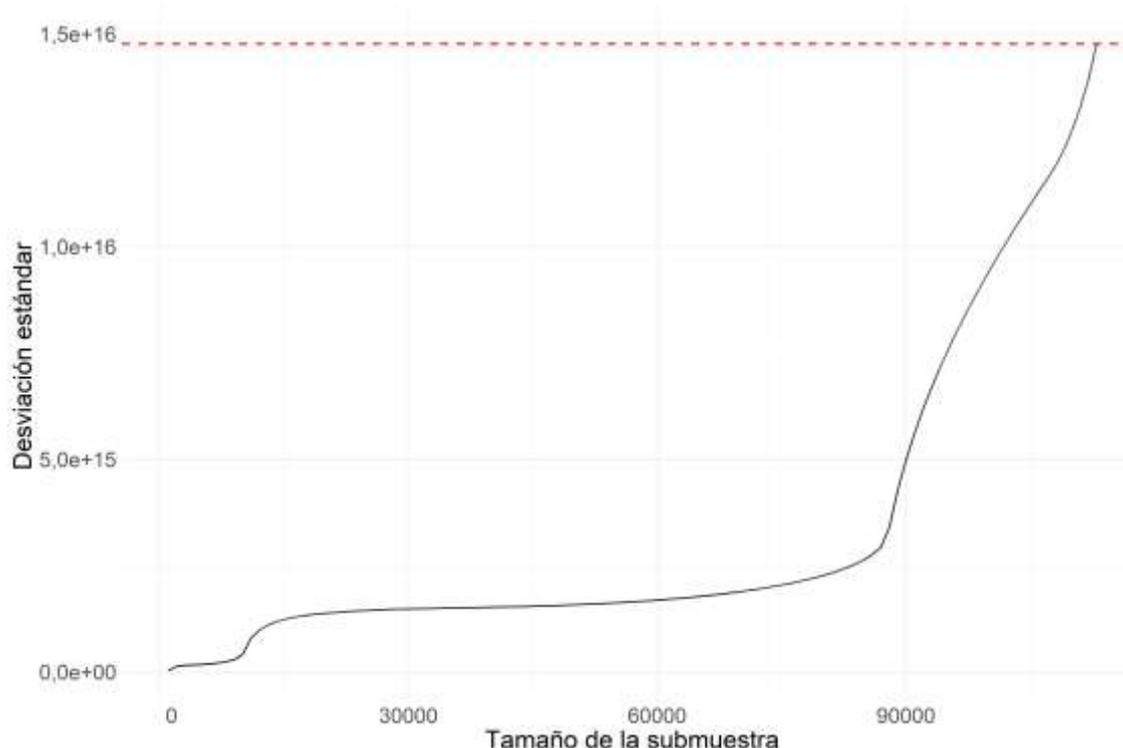
El resultado concuerda con investigaciones que señalan sesgos positivos en las distribuciones asociadas a eventos de alta energía dentro del ámbito de la física de partículas. Los datos obtenidos en la investigación afirman la elección de utilizar metodologías de análisis no paramétricas en estimación de intervalos de confianza en la interpretación de precisión y no al uso de sistemas de distribución normal.

- **Desviación estándar para cada submuestra**

A partir del conjunto de datos total se extrajeron submuestras crecientes para observar el comportamiento de la desviación estándar estimada y el ancho de los intervalos de confianza en cada caso.

Figura 2

Desviación estándar para cada submuestra.



Nota. La gráfica representa el comportamiento de la desviación estándar del momento transversal del jet (Jet_pt) en función del tamaño de la submuestra ($n = 1000$ a 113000), medido en unidades de GeV/c.

En la figura 2 se puede observar el análisis de la relación entre el tamaño de la submuestra de la toma total de datos y la desviación muestral (\hat{s}). La línea segmentada de color rojo es la desviación estándar total de la muestra original ($s_{total} = 1,5 \times 10^{16}$). También se puede determinar el comportamiento asimétrico y el ajuste de tendencia no lineal en la variabilidad de estimador, especialmente al superar los 90 000.

Este comportamiento respalda la existencia de heterocedasticidad, ya que la variación del estimador no se mantiene constante a lo largo de los diferentes tamaños de submuestra. El patrón observado, que incluye una fase inicial de estabilización seguida de un aumento abrupto de la dispersión, sugiere una estructura de varianza segmentada o no homogénea. Este tipo de comportamiento es coherente con fenómenos reportados en física

de partículas, donde las energías asociadas a ciertos eventos tienden a concentrarse en regiones extremas del dominio, generando mayor variabilidad estructural en los datos (Khachatryan et al., 2017).

La información proporcionada por esta figura resulta fundamental para establecer que los métodos de estimación seleccionados deben poseer la capacidad de ajustarse a una dinámica de dispersión creciente.

4.2 Análisis comparativo de los métodos de estimación

El objetivo central de esta sección es evaluar y comparar el comportamiento de los tres métodos estadísticos no paramétricos aplicados para la estimación de intervalos de confianza de la desviación estándar del momento transversal (Jet_pt) en condiciones de heterocedasticidad.

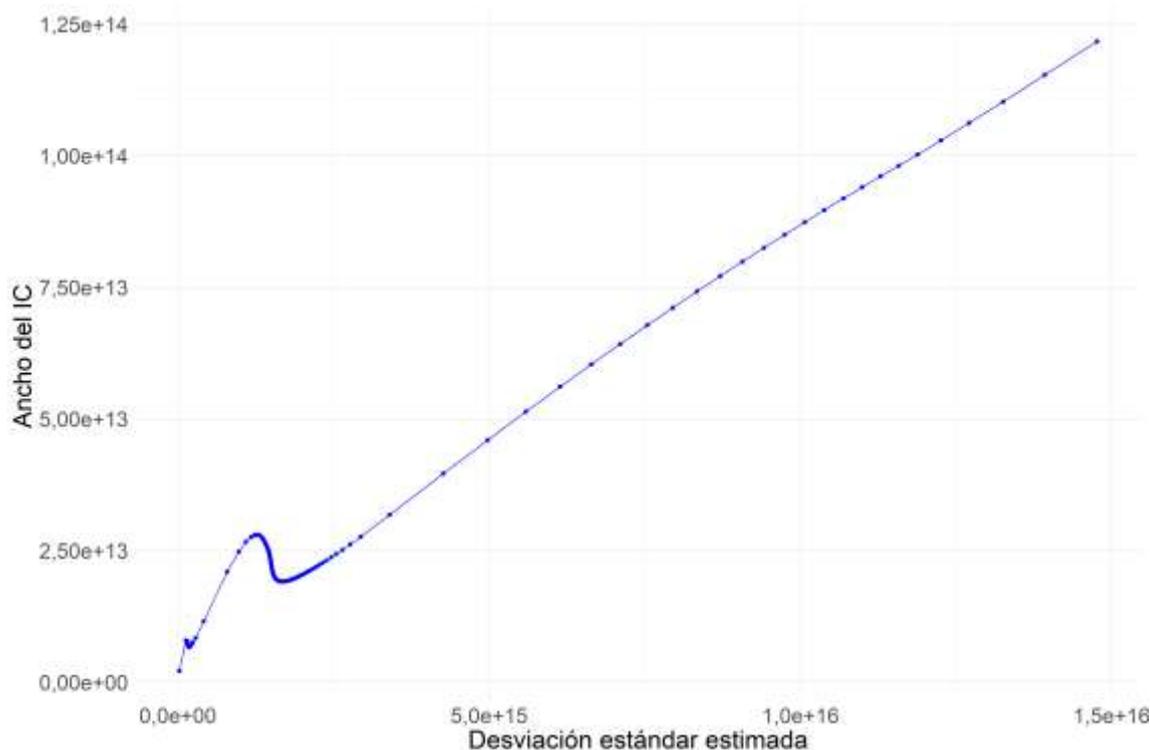
Cada enfoque fue aplicado sobre submuestras extraídas del conjunto de datos del CMS, permitiendo observar la evolución de la precisión y la eficiencia del intervalo en función del tamaño muestral y la dispersión.

4.2.1 Ancho del IC según la desviación estándar estimada - Método de transformación logarítmica ($\log - x^2$)

En la figura 3 se puede determinar la fluencia de la desviación estándar en concordancia al ancho de los intervalos de confianza mediante la técnica de transformación logarítmica para datos no normales y con heterocedasticidad. Esta técnica trabaja a través de logaritmos naturales en la desviación estándar previamente del cálculo de los intervalos de confianza.

Figura 3

Ancho del IC según la desviación estándar estimada - Método de transformación logarítmica



Nota. Tanto la desviación estándar como el ancho del intervalo de confianza están expresados en unidades de GeV/c, correspondientes al momento transversal del jet (Jet_{pt}) medido en el experimento CMS del LHC.

En el sector izquierdo del gráfico, se encuentra asociado a los valores de las desviaciones estándar siendo estos de baja magnitud, considerando una gran variabilidad de la heterogeneidad en la amplitud de los intervalos. Esta característica denota una mayor inestabilidad en la estimación cuando el tamaño de las submuestras es reducido o los valores de dispersión son notablemente bajos. En esta región, incluso leves oscilaciones en los datos pueden ocasionar fluctuaciones significativas en el intervalo, lo cual pone de manifiesto la susceptibilidad del método ante escenarios de escasa información muestral o valores atípicos.

Cuando la dispersión que prevemos (la desviación estándar) se hace más grande, la forma en que el ancho del margen de error crece se vuelve más gradual y fácil de anticipar.

En esta parte central del gráfico, parece que las cosas se estabilizan estadísticamente; nuestras estimaciones se vuelven más fiables y los márgenes de error se ajustan mejor. Aquí es donde la técnica logarítmica realmente brilla, dando resultados sólidos cuando la variabilidad no es ni muy baja ni muy alta.

Pero, al final del gráfico, se observa que el margen de confianza se ensancha significativamente. Esto no es sorprendente en absoluto ya que estadísticamente, es lógico que, si los datos están más esparcidos existan duda en la estimación.

4.2.2 Ancho del IC según la desviación estándar estimada - Método de cuantiles empíricos

La Figura 4 muestra el comportamiento del ancho de los intervalos de confianza (IC) en función de la desviación estándar estimada mediante el método de cuantiles empíricos (percentil directo). Este método, al no requerir supuestos de normalidad ni varianza constante, se adapta bien a contextos no paramétricos como el presente, en el que la distribución de la variable `Jet_pt` del CMS presenta marcada heterocedasticidad.

El gráfico demuestra que, en niveles donde la desviación típica exhibe valores mínimos, la magnitud del intervalo de confianza se mantiene comparativamente acotada, lo que implica una mayor exactitud en las valoraciones primarias. Sin embargo, a medida que se acrecienta la heterogeneidad entre las submuestras (5×10^{15}), se aprecia una expansión notable en la extensión del intervalo. Esta tendencia denota un incremento en la indeterminación vinculada a la valoración.

Aparte de esto, se observa un período de estabilidad relativa en los valores promedio de la desviación estándar. En esta etapa, el tamaño de los rangos varía, pero sin mostrar un incremento constante. Esta situación dentro de un cierto límite, el método de cuantiles

esa zona, lo que implica que está cubriendo los valores reales de forma razonable. Sin embargo, a medida que la desviación estándar crece, el intervalo se ensancha, lo que podría indicar que el método pierde fuerza cuando hay dispersión en los datos.

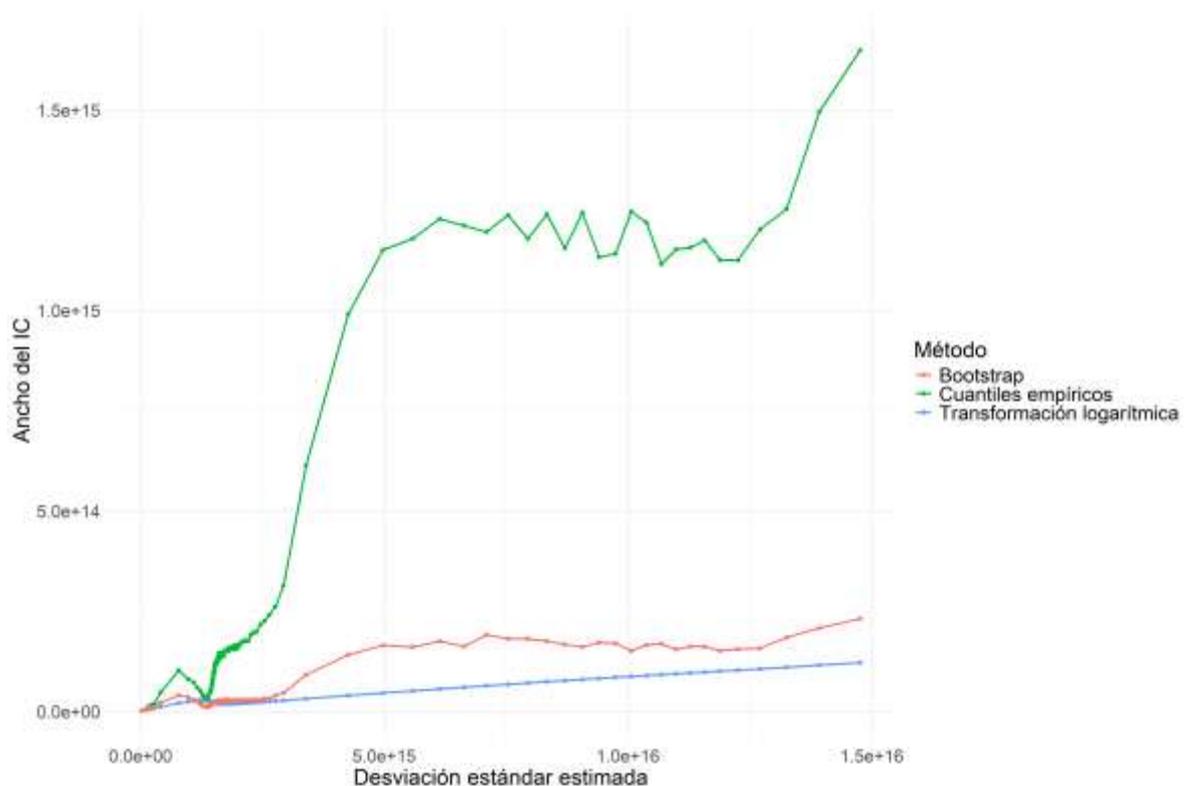
En resumen, lo visto muestra que el método bootstrap es una herramienta valiosa en muchos casos, pero debemos ser cuidadosos al usarla cuando la variabilidad es alta y las distribuciones son complicadas.

- Relación entre desviación estándar y ancho del IC.

Esta visualización permite realizar una comparación directa en cuanto a la precisión y estabilidad de cada técnica a lo largo del espectro de variabilidad observada en los datos.

Figura 6

Relación entre desviación estándar estimada y ancho del IC



Nota. Tanto la desviación estándar estimada como el ancho del intervalo de confianza están expresados en GeV/c, unidad correspondiente al momento transversal del jet (Jet_pt) registrado por el experimento CMS del LHC.

La gráfica número 6 ilustra cómo se vinculan la dispersión calculada (desviación estándar) y la amplitud del rango de valores creíbles (intervalo de confianza) para los tres enfoques analizados: remuestreo bootstrap (línea roja), percentiles basados en los datos observados (línea verde) y la técnica de transformación logarítmica (línea azul).

El método de los cuantiles empíricos, hay que decirlo, es el que más sufre cuando la variabilidad de los datos crece. Al principio puede parecer prometedor con intervalos de confianza ajustados, pero ante mucha dispersión, se dispara y genera intervalos muy amplios, revelando sus limitaciones en contextos donde la variabilidad no es uniforme. En los gráficos, se puede ver cómo se comporta de forma errática, con subidas y bajadas bruscas, lo que indica problemas para estimar correctamente la cobertura real. Esta fragilidad es especialmente evidente cuando los datos no están distribuidos de forma simétrica o presentan valores extremos atípicos.

El método bootstrap se observa que su comportamiento es más uniforme, aunque aumenta cuando los datos se dispersan ($>$ desviación estándar). Esto sugiere que se adapta bien a los problemas de sesgo y variabilidad en los datos. Siendo una buena opción cuando no se puede dar por sentado que los datos se comportan de forma "normal" ni que tienen una varianza constante.

Por otro lado, la conversión logarítmica tiende a darnos los rangos de valores más ajustados, casi siempre superando a los otros métodos analizados. Lo bueno es que se comporta de manera consistente y bastante lineal a medida que la variabilidad de los datos aumenta, lo que sugiere que es una herramienta estadística muy efectiva para lograr precisión.

En resumen, si nos basamos en la gráfica, podemos decir que la transformación logarítmica es la que mejor funciona. Los intervalos de confianza que crea son más precisos y consistentes, incluso cuando los datos varían mucho. El método bootstrap se porta bastante bien y es adaptable, especialmente útil cuando no podemos asumir que los datos siguen una distribución normal. Por otro lado, el método de cuantiles empíricos es menos eficiente y los intervalos que genera cambian mucho, sobre todo cuando la variabilidad no es uniforme.

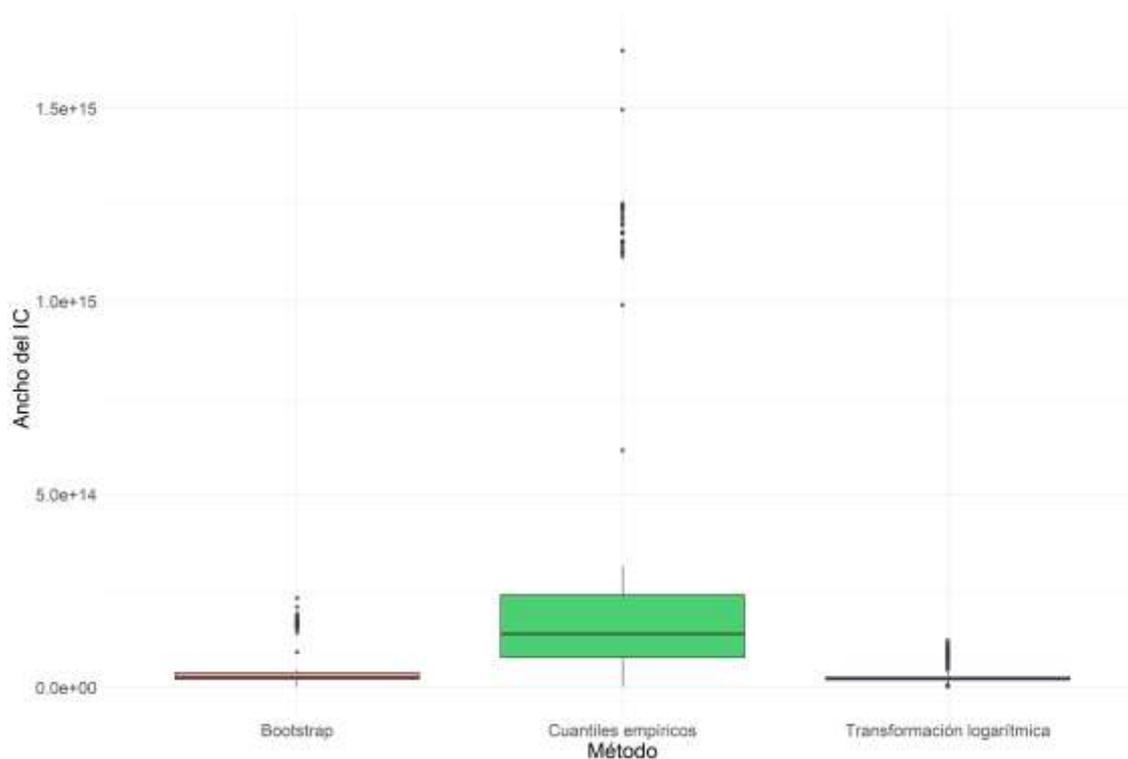
Estos hallazgos son coherentes con las propiedades teóricas esperadas de cada técnica y respaldan la elección de la transformación logarítmica cuando se prioriza la precisión en la estimación, y del bootstrap como alternativa robusta ante distribuciones complejas y no normales.

- **Diagrama de cajas del ancho del IC por método**

El diagrama de cajas nos da una imagen clara de entender de cómo varían los anchos de los intervalos de confianza calculados con los tres enfoques que probamos: bootstrap, cuantiles empíricos y la transformación logarítmica.

Figura 7

Diagrama de cajas del ancho del IC por método



Nota. El ancho del intervalo de confianza está expresado en unidades de GeV/c , correspondientes al momento transversal del jet (Jet_{pt}) registrado en el experimento CMS del LHC.

Al examinar los intervalos de confianza calculados con el método de los cuantiles empíricos, salta a la vista una variabilidad importante en su amplitud. La mediana de estos anchos es alta y, además, se observan valores que se desvían mucho de la norma (valores atípicos) en la parte superior, lo que sugiere que este método no se comporta de manera fiable cuando hay heterocedasticidad (o sea, cuando la variabilidad no es constante).

Por otro lado, el método bootstrap tiene un panorama diferente. Su mediana es notablemente inferior y el rango intercuartílico es ajustado, lo que apunta a una mayor constancia y exactitud en la estimación. Además, tiene pocos valores extremos que distorsionen la imagen.

Tabla 2

Resumen de resultados de intervalos de confianza (IC) para la desviación estándar muestral usando los métodos estudiados

Mues- tra	Método	Tamaño muestra	Desviación estimada	IC inferior	IC superior	Ancho IC
1	Transformación logarítmica	1000	2,438e+13	2,334e+13	2,548e+13	2,139e+12
1	Cuantiles empíricos	1000	2,438e+13	2,350e+13	2,521e+13	1,705e+12
1	Bootstrap	1000	2,438e+13	2,332e+13	2,555e+13	2,227e+12
2	Transformación logarítmica	2000	1,281e+14	1,241e+14	1,321e+14	7,940e+12
2	Cuantiles empíricos	2000	1,281e+14	1,260e+14	1,299e+14	3,892e+12
2	Bootstrap	2000	1,281e+14	1,251e+14	1,303e+14	5,178e+12
3	Transformación logarítmica	3000	1,541e+14	1,503e+14	1,581e+14	7,803e+12
3	Cuantiles empíricos	3000	1,541e+14	1,521e+14	1,561e+14	3,918e+12
3	Bootstrap	3000	1,541e+14	1,522e+14	1,562e+14	4,005e+12
4	Transformación logarítmica	4000	1,638e+14	1,602e+14	1,674e+14	7,179e+12
4	Cuantiles empíricos	4000	1,638e+14	1,601e+14	1,673e+14	7,226e+12
4	Bootstrap	4000	1,638e+14	1,611e+14	1,666e+14	5,487e+12
5	Transformación logarítmica	5000	1,722e+14	1,689e+14	1,757e+14	6,753e+12
5	Cuantiles empíricos	5000	1,722e+14	1,680e+14	1,771e+14	9,148e+12
5	Bootstrap	5000	1,722e+14	1,685e+14	1,752e+14	6,764e+12
6	Transformación logarítmica	6000	1,852e+14	1,819e+14	1,886e+14	6,629e+12
6	Cuantiles empíricos	6000	1,852e+14	1,792e+14	1,910e+14	1,180e+13
6	Bootstrap	6000	1,852e+14	1,813e+14	1,893e+14	8,056e+12
7	Transformación logarítmica	7000	2,085e+14	2,050e+14	2,120e+14	6,908e+12
7	Cuantiles empíricos	7000	2,085e+14	2,011e+14	2,156e+14	1,451e+13
7	Bootstrap	7000	2,085e+14	2,037e+14	2,130e+14	9,310e+12
...
113	Transformación logarítmica	113000	1,477e+16	1,471e+16	1,483e+16	1,218e+14
113	Cuantiles empíricos	113000	1,477e+16	1,390e+16	1,555e+16	1,649e+15
113	Bootstrap	113000	1,477e+16	1,464e+16	1,487e+16	2,316e+14

Nota. Las cantidades expresadas en la tabla (desviación estándar, límites del intervalo de confianza y amplitud) corresponden al momento transversal del jet (Jet_pt), medido en unidades de GeV/c, registrado por el experimento CMS del LHC del CERN durante colisiones protón-protón a $\sqrt{s} = 7$ TeV según la base de datos.

El análisis comparativo de los tres métodos empleados para la estimación de intervalos de confianza de la desviación estándar muestral (transformación logarítmica, bootstrap y cuantiles empíricos) evidencia diferencias sustanciales en cuanto a precisión, amplitud de los intervalos y estabilidad del error estándar. Estas diferencias se mantienen tanto en submuestras pequeñas como grandes, lo que permite valorar la robustez

metodológica de cada enfoque ante condiciones de heterocedasticidad y desviaciones de la normalidad.

Tabla 3

Matriz comparativa final entre los tres métodos con base en los indicadores de eficiencia en las submuestras.

Método	Ancho promedio IC	Error estándar
Bootstrap	2,316e+14	3,594e+15
Cuantiles empíricos	1,649e+15	3,586e+15
Transformación logarítmica	1,218e+14	2,660e+14

Nota. Unidades expresadas en GeV/c. Ver descripción detallada en la tabla 2.

Después de probar varios caminos, se concluye que transformar los datos con logaritmos y usar la distribución chi-cuadrado fue la mejor opción para estimar con confianza la variabilidad de nuestra muestra. En las 113 submuestras analizadas nos dio los rangos más precisos, promediando apenas $1,218 \times 10^{14}$. Las estimaciones fueron consistentes con un margen de error de apenas $2,660 \times 10^{14}$, lo que dice mucho de su precisión.

El método bootstrap presenta un desempeño intermedio, equilibrando precisión y flexibilidad. Aunque sus intervalos son más amplios que los del método logarítmico (ancho promedio de $2,316 \times 10^{14}$), su error estándar de $3,594 \times 10^{15}$ revela un comportamiento más disperso en la estimación de la desviación estándar, especialmente en tamaños de muestra pequeños. No obstante, su principal fortaleza radica en su adaptabilidad a distribuciones no normales y su facilidad de implementación computacional. En contextos donde se desconoce la forma real de la distribución o existe alta variabilidad, el bootstrap se presenta como una alternativa robusta con buena cobertura empírica.

El método que usa los cuantiles empíricos en los datos observados genera intervalos más amplios, de hasta $1,649 \times 10^{15}$ y su error estándar es igualmente a $3,586 \times 10^{15}$. Esto muestra que no es tan "eficiente" como los otros métodos en presiones de intervalos. Si bien es bueno para detectar desbalances en los datos (sesgos y asimetrías), es muy sensible a qué tan dispersos están esos datos. Así que, si se trabaja con intervalos de confianza pequeños y estables, este método podría dar resultados poco fiables. Es más útil cuando los datos tienen datos más extensos para analizar.

En resumen, los hallazgos revelan que la selección del método ideal está íntimamente ligada a la naturaleza de los datos y a los objetivos del análisis. Si los datos se desvían de una distribución normal, el bootstrap se presenta como una alternativa sólida. Aunque el enfoque de cuantiles empíricos no es el más óptimo, puede ser valioso para el reconocimiento de los datos o como un adjunto a técnicas más elaboradas. Si la precisión de los resultados es primordial, la transformación logarítmica emerge como la opción superior

4.3 Discusión de los resultados

Los hallazgos obtenidos a partir de la comparación de métodos para estimar intervalos de confianza (IC) de la desviación estándar, en condiciones de heterocedasticidad, confirman importantes diferencias en desempeño entre las tres técnicas analizadas: transformación logarítmica ($\log - x^2$), cuantiles empíricos y remuestreo bootstrap. En particular, se observa que la transformación logarítmica proporciona los IC más estrechos y estables, incluso frente a valores extremos en la desviación estándar estimada, como ocurre en las primeras submuestras donde $\hat{s} > 1,5 \times 10^{16}$. Esto es consistente con lo reportado en la literatura, donde este enfoque se recomienda por su corrección de sesgo y eficiencia bajo distribuciones no normales (Justus et al., 2024; Wilcox, 2020).

El método bootstrap se fiable y consistente al analizar diferentes porciones de los datos. Una de sus ventajas es que los intervalos de confianza que genera no varían mucho, y generalmente aciertan a cubrir el valor real que intentamos estimar. Esto concuerda con lo que han encontrado otros investigadores recientemente, que destacan lo bien que funciona, incluso cuando los datos están muy dispersos o no cumplen supuestos estrictos como la normalidad. Eso sí, cuando trabajamos con pocos datos, el bootstrap puede mostrar pequeñas variaciones, lo que indica que se ve afectado tanto por la "forma" interna de los datos como por la cantidad de veces que lo volvemos a muestrear.

Por otro lado, el método de los cuantiles empíricos tiende a ser más volátil en sus resultados. Genera intervalos de confianza más grandes, y cuanto más se dispersan los datos, más grandes se hacen estos intervalos. Esta inestabilidad no corregir errores a satisfacción cuando la variabilidad de los datos es muy alta. Así lo advierte Veljkovic (2024), quien señala que, aunque este método puede ser útil para tratar distribuciones asimétricas, su eficiencia disminuye. A través del test de Lilliefors como el gráfico Q-Q nos muestran sin lugar a dudas que los datos de Jet_pt no siguen una distribución normal.

Según lo observado en la práctica en comparación a los fundamentos teóricos, se determina que transformar los datos usando el método de logaritmos percentiles es la mejor opción de calcular intervalos de confianza para la desviación estándar cuando hay mucha variabilidad en datos no paramétricos. El método bootstrap percentil es una alternativa buena y adaptable, sobre todo cuando no podemos asumir que los datos son normales. El método que se basa en observar los valores más comunes (cuantiles empíricos) funciona bien en ciertos casos, pero hay que tener cuidado al interpretarlo, especialmente cuando la variabilidad de los datos no es constante. Con esta evidencia, se puede ayudar para futuros estudios que analicen problemas estadísticos donde hay mucha variabilidad, como suele

pasar al analizar datos experimentales en física de partículas. Los hallazgos de este estudio nos dan la oportunidad de poner a prueba, con datos reales, las ideas principales que se plantearon al principio.

En resumen, los datos que obtenidos nos conduce a rechazar de la hipótesis nula general (H_0), apoyando la hipótesis alternativa (H_1) que sostiene la existencia de diferencias significativas entre los métodos estadísticos comparados.

Capítulo 5

Marco Propositivo

5.1 Planificación de la actividad preventiva

Los hallazgos obtenidos a lo largo de esta investigación han evidenciado las limitaciones de los métodos tradicionales de estimación de intervalos de confianza bajo condiciones de heterocedasticidad, particularmente en contextos donde se analizan variables experimentales como el momento transversal (Jet_{pt}) de los jets de partículas. Ante esta problemática metodológica, se propone una estrategia preventiva basada en la selección óptima de técnicas de estimación adaptadas a la naturaleza no paramétrica de los datos.

La meta principal es crear una guía paso a paso, fácil de entender, para que cualquiera pueda calcular rangos de confianza sólidos, incluso si tus datos no siguen una distribución normal o si la variabilidad es diferente a lo largo del conjunto de datos. Pensar que esta guía puede facilitar a evaluación de rangos de confianza cuando se trabaja con datos no paramétricos. Explorando diferentes caminos, como usar transformaciones logarítmicas, basarnos en los propios datos para construir intervalos (intervalos empíricos) o emplear el método de remuestreo bootstrap percentil. Establecer vivencias para elegir la mejor ruta, dependiendo de cómo se comporten los datos. Obtener rangos de confianza que sean exactos (lo más estrechos posible), seguros (con una baja probabilidad de error) y efectivos (que "atrapen" el valor real la mayor parte del tiempo). Básicamente, una herramienta que te ayude a confiar en tus resultados.

Para evitar problemas más adelante, se sugiere usar estas técnicas al principio del análisis de datos experimentales con muchos datos como en física de partículas, especialmente cuando no sigue un patrón predefinido. De esta manera, se puede detectar la

variabilidad no constante (heterocedasticidad) antes de llegar a conclusiones definitivas. Saber esto desde el principio permite elegir el método de estimación más adecuado y evitar aplicar a ciegas técnicas tradicionales, como el intervalo de Wald, que no tienen gran precisión cuando los datos no siguen una distribución normal o la varianza cambia.

Esta propuesta tiene también un valor académico y profesional. Desde el punto de vista formativo, permite fortalecer las competencias en estadística avanzada y análisis computacional en estudiantes e investigadores de ciencias exactas, proporcionando herramientas prácticas y replicables en entornos reales. En el plano institucional, su adopción puede mejorar significativamente la calidad del análisis de datos científicos en laboratorios, universidades y centros de investigación, promoviendo decisiones basadas en evidencias más confiables y ajustadas a la estructura de los datos observados.

Esta guía no se limita a este problema específico. Podemos usarla en muchos otros campos donde la variabilidad de los datos (heterocedasticidad) es algo común, como en economía, el clima, la biología con computadoras, o incluso para mejorar cómo funcionan las fábricas. Esto demuestra que esta idea es útil en diferentes áreas y para distintos profesionales. Si se empieza a usar estos métodos de manera constante, a la larga mejorará la forma en que hace estadística, animando a los investigadores a ser más cuidadosos, a pensar con más criterio y a ser conscientes de que los métodos tradicionales no siempre son suficientes cuando las cosas se complican.

Conclusiones

Se cumplió con el objetivo general al contrastar tres técnicas estadísticas no paramétricas para calcular rangos de confianza bajo la dispersión de datos (desviación estándar) cuando los datos no son uniformes (heterocedasticidad). Usando datos reales de la medición de momento transversal (Jet_pt) del experimento CMS en el CERN. Al estudiar los resultados, se encontraron diferencias notables en la precisión, el tamaño de los rangos y la consistencia de los métodos. Esto nos dio la evidencia necesaria para validar nuestra hipótesis alternativa (H_1), y rechazar la hipótesis nula (H_0).

Se confirmó la presencia de heterocedasticidad en los datos mediante la prueba de normalidad de Kolmogorov–Smirnov con corrección de Lilliefors, cuyo estadístico $D = 0,29127$ y valor $p < 2,2 \times 10^{-16}$ permitieron rechazar con contundencia la hipótesis de normalidad por tratarse de datos no paramétricos. Esta evidencia justificó el uso de los métodos planteados frente a modelos clásicos que suponen homocedasticidad, reforzando la necesidad de enfoques robustos ante variabilidad estructural.

La estrategia metodológica basada en submuestras progresivas de tamaño creciente ($n = 1000$ a 113000) permitió observar la evolución del comportamiento de los estimadores y la estabilización de los intervalos de confianza conforme aumentaba el tamaño muestral. Esta aproximación aportó evidencia empírica sólida sobre la eficiencia relativa de los métodos evaluados.

Con base en los resultados obtenidos, el método de transformación logarítmica se posicionó como el más preciso y eficiente, al presentar los intervalos de confianza más estrechos y el menor error estándar en todas las submuestras analizadas; específicamente, en la submuestra de tamaño 113000 alcanzó un ancho de intervalo de $1,218 \times 10^{-14}$ y un error

estándar notablemente bajo de $2,660 \times 10^{14}$, lo que confirma su robustez teórica incluso en contextos de alta dispersión. En comparación, el método bootstrap mostró un desempeño intermedio, caracterizado por un equilibrio aceptable entre precisión y adaptabilidad ante distribuciones desconocidas, con un ancho de intervalo de $2,316 \times 10^{14}$ y un error estándar de $3,594 \times 10^{15}$, lo cual evidencia su utilidad práctica pero también una mayor dispersión. Por otro lado, el método de cuantiles empíricos fue el menos eficiente estadísticamente, con los intervalos más amplios (ancho de $1,649 \times 10^{15}$) y el mayor error estándar ($3,586 \times 10^{15}$), reflejando una sensibilidad considerable frente a la heterogeneidad de los datos.

Los métodos que empleamos y su puesta en práctica en R resultaron ser, además de sólidos desde el punto de vista estadístico, notablemente prácticos. Lo mejor de todo es que este enfoque tiene un gran futuro: puede adaptarse fácilmente a otras áreas donde lidian con datos que varían mucho (heterocedasticidad) o que no siguen una distribución normal, como en estudios biológicos, análisis económicos o investigaciones de salud. En definitiva, el uso de estos métodos ayuda a mejorar la confianza en nuestras conclusiones estadísticas cuando nos enfrentamos a datos reales que son muy variables y con patrones complejos.

Recomendaciones

Desde el enfoque metodológico, se recomienda ampliar el análisis incorporando pruebas adicionales de bondad de ajuste y heterocedasticidad, como la prueba de Breusch Pagan o el test de White, con el fin de fortalecer el diagnóstico previo a la estimación de intervalos. Esto permitiría validar aún más la adecuación de los métodos no paramétricos aplicados en diferentes estructuras de datos.

Desde una perspectiva de educación avanzada, sería muy valioso integrar esta forma de comparar y calcular rangos de valores probables en cursos como estadística aplicada o inferencia no paramétrica a nivel de posgrado. Lo valioso de este enfoque es que transmite sobre cómo examinar datos reales no paramétricos. Esta manera de proceder no solo despierta el pensamiento crítico, sino que también afianza la conexión entre los fundamentos teóricos de la estadística y su aplicación práctica en escenarios complejos.

Para hacerlo más tangible, sería desarrollar una herramienta intuitiva e interactiva, tal vez utilizando RShiny u otras plataformas parecidas, para simplificar el uso de los tres métodos que hemos explorado. Esta herramienta sería un gran apoyo para investigadores que no son expertos en estadística, ayudándoles a tomar la mejor decisión cuando los datos no siguen una distribución normal y presentan gran variabilidad.

Finalmente, se sugiere replicar este estudio utilizando otras variables experimentales provenientes del CERN por su base datos, tales como la energía total del evento o el ángulo azimutal de los jets, con el propósito de evaluar la generalización de los resultados obtenidos.

Referencias Bibliográficas

- Asamblea Nacional del Ecuador. (2004). *Ley Orgánica de Transparencia y Acceso a la Información Pública (LOTAIP)*. <https://doi.org/https://www.defensoria.gob.ec>
- Asamblea Nacional del Ecuador. (2018). *Ley Orgánica de Educación Superior (LOES)*. <https://doi.org/https://www.loes.ec>
- CERN. (2021). *CERN Open Data Portal*. <https://doi.org/https://opendata.cern.ch>
- Chatrchyan, S. (2017). Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *Journal of Instrumentation*, 7(10), P10002. <https://doi.org/10.1088/1748-0221/7/10/P10002>
- Ciucan-Rusu, L., Vasile, V., Stefan, D., & Comes, C. (2022). Consumers Behavior Determinants on Online Local Market Platforms in COVID-19 Pandemic - A Probit Qualitative Analysis. *Mathematics*, 10(22), 4281. <https://doi.org/10.3390/math10224281>
- Código Orgánico de La Economía Social de Los Conocimientos, Creatividad e Innovación (2016). <https://doi.org/https://www.educacionsuperior.gob.ec>
- Convenio de Berna Para La Protección de Las Obras Literarias y Artísticas (1971). <https://doi.org/https://www.wipo.int/treaties/es/ip/berne/>
- Creative Commons. (2024). *About the licenses – Attribution 4.0 International (CC BY 4.0)*. <https://doi.org/https://creativecommons.org/licenses/by/4.0/>
- Duarte, J. (2021). Fast inference of deep neural networks in FPGAs for particle physics. *Journal of Instrumentation*, 16(10), P10003. <https://doi.org/10.1088/1748-0221/16/10/P10003>

- Efron, B., & Hastie, T. (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science* (Cambridge University Press, Ed.).
<https://doi.org/10.1017/CBO9781316576533>
- Fernández-Casal, R., Castillo-Páez, S., & Flores, M. (2024). A Nonparametric Bootstrap Method for Heteroscedastic Functional Data. *Journal of Agricultural, Biological and Environmental Statistics*, 29(1), 169–184. <https://doi.org/10.1007/s13253-023-00561-2>
- Hayes, A. F., & Cai, L. (2007). Heteroscedasticity effects in standard deviation confidence intervals. *Behavior Research Methods*, 39, 709–722.
<https://doi.org/10.3758/BF03192961>
- Hernández, R., Fernández, C., & Baptista, P. (2021). *Metodología de la investigación* (7th ed.). McGraw-Hill Education.
- Hirschauer, N., Grüner, S., & Mußhoff, O. (2021). *Fundamentals of Statistical Inference: What is the Meaning of Random Error*. Springer. <https://doi.org/10.1007/978-3-030-99091-6>
- Jung, K., Lee, J., Gupta, V., & Cho, G. (2019). Comparison of bootstrap confidence interval methods for GSCA using a Monte Carlo simulation. *Frontiers in Psychology*, 10, 2215. <https://doi.org/10.3389/fpsyg.2019.02215>
- Justus, V. L., Rodrigues, V. B., & Sousa, A. R. (2024). Bootstrap confidence intervals: A comparative simulation study. *PLOS ONE*, 19(2).
<https://doi.org/10.1371/journal.pone.0294982>

- Khachatryan, V., Sirunyan, A. M., Tumasyan, A., & Adam, W. (2017). Event shape variables measured in $Z \rightarrow \mu^+\mu^-$ events in pp collisions at $\sqrt{s} = 7$ TeV with the CMS detector. *Physical Review D*, *95*(11). <https://doi.org/10.1103/PhysRevD.95.012011>
- Lumley, T. (2012). The importance of the normality assumption in large sample inference. *Annual Review of Public Health*, *23*, 151–169. <https://doi.org/10.1146/annurev.publhealth.23.100901.140546>
- Ma, Y., Leng, C., & Leng, C. (2023). *Optimal subsampling bootstrap for massive data*. <https://doi.org/https://arxiv.org/abs/2302.07533>
- Max Planck Society. (2003). *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. <https://doi.org/https://openaccess.mpg.de/Berlin-Declaration>
- Navidi, W. (2021). *Estadística para ingenieros y científicos* (McGraw-Hill, Ed.; 5th ed.).
- Rakhimov, Z. (2024). Simulation study on bootstrap confidence intervals in linear models: Case of heteroscedasticity. *World Journal of Advanced Research and Reviews*, *23*(3), 2250–2259. <https://doi.org/10.30574/wjarr.2024.23.3.2866>
- Rodríguez Medina, C. J., Terrazas Mata, M. M., Benítez Guadarrama, F. J., & Aldape Almilló, H. A. (2021). Pruebas de significancia en el análisis de regresión lineal simple. *Revista Iberoamericana Para La Investigación y El Desarrollo Educativo*, *11*(21), 1–26.
- Sayili, U., & Gunver, M. G. (2025). A novel logarithmic modification of the Kolmogorov–Smirnov test for large sample normality. *Scandinavian Journal of Clinical and Laboratory Investigation*, *85*(4), 287–298. <https://doi.org/10.1080/00365513.2025.2512384>

- Severiano A., Carriço J., Robinson D. A., & Ramirez M. (2011). Evaluation of Jackknife and Bootstrap for defining confidence intervals for pairwise agreement measures. *PLOS ONE*, 6(5). <https://doi.org/10.1371/journal.pone.0019539>
- Sirunyan, A. M., Tumasyan, A., Adam, W., & Ambrogi, F. (2020). Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *Journal of Instrumentation*, 12(2). <https://doi.org/10.1088/1748-0221/12/02/P02014>
- Tripathy, A., Xue, W., Andreassen, A., Feige, I., & Freytsis, M. (2019). ParticleNet: Jet tagging via particle clouds. *Physical Review D*, 100(9), 094004. <https://doi.org/10.1103/PhysRevD.100.094004>
- Veljkovic, K. (2024). Comparing Confidence Intervals for the Mean of Symmetric and Skewed Distributions. *Symmetry*, 16(11). <https://doi.org/10.3390/sym16111424>
- Wilcox, R. R. (2020). The percentile bootstrap: A primer with step-by-step instructions in R. *Advances in Methods and Practices in Psychological Science*, 3(4), 316–335. <https://doi.org/10.1177/2515245920911881>
- Yang, Y., & Xie, M. (2020). A review of statistical methods for constructing confidence intervals. *Quality and Reliability Engineering International*, 36(1), 283–301. <https://doi.org/10.1002/qre.2560>
- Zhang, Y., Lin, H., & Shao, X. (2023). Bootstrap inference for functional linear models with heteroscedastic errors. *ArXiv Preprint*. <https://doi.org/https://arxiv.org/abs/2311.07031>
- Zitzmann, S., Weirich, S., & Hecht, M. (2023). Accurate standard errors in multilevel modeling with heteroscedasticity: A computationally more efficient jackknife

technique. *Behavior Research Methods*, 55(2), 521–535.

<https://doi.org/10.3758/s13428-022-01917-3>

Anexos

Anexo 1. Código de programación en RStudio de los métodos analizados.

```
install.packages("readxl")
install.packages("ggplot2")
install.packages("systemfonts")
install.packages("nortest")

library(readxl)
library(ggplot2)
library(systemfonts)
library(nortest)

datos_jet <- read_excel("C:/Users/USER/OneDrive/Desktop/Datos_Jet_pt.xlsx")
print(datos_jet)
media <- mean(datos_jet$Jet_pt)
desv <- sd(datos_jet$Jet_pt)
print(media)
print(desv)

ks.test(datos_jet, "pnorm", mean = media, sd = desv)
lillie.test(datos_jet$Jet_pt)
str(datos_jet)

# Extraer el vector de la columna
jet_pt <- datos_jet$Jet_pt
jet_pt <- jet_pt[!is.na(jet_pt)] # Si hay NAs, deben eliminarse

# Crear el data.frame
df <- data.frame(valores = datos_jet)
str(df)
```

Q-Q plot con ggplot2

```
ggplot(df, aes(sample = Jet_pt)) +  
  stat_qq(size = 0.3) + # puntos Q-Q  
  stat_qq_line(color = "red", linetype = "dashed", size = 0.8) + # línea referencia  
  labs(  
    x = "Cuantiles teóricos (Normal)",  
    y = "Cuantiles muestrales",  
  ) +  
  scale_x_continuous(labels = function(x) format(x, decimal.mark = ",")) +  
  scale_y_continuous(labels = function(x) format(x, decimal.mark = ",")) +  
  theme_minimal()+  
  theme(  
    text = element_text(family = "Times New Roman", size = 24),  
    plot.title = element_text(size = 28, face = "bold"),  
    axis.title = element_text(size = 24),  
    axis.text = element_text(size = 20)  
  )  
  # Guardar el gráfico en formato BPM (vectorial)  
  ggsave(  
    filename = "Q-Q.bmp",  
    plot = last_plot(), # o plot = p si guardaste el gráfico  
    device = "bmp",  
    width = 15, height = 10,  
    dpi = 400  
  )
```

#Método Transformación logarítmica

```
# Definir tamaños de submuestra  
tamaños <- seq(1000, length(jet_pt), by = 1000)  
  
# Crear vectores para guardar resultados  
sd_est <- numeric(length(tamaños))
```

```

ic_inf <- numeric(length(tamaños))
ic_sup <- numeric(length(tamaños))

# Nivel de confianza
alpha <- 0.05
z_alpha <- qnorm(1 - alpha/2) # z = 1.96 para 95%

# Bucle sobre submuestras acumulativas
for (i in seq_along(tamaños)) {
  muestra <- jet_pt[1:tamaños[i]]
  n <- length(muestra)

  # Estimación puntual de la desviación estándar
  s <- sd(muestra)
  sd_est[i] <- s

  # Error estándar de log(s)
  se_log_s <- 1 / sqrt(2 * (n - 1))

  # IC para log(s)
  log_s <- log(s)
  li_log <- log_s - z_alpha * se_log_s
  ls_log <- log_s + z_alpha * se_log_s

  # Transformar de nuevo a escala original
  ic_inf[i] <- exp(li_log)
  ic_sup[i] <- exp(ls_log)
}

df_sd <- data.frame(
  tamaño_muestra = tamaños,
  desviacion_estimada = sd_est,

```

```

ic_inferior = ic_inf,
ic_superior = ic_sup
)

ggplot(df_sd, aes(x = tamaño_muestra, y = desviacion_estimada)) +
  geom_line(color = "black", size = 0.8) + # línea de estimaciones
  geom_ribbon(aes(ymin = ic_inferior, ymax = ic_superior), fill = "lightblue", alpha =
0.4) + # banda de IC
  geom_hline(yintercept = desv, color = "red", linetype = "dashed", size = 1) + # línea
de referencia global
  labs(x = "Tamaño de la submuestra",
  y = "Desviación estándar"
) +
  scale_x_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  scale_y_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  theme_minimal() +
  theme(
  text = element_text(family = "Times New Roman", size = 24),
  plot.title = element_text(size = 28, face = "bold"),
  axis.title = element_text(size = 24),
  axis.text = element_text(size = 20)
)

# Guardar el gráfico en formato BPM (vectorial)
ggsave(
  filename = "sd-subm.bmp",
  plot = last_plot(), # o plot = p si guardaste el gráfico
  device = "bmp",
  width = 15, height = 10,
  dpi = 400
)

```

```

df_sd$ancho_ic <- df_sd$ic_sup - df_sd$ic_inf
ggplot(df_sd, aes(x = desviacion_estimada, y = ancho_ic)) +
  geom_line(color = "blue") +
  geom_point(color = "blue", size = 1.5) +
  labs( x = "Desviación estándar estimada",
        y = "Ancho del IC"
  ) +
  scale_x_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  scale_y_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman", size = 24),
    plot.title = element_text(size = 28, face = "bold"),
    axis.title = element_text(size = 24),
    axis.text = element_text(size = 20)
  )

```

```
# Guardar el gráfico en formato BPM (vectorial)
```

```

ggsave(
  filename = "sd-ic.bmp",
  plot = last_plot(), # o plot = p si guardaste el gráfico
  device = "bmp",
  width = 15, height = 10,
  dpi = 400
)

```

```
#Método Método de Cuantiles empíricos (Percentil directo de submuestras)
```

```

set.seed(123)
tamaños <- seq(1000, length(jet_pt), by = 1000)
sd_est_1 <- numeric(length(tamaños))
ic_inf_1 <- numeric(length(tamaños))
ic_sup_1 <- numeric(length(tamaños))

```

```

num_submuestras <- 1000

for (i in seq_along(tamaños)) {
  t <- tamaños[i]
  muestra <- jet_pt[1:t]

  sd_est_1[i] <- sd(muestra)

  # Tamaño de submuestra: la mitad de t, pero máximo 1000 para que no sea muy grande
  tamaño_submuestra <- min(1000, floor(t / 2))

  sd_submuestras <- numeric(num_submuestras)

  for (j in 1:num_submuestras) {
    # Submuestra aleatoria sin reemplazo
    submuestra <- sample(muestra, size = tamaño_submuestra, replace = FALSE)
    sd_submuestras[j] <- sd(submuestra)
  }

  ic_inf_1[i] <- quantile(sd_submuestras, 0.025)
  ic_sup_1[i] <- quantile(sd_submuestras, 0.975)
}

df_sd_1 <- data.frame(
  tamaño_muestra = tamaños,
  desviacion_estimada = sd_est_1,
  ic_inferior = ic_inf_1,
  ic_superior = ic_sup_1
)

print(df_sd_1)

```

```

df_sd_1$ancho_ic <- df_sd_1$ic_superior - df_sd_1$ic_inferior
ggplot(df_sd_1, aes(x = desviacion_estimada, y = ancho_ic)) +
  geom_line(color = "blue") +
  geom_point(color = "blue", size = 1.5) +
  labs( x = "Desviación estándar estimada",
        y = "Ancho del IC"
  ) +
  scale_x_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  scale_y_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman", size = 24),
    plot.title = element_text(size = 28, face = "bold"),
    axis.title = element_text(size = 24),
    axis.text = element_text(size = 20)
  )

# Guardar el gráfico en formato BPM (vectorial)
ggsave(
  filename = "sd-ic_1.bmp",
  plot = last_plot(), # o plot = p si guardaste el gráfico
  device = "bmp",
  width = 15, height = 10,
  dpi = 400
)

#Método Bootstrap.
set.seed(123) # Reproducibilidad

# Reducimos el rango para evitar sobrecargar la memoria
tamaños <- seq(1000, 113000, by = 1000)

```

```

sd_est_2 <- numeric(length(tamaños))
ic_inf_2 <- numeric(length(tamaños))
ic_sup_2 <- numeric(length(tamaños))

num_bootstraps <- 300

for (i in seq_along(tamaños)) {
  t <- tamaños[i]
  muestra <- jet_pt[1:t]

  sd_est_2[i] <- sd(muestra)

  # Bootstrap con remuestreos más pequeños (50%)
  sd_bootstraps <- numeric(num_bootstraps)

  for (b in 1:num_bootstraps) {
    remuestreo <- sample(muestra, size = floor(t * 0.5), replace = TRUE)

    # Validar que no sean todos iguales
    if (length(unique(remuestreo)) > 1) {
      sd_bootstraps[b] <- sd(remuestreo)
    } else {
      sd_bootstraps[b] <- NA
    }
  }

  # Verificamos que haya datos válidos
  if (sum(!is.na(sd_bootstraps)) >= 30) {
    ic_inf_2[i] <- quantile(sd_bootstraps, 0.025, na.rm = TRUE)
    ic_sup_2[i] <- quantile(sd_bootstraps, 0.975, na.rm = TRUE)
  } else {

```

```

ic_inf_2[i] <- NA
ic_sup_2[i] <- NA
cat("Pocos datos válidos en muestra de tamaño", t, "\n")
}

# Liberar memoria cada 10 iteraciones
if (i %% 10 == 0) gc()
}

# Crear data.frame
df_sd_2 <- data.frame(
  tamaño_muestra = tamaños,
  desviacion_estimada = sd_est_2,
  ic_inferior = ic_inf_2,
  ic_superior = ic_sup_2
)

print(df_sd_2)

df_sd_2$ancho_ic <- df_sd_2$ic_superior - df_sd_2$ic_inferior

ggplot(df_sd_2, aes(x = desviacion_estimada, y = ancho_ic)) +
  geom_line(color = "blue") +
  geom_point(color = "blue", size = 1.5) +
  labs(x = "Desviación estándar estimada",
       y = "Ancho del IC"
  ) +
  scale_x_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  scale_y_continuous(labels = function(x) format(x, decimal.mark = ",")) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman", size = 24),

```

```

    plot.title = element_text(size = 28, face = "bold"),
    axis.title = element_text(size = 24),
    axis.text = element_text(size = 20)
  )

# Guardar el gráfico en formato BPM (vectorial)
ggsave(
  filename = "sd-ic_2.bmp",
  plot = last_plot(), # o plot = p si guardaste el gráfico
  device = "bmp",
  width = 15, height = 10,
  dpi = 400
)

#Gráfica de puntos 3 en 1 para IC.
df_sd$metodo <- "Transformación logarítmica"
df_sd_1$metodo <- "Cuantiles empíricos"
df_sd_2$metodo <- "Bootstrap"

df_comparado <- rbind(
  df_sd[, c("desviacion_estimada", "ancho_ic", "metodo")],
  df_sd_1[, c("desviacion_estimada", "ancho_ic", "metodo")],
  df_sd_2[, c("desviacion_estimada", "ancho_ic", "metodo")]
)

ggplot(df_comparado, aes(x = desviacion_estimada, y = ancho_ic, color = metodo)) +
  geom_line(linewidth = 1) + # Une los puntos con línea
  geom_point(size = 1.5) + # Agrega puntos si deseas visibilidad
  labs(x = "Desviación estándar estimada",
    y = "Ancho del IC",
    color = "Método"
  ) +

```

```

scale_x_continuous(labels = scales::scientific) +
scale_y_continuous(labels = scales::scientific) +
theme_minimal() +
theme(
  text = element_text(family = "Times New Roman", size = 20),
  plot.title = element_text(size = 24, face = "bold"),
  legend.title = element_text(size = 20),
  legend.text = element_text(size = 18),
  axis.title = element_text(size = 20),
  axis.text = element_text(size = 16)
)

```

```

# Guardar el gráfico en formato BPM (vectorial)
ggsave(
  filename = "3_en_1.bmp",
  plot = last_plot(), # o plot = p si guardaste el gráfico
  device = "bmp",
  width = 15, height = 10,
  dpi = 400
)

```

#Gráfica de cajas 3 en 1 para desviación estándar estimada.

```
# Método 1: Transformación logarítmica
```

```
df_sd$ancho_ic <- df_sd$ic_sup - df_sd$ic_inf
```

```
df_sd$metodo <- "Transformación logarítmica"
```

```
# Método 2: Cuantiles empíricos
```

```
df_sd_1$ancho_ic <- df_sd_1$ic_superior - df_sd_1$ic_inferior
```

```
df_sd_1$metodo <- "Cuantiles empíricos"
```

```
# Método 3: Bootstrap
```

```
df_sd_2$ancho_ic <- df_sd_2$ic_superior - df_sd_2$ic_inferior
```

```

df_sd_2$metodo <- "Bootstrap"

df_box <- rbind(
  df_sd[, c("ancho_ic", "metodo")],
  df_sd_1[, c("ancho_ic", "metodo")],
  df_sd_2[, c("ancho_ic", "metodo")]
)

ggplot(df_box, aes(x = metodo, y = ancho_ic, fill = metodo)) +
  geom_boxplot(alpha = 0.7) +
  labs(
    x = "Método",
    y = "Ancho del IC"
  ) +
  scale_y_continuous(labels = scales::scientific) + # o usa scale_y_log10() si es mejor
  con log
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman", size = 20),
    plot.title = element_text(size = 24, face = "bold"),
    axis.title = element_text(size = 20),
    axis.text = element_text(size = 16),
    legend.position = "none"
  )
# Guardar el gráfico en formato BPM (vectorial)
ggsave(
  filename = "3_en_1_box.bmp",
  plot = last_plot(), # o plot = p si guardaste el gráfico
  device = "bmp",
  width = 15, height = 10,
  dpi = 400
)

```