



Facultad
de
Ciencias

**Intervalos de confianza para medidas de
parámetros en teorías de campo efectivo EFT**
(Confidence Intervals for Parameter Measurements
in Effective Field Theories (EFT))

Trabajo de Fin de Grado
para acceder al
GRADO EN FÍSICA

Autor: Joaquín Pardo Ortiz
Tutor: Francisco Matorras Weining

Julio - 2025

Resumen

Las teorías efectivas de campo (EFT) son una herramienta fundamental en física de partículas para parametrizar de forma sistemática posibles desviaciones respecto al Modelo Estándar mediante operadores de dimensión superior. Estas teorías permiten encapsular los efectos de nueva física en unos pocos parámetros, cuya estimación o acotación se realiza a partir de medidas experimentales.

En la práctica, la inferencia estadística de estos parámetros suele basarse en métodos aproximados que presuponen condiciones como alta estadística, linealidad en los parámetros o la validez de teoremas asintóticos, hipótesis que no siempre se cumplen.

Este trabajo tiene como objetivo analizar críticamente la validez de dichas aproximaciones mediante simulaciones Monte Carlo, explorando en modelos simplificados alternativas tanto numéricas como analíticas para la construcción de intervalos de confianza. Se estudian especialmente los casos en los que fallan métodos tradicionales como el de Wilks, evaluando el rendimiento comparativo de otros enfoques más robustos, como las construcciones exactas de Neyman o Feldman-Cousins.

Palabras clave: Teorías efectivas de campo, Modelo Estándar, coeficientes de Wilson, intervalos de confianza, estadística frecuentista, simulaciones Monte Carlo, método de Wilks, construcción de Neyman, Feldman-Cousins.

Abstract

Effective Field Theories (EFTs) are a key tool in particle physics to systematically parametrize possible deviations from the Standard Model using higher-dimensional operators. These theories allow the effects of new physics to be encapsulated in a small number of parameters, whose estimation or constraint is based on experimental measurements.

In practice, statistical inference of these parameters often relies on approximate methods that assume conditions such as high statistics, linearity in the parameters, or the validity of asymptotic theorems, assumptions that are not always satisfied.

This work aims to critically assess the validity of such approximations using Monte Carlo simulations and to explore, through simplified models, both numerical and analytical alternatives for the construction of confidence intervals. Particular attention is given to scenarios where traditional methods, such as Wilks' theorem, fail, and a comparative evaluation is carried out with more robust approaches like the exact constructions of Neyman and Feldman-Cousins.

Key words: Effective Field Theories, Standard Model, Wilson coefficients, confidence intervals, frequentist statistics, Monte Carlo simulations, Wilks' method, Neyman construction, Feldman-Cousins.

Agradecimientos

Me gustaría agradecer en especial a mis padres, hermana y abuela, ya que ellos son los que han convivido conmigo durante toda la carrera y me han apoyado y ayudado en todo momento para poder superarla satisfactoriamente.

Otra parte fundamental a la que estoy muy agradecido es a mi grupo de amigos de la facultad, "los chipis", que me han brindado momentos inolvidables y han sabido sacarme una sonrisa cuando más lo necesitaba, así como ayudarme cuando era necesario.

Por último, agradecer a mi tutor Francisco Matorras que no solo me ha sabido guiar a lo largo de este trabajo, sino que también lo ha hecho a lo largo de la carrera.

Índice

1	Introducción	1
2	Fundamentos estadísticos	3
2.1	Probabilidades y enfoques estadísticos	3
2.2	Intervalos de confianza y construcción estadística	3
2.3	Coverage y propiedades de los intervalos	4
2.3.1	Coverage (Cobertura)	4
2.3.2	Evaluación del coverage	4
2.3.3	Relación entre coverage, potencia y conservadurismo	5
2.3.4	Coverage en regímenes no asintóticos	5
2.3.5	Ejemplo ilustrativo	5
2.3.6	Importancia práctica en física de partículas	6
3	Métodos Para La estimación de intervalos de confianza	7
3.1	Wilks	7
3.1.1	Fundamento teórico	7
3.1.2	Ventajas del método	8
3.1.3	Limitaciones para el problema estudiado	8
3.2	Construcción de Neyman	9
3.2.1	Ventajas y limitaciones	10
3.2.2	Aplicación en física experimental	11
3.3	Feldman-Cousins	11
3.3.1	Motivación y contexto físico	11
3.3.2	Fundamento del método	11
3.3.3	Ventajas frente a otros métodos	12
3.3.4	Limitaciones y consideraciones	12
3.3.5	Importancia en este trabajo	12
4	Primera aplicación del método	13
4.1	Modelo lineal $a \cdot g$	14
4.2	Dependencia cuadrática sin término lineal $b \cdot g^2$	18
4.3	Caso $ag + bg^2$	20

5 Cálculo del intervalo de confianza para el caso de histogramas	26
5.1 Estudio del coverage para diferentes casos	31
5.2 Estudio de la eficiencia computacional	33
6 Conclusiones	38
Referencias	40

CAPÍTULO 1

INTRODUCCIÓN

El Modelo Estándar (SM) de la física de partículas constituye la teoría más completa y comprobada hasta la fecha para describir las interacciones fundamentales entre partículas elementales. A pesar de su éxito, el SM no puede explicar una serie de fenómenos observados, como la materia oscura, la energía oscura, las oscilaciones de neutrinos o la asimetría materia-antimateria. Estas limitaciones motivan la búsqueda de nuevas teorías que extiendan el marco del SM.

Una de las aproximaciones más utilizadas para estudiar de manera indirecta posibles efectos de nueva física es el formalismo de las teorías de campo efectivo (EFT, por sus siglas en inglés). Las EFT permiten extender el lagrangiano del SM mediante la incorporación de operadores de dimensión superior, construidos con campos del SM que respetan sus simetrías gauge. El lagrangiano efectivo puede escribirse como:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_{d>4} \sum_i \frac{c_i^{(d)}}{\Lambda^{d-4}} \mathcal{O}_i^{(d)}, \quad (1.1)$$

donde Λ es la escala de nueva física, $\mathcal{O}_i^{(d)}$ son operadores de dimensión d , y $c_i^{(d)}$ son los coeficientes de Wilson que cuantifican su intensidad. En muchos análisis se restringe la expansión a operadores de dimensión seis, ya que los de dimensión cinco violan el número leptónico y los de dimensión ocho están más suprimidos por Λ [9, 2].

Este marco, conocido como SMEFT (SM Effective Field Theory), se ha aplicado extensamente en análisis experimentales del LHC, incluyendo estudios en sectores como el de quarks top, bosones electrodébiles o el Higgs. En particular, procesos de producción asociada como $t\bar{t}H$, $t\bar{t}Z$, tZq o tHq son especialmente sensibles a desviaciones respecto al SM debido al gran acoplamiento del top y la limpieza experimental de los estados finales leptónicos [4].

En estos análisis, la dependencia de los observables respecto a los coeficientes de Wilson es en general no lineal. Esto se debe a que las amplitudes de los procesos dependen linealmente de los coeficientes, pero las secciones eficaces (y por tanto las distribuciones de eventos) dependen cuadráticamente. De hecho, como se señala en [4], los eventos simulados con EFT se repondrán mediante funciones cuadráticas en el espacio de 26 coeficientes de Wilson. Esta expansión hasta segundo orden es necesaria para capturar interferencias entre el SM y los nuevos operadores, así como entre operadores distintos, y representa el nivel de precisión habitual en los estudios actuales de EFT.

Sin embargo, esta estructura cuadrática tiene implicaciones importantes desde el punto de vista estadístico. La construcción de intervalos de confianza sobre los coeficientes de Wilson,

o combinaciones de ellos, requiere métodos estadísticos adecuados. Tradicionalmente, el teorema de Wilks se ha utilizado para construir intervalos de confianza, ya que establece que el estadístico de razón de verosimilitud sigue una distribución χ^2 bajo ciertas condiciones. Estas condiciones incluyen: un número suficientemente grande de observaciones (alta estadística), estimadores consistentes, que los parámetros del modelo se encuentren dentro del espacio permitido y, fundamentalmente, que el modelo sea lineal en los parámetros. [11, 5].

En el caso de las EFT, la dependencia cuadrática rompe esta linealidad, lo que puede hacer que el estadístico de Wilks no siga la distribución esperada. Esto puede llevar a intervalos de confianza con cobertura incorrecta: es decir, que incluyan el valor verdadero con una probabilidad distinta a la nominal. Trabajos como [4] aplican directamente el método de Wilks en contextos EFT, incluso cuando los efectos buscados son pequeños y la estadística limitada. Aunque no se discute en detalle la validez de la cobertura obtenida, estudios como el presente muestran que en tales condiciones pueden producirse desviaciones significativas respecto al nivel nominal.

Para abordar esta cuestión, existen alternativas más robustas como la construcción de Neyman basada en pseudoexperimentos, o el método de Feldman-Cousins, que define una regla de ordenación óptima de los resultados experimentales y evita problemas como el "flip-flopping" [6, 8]. Estos métodos, si bien más costosos computacionalmente, permiten obtener intervalos con coverage exacto o controlado, incluso en presencia de no linealidades o espacios paramétricos con simetrías y degeneraciones (como la equivalencia de g y $-g$ en modelos simétricos).

Este trabajo tiene como objetivo estudiar de forma sistemática la validez de los métodos estadísticos tradicionales y alternativos para la construcción de intervalos de confianza en modelos con estructura funcional inspirada en EFT. Para ello, se analizan modelos simplificados donde el observable depende de un parámetro g con distintas formas funcionales: lineal (ag), cuadrática (bg^2) y combinada ($ag + bg^2$). A través de simulaciones Monte Carlo se evalúa la cobertura de cada método, identificando los regímenes donde fallan las aproximaciones asintóticas y proponiendo soluciones más fiables.

El estudio se estructura como sigue: primero se revisan los fundamentos teóricos de los métodos frecuentistas para la construcción de intervalos; luego se aplican a modelos analíticos simples para comprobar sus propiedades estadísticas; posteriormente se consideran modelos más realistas que reproducen las no linealidades típicas de los análisis EFT; y finalmente se extiende el análisis a distribuciones más complejas como histogramas binned, evaluando cómo varían las conclusiones según el tipo de observable. Con ello se busca proporcionar una guía práctica para físicos de partículas sobre cuándo confiar en las aproximaciones estadísticas habituales y cuándo es necesario recurrir a métodos exactos.

CAPÍTULO 2

FUNDAMENTOS ESTADÍSTICOS

En este capítulo se presentan las herramientas estadísticas básicas necesarias para entender y aplicar los métodos utilizados en física de partículas, en particular para la interpretación de resultados experimentales mediante modelos efectivos y la estimación de parámetros físicos. El objetivo es proporcionar una base sólida para los capítulos posteriores, donde se detallarán los métodos de construcción de intervalos de confianza y los enfoques estadísticos específicos para problemas físicos reales.

2.1. Probabilidades y enfoques estadísticos

La física experimental se enfrenta continuamente a datos con incertidumbre. El marco para manejar esta incertidumbre es la teoría de la probabilidad. Existen dos enfoques principales para interpretarla:

- El enfoque **frecuentista**, donde las probabilidades se interpretan como frecuencias relativas en un número infinito de repeticiones de un experimento.
- El enfoque **bayesiano**, donde las probabilidades representan un grado de creencia o conocimiento sobre un parámetro desconocido, basado en la información disponible.

Ambos enfoques tienen ventajas y limitaciones. El frecuentista domina en los análisis de alta energía, donde las construcciones de Neyman y los teoremas asintóticos (como el teorema de Wilks) proporcionan herramientas potentes, especialmente cuando se dispone de grandes cantidades de datos. El bayesiano permite incorporar información previa y es especialmente útil cuando los datos son escasos o hay restricciones físicas complicadas.

2.2. Intervalos de confianza y construcción estadística

El concepto de intervalo de confianza juega un papel central en la presentación de resultados ya que es el rango de valores del parámetro de interés que es compatible con los datos observados, bajo un cierto nivel de confianza (típicamente 68 %, 90 % o 95 %).

En el enfoque frecuentista, los intervalos se construyen de modo que, si el experimento se repitiera muchas veces, un porcentaje dado de esos intervalos incluiría el valor verdadero del parámetro. En cambio, en el enfoque bayesiano, se habla de *credibility intervals*, que reflejan la probabilidad condicionada del parámetro dado el dato observado y una distribución a priori.

En ambos casos, aparecen desafíos importantes:

- Las restricciones físicas (valores acotados por consideraciones físicas) pueden producir intervalos no triviales.
- La baja estadística puede hacer que los teoremas asintóticos no sean aplicables, exigiendo soluciones específicas como los métodos de Feldman-Cousins.
- La presencia de parámetros no deseados (*nuisance parameters*), que influyen en los datos pero no son el objetivo del análisis, requiere técnicas como el perfilado o la marginalización.

2.3. Coverage y propiedades de los intervalos

2.3.1. Coverage (Cobertura)

Un aspecto central en los métodos frecuentistas es el concepto de **coverage** (cobertura). Este concepto se refiere a la proporción de veces que un intervalo de confianza, construido según un procedimiento determinado, incluye el valor verdadero del parámetro de interés, cuando se repite el experimento bajo las mismas condiciones.

Formalmente, sea θ el parámetro desconocido y sea $[L(x), U(x)]$ el intervalo de confianza construido a partir de los datos x , obtenidos de una distribución $f(x | \theta)$. El coverage para un valor dado de θ se define como:

$$\text{Coverage}(\theta) = P_{\theta}(\theta \in [L(x), U(x)]),$$

donde la probabilidad se toma sobre la distribución de los datos, manteniendo θ fijo.

El nivel de confianza nominal (por ejemplo, $1 - \alpha = 0.90$) es el nivel para el cual se diseña el procedimiento: se espera que, en promedio, un 90 % de los intervalos construidos contengan el valor verdadero de θ . Sin embargo, en la práctica, los métodos pueden tener:

- **Coverage exacto:** cuando el coverage real coincide exactamente con el nivel nominal para todos los valores de θ .
- **Overcoverage:** cuando el coverage real es mayor al nivel nominal, es decir, los intervalos son más conservadores de lo necesario. Aunque esto garantiza que no se infringe el nivel de confianza, puede implicar intervalos excesivamente amplios y menor potencia para excluir hipótesis alternativas.
- **Undercoverage:** cuando el coverage real es menor al nivel nominal, es decir, los intervalos fallan en contener el valor verdadero más veces de lo prometido. Esto es problemático porque genera interpretaciones engañosas: los resultados aparentan ser más precisos de lo que realmente son.

2.3.2. Evaluación del coverage

El coverage se puede calcular de forma exacta solo en casos muy simples. En general, se evalúa mediante simulaciones de Monte Carlo:

1. Se fija un valor de θ (el valor verdadero).

2. Se generan muchas réplicas de datos x según la distribución $f(x | \theta)$.
3. Para cada conjunto de datos, se construye el intervalo $[L(x), U(x)]$ según el procedimiento elegido.
4. Se cuenta qué fracción de los intervalos generados contienen el valor θ .

Este procedimiento se repite para varios valores de θ , obteniendo un perfil de coverage como función del parámetro. Este perfil es fundamental para validar el uso de métodos aproximados.

2.3.3. Relación entre coverage, potencia y conservadurismo

Un método de construcción de intervalos no solo debe tener buen coverage, sino que también debe ser eficiente. Un método demasiado conservador (overcoverage) produce intervalos muy amplios que, aunque garantizan contener el valor verdadero, pierden capacidad para restringir hipótesis alternativas. Por otro lado, un método demasiado agresivo (undercoverage) puede parecer más potente, pero en realidad genera un número inaceptable de intervalos que no cumplen con el nivel de confianza prometido, lo que lleva a conclusiones estadísticamente inválidas.

En física experimental, mantener el coverage al nivel nominal es especialmente importante cuando los resultados se usan para excluir regiones del espacio de parámetros (por ejemplo, imponer límites superiores sobre una sección eficaz de producción o un acoplamiento). Si se usa un método con undercoverage, se corre el riesgo de excluir parámetros que en realidad son compatibles con los datos, lo que puede llevar a rechazar hipótesis válidas.

2.3.4. Coverage en regímenes no asintóticos

Muchos de los métodos frecuentistas usados en análisis de datos, como el teorema de Wilks, se basan en la linealidad, aproximaciones asintóticas válidas solo para tamaños de muestra grandes y parámetros en el interior del espacio permitido. Cuando estas condiciones no se cumplen (por ejemplo, cuando se trabaja cerca de límites físicos como $\mu \geq 0$, o con pocos eventos), el coverage real puede desviarse significativamente del valor nominal, generando undercoverage.

Por eso, en análisis críticos como búsquedas de nueva física con poca estadística, se recomienda usar métodos exactos (como la construcción de Neyman o el método de Feldman-Cousins) o, al menos, validar las aproximaciones mediante simulaciones Monte Carlo para asegurar que el coverage se mantiene en los rangos relevantes.

2.3.5. Ejemplo ilustrativo

Supongamos que queremos construir un intervalo de confianza del 90 % para una tasa de eventos μ a partir de un número observado n de sucesos Poisson. Si usamos un método aproximado (como Wilks) sin comprobar el coverage real, podríamos encontrarnos con que, para $\mu \approx 0$, el intervalo excluye el valor verdadero más de un 10 % de las veces. Esto significa que, aunque formalmente anunciemos un intervalo al 90 % C.L., en realidad estamos construyendo un procedimiento con undercoverage. En contraste, un método exacto, diseñado para respetar los límites físicos, garantizaría que al menos un 90 % de los intervalos contienen μ , tal como promete.

2.3.6. Importancia práctica en física de partículas

El coverage no es solo una propiedad matemática abstracta: tiene implicaciones directas en la interpretación de resultados. Un experimento que reporta un límite superior al 95 % C.L. sobre un parámetro físico está diciendo, en términos frecuentistas, que si se repitiera ese análisis muchas veces, en al menos un 95 % de los casos el límite superior incluiría el valor verdadero. Si ese límite está mal calibrado y subcubre, se corre el riesgo de excluir regiones del parámetro que, en realidad, no están excluidas. Por eso, los análisis experimentales robustos incluyen validaciones explícitas del coverage, especialmente cuando se trabaja cerca de bordes físicos, con conteos pequeños o usando métodos aproximados.

CAPÍTULO 3

MÉTODOS PARA LA ESTIMACIÓN DE INTERVALOS DE CONFIANZA

Este capítulo se centra en los métodos estadísticos utilizados para construir intervalos de confianza en el contexto de la teoría efectiva de campos (EFT). Se revisan los fundamentos y aplicaciones del teorema de Wilks, el método de Neyman y el método de Feldman-Cousins, destacando cómo cada uno aborda las particularidades de los modelos no lineales en EFT. Se explica por qué el método de Wilks no es adecuado para este tipo de modelos y se analiza cómo las alternativas proponen soluciones para manejar la dependencia no lineal de los parámetros, asegurando intervalos de confianza más precisos y fiables.

3.1. Wilks

El método de Wilks, basado en el teorema formulado por Samuel S. Wilks en 1938 [11], es uno de los pilares de la inferencia estadística frecuentista moderna. Este resultado establece que, bajo ciertas condiciones generales, el estadístico de razón de verosimilitud $-2 \ln \lambda$ sigue asintóticamente una distribución χ^2 con grados de libertad igual al número de restricciones impuestas. Gracias a este resultado, se pueden construir intervalos de confianza y regiones de exclusión de forma sencilla, evitando tener que calcular la distribución exacta de los datos bajo cada hipótesis. El método de Wilks ha sido ampliamente adoptado en física de partículas, bioestadística y econometría, dado que permite transformar problemas complejos en cálculos relativamente simples.

3.1.1. Fundamento teórico

Supongamos que se observa un conjunto de datos x generado según una distribución dependiente de un parámetro θ . La función de verosimilitud se define como $L(\theta) = P(x | \theta)$. Queremos contrastar la hipótesis nula $H_0 : \theta = \theta_0$ frente a la alternativa $H_1 : \theta \neq \theta_0$.

El estadístico de razón de verosimilitud se construye como:

$$\lambda(x) = \frac{L(\theta_0)}{L(\hat{\theta})},$$

donde $\hat{\theta}$ es el estimador de máxima verosimilitud (MLE).

El teorema de Wilks establece que, bajo condiciones generales (tamaño de muestra grande, modelo regular, parámetros interiores), el estadístico:

$$-2 \ln \lambda(x) = -2[\ln L(\theta_0) - \ln L(\hat{\theta})],$$

se distribuye asintóticamente como una variable χ^2 con k grados de libertad, donde k es la diferencia de grados de libertad entre H_1 y H_0 . Esta aproximación permite construir intervalos de confianza al resolver:

$$-2 \ln \lambda(\theta) \leq \chi_1^2(1 - \alpha),$$

lo que define el conjunto de valores plausibles de θ al nivel de confianza deseado.

En la práctica, es muy común aplicar una simplificación derivada del teorema: para construir intervalos al 68.3% de confianza, basta encontrar los valores de θ para los cuales el logaritmo de la verosimilitud cae no más de 0.5 unidades respecto al máximo.

3.1.2. Ventajas del método

El método de Wilks es extraordinariamente práctico:

- Permite evitar simulaciones Monte Carlo costosas.
- Es aplicable a problemas multi-dimensionales.
- Su interpretación geométrica en el espacio de parámetros es intuitiva.
- Ha sido validado empíricamente en muchos contextos experimentales.

Por estas razones, es ampliamente utilizado en los análisis estadísticos de física de altas energías, donde el tamaño de los datasets suele ser grande y las aproximaciones asintóticas funcionan bien.

3.1.3. Limitaciones para el problema estudiado

El método de Wilks es altamente efectivo cuando se cumplen las condiciones asintóticas y de regularidad que establece el teorema original. Estas condiciones incluyen que los parámetros del modelo sean lineales en los datos y que el número de observaciones sea lo suficientemente grande como para que los resultados asintóticos sean una buena aproximación. Sin embargo, en nuestro caso de estudio, estas condiciones no se cumplen debido a la dependencia no lineal entre los parámetros y los observables.

En teorías efectivas de campo (EFT), los coeficientes de Wilson se incluyen a través de operadores de dimensión seis y modifican los observables de manera no lineal. Específicamente, los observables dependen de los parámetros de forma cuadrática (y potencialmente de orden superior), lo que introduce una no linealidad en la relación entre los parámetros y los datos experimentales. Esta dependencia cuadrática entre los parámetros de Wilson y los observables es la causa fundamental por la que el teorema de Wilks no es aplicable en este contexto.

En EFT, los coeficientes de Wilson no son lineales en el espacio paramétrico, lo que provoca que los estadísticos de verosimilitud no sigan una distribución χ^2 estándar. En lugar de esto, suelen mostrar una distribución sesgada y no centrada en los valores esperados para los parámetros. Esta situación altera significativamente el coverage de los intervalos de confianza construidos utilizando el método de Wilks, llevando al undercoverage (los intervalos de confianza son más pequeños de lo que deberían) o al overcoverage (los intervalos incluyen el valor

verdadero más veces de lo esperado).

Por lo tanto, en nuestro caso, el fallo del método de Wilks no se debe a problemas relacionados con bordes físicos o datos discretos, sino a la no linealidad de los parámetros. Cuando los parámetros en un modelo dependen de manera cuadrática de los observables, como sucede en EFT, la distribución de los estadísticos de verosimilitud se aleja de la distribución χ^2 , lo que hace que el teorema de Wilks y su aproximación no sean válidos. Esto requiere el uso de métodos estadísticos alternativos que puedan manejar este tipo de dependencias, como Neyman o Feldman-Cousins.

3.2. Construcción de Neyman

El método de Neyman se basa en la idea de que el parámetro de interés θ es una cantidad fija (aunque desconocida), mientras que los datos x son aleatorios. El objetivo es construir, para cada posible valor de θ , una región de aceptación en el espacio de los datos tal que:

$$P(x \in R(\theta) \mid \theta) = 1 - \alpha.$$

En el contexto de la construcción de intervalos de confianza, $1 - \alpha$ corresponde al nivel de confianza, que es la probabilidad de que el intervalo contenga el valor verdadero del parámetro. Por ejemplo, un nivel de confianza del 95 % corresponde a $\alpha = 0.05$, lo que implica que hay un 5 % de probabilidad de que el intervalo no contenga el valor verdadero del parámetro.

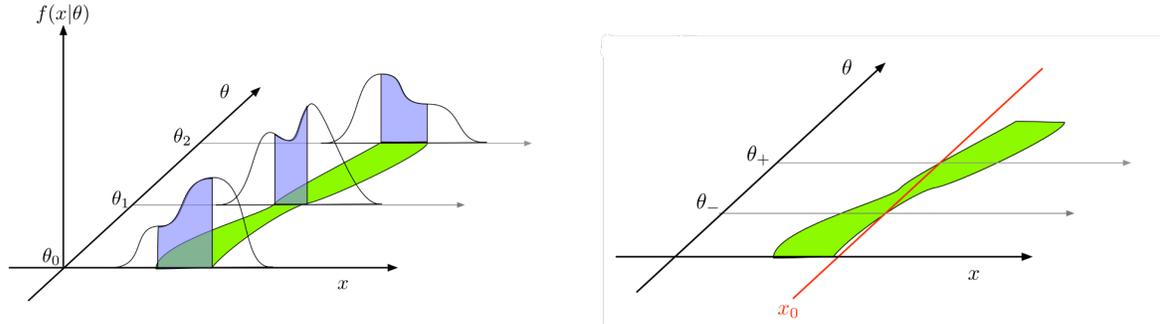
Estas regiones de aceptación se conectan para formar el llamado *cinturón de confianza* (*confidence belt*), una estructura geométrica en el plano (x, θ) que describe la relación entre datos y parámetros. Una vez realizado el experimento y obtenido un valor observado x_{obs} , se traza una línea vertical en x_{obs} y se identifican todos los valores de θ para los cuales x_{obs} pertenece a la región de aceptación. Ese conjunto constituye el intervalo de confianza asociado a la observación.

El procedimiento se puede resumir en los siguientes pasos:

1. Para cada valor fijo de θ , determinar la distribución de los datos x bajo la hipótesis de que θ es el valor verdadero.
2. Seleccionar una región $R(\theta)$ en el espacio de los datos que acumule una probabilidad de al menos $1 - \alpha$.
3. Reunir todas las regiones para formar el cinturón de confianza.
4. Una vez obtenido el dato observado x_{obs} , extraer el conjunto de valores de θ que lo mantienen dentro de su región de aceptación.

La generalidad del método permite aplicarlo tanto a problemas continuos como discretos. Sin embargo, requiere conocer bien la distribución de los datos y definir cuidadosamente las regiones de aceptación, especialmente en situaciones con restricciones físicas o distribuciones asimétricas.

Un aspecto importante es que la regla de ordenación (*ordering rule*), es decir, el criterio para seleccionar qué datos incluir en la región $R(\theta)$, no está determinada por el método. Diferentes elecciones pueden producir intervalos distintos, aunque todos respeten el coverage. Por ejemplo, en distribuciones gaussianas es común usar intervalos centrales, mientras que en problemas con límites físicos (como $\mu \geq 0$) puede ser más adecuado usar límites unilaterales.



(a) Esquema ilustrativo del cinturón de confianza (*confidence belt*) según el método de Neyman. Para cada valor del parámetro θ , se construye una región de aceptación sobre los datos x , que se conecta formando la banda sombreada. Adaptado de [7].

(b) Determinación del intervalo de confianza para un valor observado x_0 . Trazando una línea horizontal en x_0 sobre el cinturón de confianza, se identifican los valores de θ que quedan dentro de la banda sombreada, definiendo el intervalo $[\theta_-, \theta_+]$. Adaptado de [7].

Figura 3.1: Visualización del método de Neyman: a la izquierda, la construcción general del cinturón de confianza; a la derecha, la obtención del intervalo para un dato observado.

3.2.1. Ventajas y limitaciones

El método de Neyman presenta claras ventajas:

- Garantiza cobertura exacta al nivel nominal, sin depender de aproximaciones asintóticas ni suposiciones sobre tamaños de muestra grandes.
- Es aplicable en contextos de baja estadística y puede manejar distribuciones no simétricas o con restricciones físicas.
- Ofrece una base general que puede extenderse a problemas multi-dimensionales.

Sin embargo, también tiene limitaciones:

- La elección de la regla de ordenación puede introducir arbitrariedad, afectando la forma y las propiedades del intervalo.
- En modelos complejos, la construcción explícita del cinturón es computacionalmente costosa y suele requerir simulaciones numéricas.
- En problemas discretos, como los conteos Poisson, pueden aparecer discontinuidades en el coverage, generando overcoverage en algunas regiones y reduciendo la potencia estadística.
- El método es puramente frecuentista: no asigna una interpretación probabilística al parámetro dado el dato observado, a diferencia del enfoque bayesiano.

3.2.2. Aplicación en física experimental

En física de partículas, el método de Neyman es fundamental para establecer límites superiores, inferiores o bilaterales sobre parámetros como tasas de producción, secciones eficaces, acoplamientos o coeficientes de operadores efectivos en teorías EFT. Un ejemplo habitual es la determinación de un límite superior al 95 % C.L. sobre una tasa de producción cuando no se observa ninguna señal: incluso si el número de eventos observados es cero, el método permite excluir valores altos de la tasa, garantizando que el coverage se respeta.

En medidas de precisión, los intervalos bilaterales construidos según Neyman permiten reportar valores centrales acompañados de incertidumbres estadísticas bien definidas. Además, en regímenes de baja estadística, la aplicación exacta del método permite evitar los problemas derivados de usar aproximaciones asintóticas (como el teorema de Wilks), que pueden fallar cerca de los bordes del espacio de parámetros.

3.3. Feldman-Cousins

El método de Feldman-Cousins, propuesto por Gary J. Feldman y Robert D. Cousins en 1998 [8], se ha convertido en uno de los procedimientos más robustos para la construcción de intervalos de confianza en física de partículas. Este enfoque surgió para resolver problemas prácticos detectados en los métodos frecuentistas tradicionales, como la construcción clásica de Neyman, y en el uso de aproximaciones asintóticas como las del teorema de Wilks, especialmente en experimentos con baja estadística, límites físicos y datos discretos.

3.3.1. Motivación y contexto físico

En muchos experimentos de física, como las búsquedas de nueva física, los análisis de oscilaciones de neutrinos o la detección de partículas exóticas, los datos presentan características desafiantes, tales como conteos Poisson con fondo conocido, restricciones físicas (como tasas de producción no negativas) y tamaños muestrales pequeños. En estos contextos, los métodos tradicionales pueden no ser adecuados. Por ejemplo, los intervalos de confianza construidos mediante el método de Neyman pueden ser vacíos o no físicos, y los construidos con el teorema de Wilks tienden a subcubrir cerca de los bordes del espacio paramétrico [8]. El método de Feldman-Cousins, a diferencia de los métodos anteriores, resuelve estos problemas al proporcionar una regla de ordenación natural para la construcción de intervalos de confianza, garantizando que los resultados sean físicamente consistentes y cubran adecuadamente el espacio paramétrico. Este enfoque elimina la necesidad de tomar decisiones a posteriori sobre los límites unilaterales o bilaterales, asegurando un coverage preciso y fiable en los análisis experimentales.

3.3.2. Fundamento del método

El enfoque de Feldman-Cousins puede considerarse como una extensión de la construcción de Neyman, pero incorpora un ingrediente clave: una regla de ordenación basada en la razón de verosimilitudes. En lugar de seleccionar los datos a incluir en la región de aceptación

únicamente por probabilidad acumulada, se utiliza el estadístico:

$$R(x) = \frac{P(x | \mu)}{P(x | \mu_{\text{best}})},$$

donde μ_{best} es el valor del parámetro que maximiza la verosimilitud para el dato observado x (restringido al dominio físico permitido).

Esta regla tiene consecuencias importantes. Cuando los datos muestran una desviación significativa, el método produce automáticamente límites unilaterales; cuando los datos son compatibles con la hipótesis nula, genera intervalos bilaterales. Así se garantiza el coverage nominal para todos los valores del parámetro [8].

3.3.3. Ventajas frente a otros métodos

Feldman y Cousins demostraron que su método tiene varias ventajas notables sobre alternativas como Neyman puro o Wilks [8]:

- Mantiene cobertura frecuentista exacta incluso con datos discretos y parámetros en el borde del espacio permitido.
- Genera intervalos físicamente consistentes, evitando resultados no físicos o intervalos vacíos.
- Unifica los límites superiores y los intervalos bilaterales en un solo procedimiento, eliminando decisiones subjetivas basadas en los datos observados incluso en casos donde la región de confianza son la unión de varios intervalos disjuntos.

3.3.4. Limitaciones y consideraciones

A pesar de su solidez conceptual, el método de Feldman-Cousins no está exento de ciertas dificultades en su aplicación práctica. En problemas complejos, como modelos con múltiples parámetros, la construcción numérica de los cinturones de confianza puede ser computacionalmente intensiva, ya que requiere generar simulaciones Monte Carlo para evaluar las regiones de aceptación.

3.3.5. Importancia en este trabajo

En el contexto de este análisis, el método de Feldman-Cousins se adapta de manera óptima a las necesidades del problema, ya que resuelve de forma sistemática los problemas derivados de la no linealidad de los parámetros en modelos de teoría efectiva de campos (EFT). Mientras que otros enfoques como el de Wilks pueden presentar dificultades como ya se mencionó anteriormente, el método de Feldman-Cousins nos ayuda precisamente por ofrecernos esta regla de ordenación natural para construir intervalos de confianza. Este enfoque garantiza una coverage preciso y una interpretación físicamente consistente de los resultados experimentales, incluso cuando se enfrentan a distribuciones no simétricas o restricciones físicas. Su aplicación nos permite establecer límites más confiables, asegurando la coherencia física de los intervalos de confianza sin depender de suposiciones asintóticas.

CAPÍTULO 4

PRIMERA APLICACIÓN DEL MÉTODO

En este capítulo se presentan los principales resultados obtenidos a lo largo del desarrollo del trabajo, cuyo objetivo es comparar una medida experimental con un modelo de teoría efectiva de campos (EFT). El parámetro de interés en este análisis es el acoplamiento g , que representa la interacción entre los campos en el modelo EFT. El análisis comienza con una medida sencilla, donde se supone que los datos siguen una distribución gaussiana con anchura 1, y se ajusta un modelo de EFT de un único parámetro, el acoplamiento g . A medida que avanzamos, el modelo se complica introduciendo dependencias no lineales de g , como $a \cdot g + b \cdot g^2$, lo que permite estudiar cómo diferentes términos en el modelo afectan la estimación de g y los intervalos de confianza. Cabe destacar que los códigos de Python empleados para obtener todos los resultados de este trabajo se encuentran en [10].

Para este análisis, el likelihood es fundamental, ya que es utilizado para ajustar los parámetros del modelo a los datos experimentales y calcular los intervalos de confianza correspondientes. En cada uno de los casos estudiados, la función de pérdida se definió como el negativo del logaritmo del likelihood, $-2 \ln L(g)$, y el parámetro g se ajustó mediante el método de máxima verosimilitud (MLE). Como se describió en el capítulo anterior, este enfoque busca maximizar la probabilidad de los datos dados los parámetros del modelo.

Una vez ajustado el modelo, se calculan los intervalos de confianza utilizando el estadístico de razón de verosimilitud $-2 \ln \lambda(x)$. Este estadístico se comparó con una distribución χ^2 para obtener los límites de los intervalos de confianza para g .

En los casos con modelos más complejos, como los que incluyen términos no lineales de la forma $a \cdot g + b \cdot g^2$, la aplicación del método de Wilks mostró problemas de overcoverage cerca del mínimo, donde el modelo cuadrático tiene un comportamiento menos intuitivo. Para estos casos, se utilizó el método de Feldman-Cousins, que proporcionó intervalos más confiables y con un mejor coverage, especialmente en las regiones donde la dependencia no lineal entre los parámetros es más pronunciada.

Los resultados obtenidos muestran que los métodos de Neyman y Feldman-Cousins son más adecuados para modelos con dependencias no lineales, garantizando un coverage más preciso y evitando los problemas de undercoverage observados con el método de Wilks.

4.1. Modelo lineal $a \cdot g$

Para este primer caso, se asume que el observable y sigue una distribución normal centrada en $a \cdot g$ con varianza unitaria, es decir, $y \sim \mathcal{N}(a \cdot g, 1)$. La función de verosimilitud para una observación dada es entonces:

$$L(g) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y - a \cdot g)^2\right). \quad (4.1)$$

El estadístico basado en la razón de verosimilitud se define como $-2 \ln \lambda(g)$, donde $\lambda(g)$ es la razón de verosimilitud entre el valor hipotético g y el estimador de máxima verosimilitud \hat{g} . En este caso, se tiene:

$$-2 \ln \lambda(g) = (y - a \cdot g)^2 - (y - a \cdot \hat{g})^2, \quad (4.2)$$

donde $\hat{g} = y/a$. Como $(y - a \cdot \hat{g})^2 = 0$, se concluye que:

$$-2 \ln \lambda(g) = (y - a \cdot g)^2. \quad (4.3)$$

De este modo, la función de pérdida $L(g) = (y - a \cdot g)^2$ utilizada en esta sección se identifica con el estadístico $-2 \ln \lambda(g)$, que es la base para aplicar el método de Wilks en la construcción del intervalo de confianza.

Se representó entonces gráficamente la función L frente al parámetro g , observándose el comportamiento esperado de la pérdida en función de los valores de g .

Cabe destacar que este modelo tiene una resolución analítica, lo que significa que conocemos los resultados teóricos y podemos validar con certeza la exactitud de los métodos estadísticos empleados. Aunque sabemos que el método de Wilks es exacto en este caso, lo usamos principalmente como herramienta para desarrollar y probar los programas que luego serán aplicados a modelos más complejos, como aquellos con dependencias no lineales.

Considerando la función de pérdida anteriormente mostrada, queremos encontrar el intervalo de confianza al 68.3% utilizando el criterio de Wilks, es decir, identificando los valores de g que verifican:

$$\Delta L(g) = L(g) - L_{\min} = 1$$

La función $L(g)$ alcanza su mínimo cuando se anula su derivada:

$$\frac{dL}{dg} = \frac{d}{dg}(y - ag)^2 = -2a(y - ag)$$

Igualando a cero:

$$-2a(y - ag) = 0 \quad \Rightarrow \quad g_{\min} = \frac{y}{a}$$

El valor mínimo de L es entonces:

$$L_{\min} = (y - a \cdot g_{\min})^2 = (y - a \cdot \frac{y}{a})^2 = 0$$

Queremos ahora encontrar todos los valores de g tales que:

$$L(g) = (y - ag)^2 = 1$$

Desarrollando:

$$(y - ag)^2 = 1 \quad \Rightarrow \quad y - ag = \pm 1$$

De aquí se obtienen dos soluciones:

$$ag = y \pm 1 \quad \Rightarrow \quad g = \frac{y \pm 1}{a}$$

Por tanto, el intervalo de confianza al 68.3% C.L. está dado por:

$$g \in \left[\frac{y-1}{a}, \frac{y+1}{a} \right]$$

Como se ve, este resultado es simétrico respecto del estimador $g_{\text{mín}} = y/a$, ya que la función de pérdida es cuadrática. El hecho de que el mínimo de $L(g)$ sea cero simplifica considerablemente la obtención del intervalo, al reducir el problema a resolver una ecuación cuadrática con solución directa.

Haciendo todo este desarrollo en un programa de Python obtenemos una representación clara de este ejercicio, permitiendo visualizar el intervalo de confianza calculado, donde se puede ver claramente que los puntos de corte obtenidos se corresponden con los esperados.

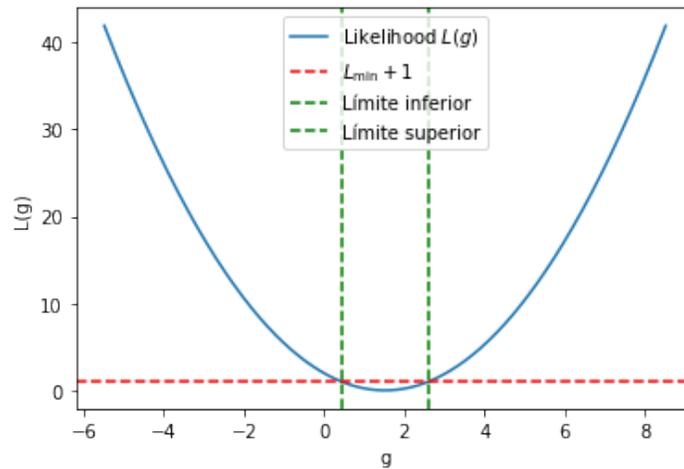


Figura 4.1: Obtención de intervalo de confianza por el método de Wilks para el caso $a \cdot g$

Para evaluar empíricamente la validez estadística del intervalo de confianza construido en este escenario lineal, se llevó a cabo un estudio de coverage mediante simulaciones Monte Carlo. Se fijó un valor concreto del parámetro a , y se generaron $N = 10^4$ réplicas del dato y siguiendo la distribución $y \sim \mathcal{N}(0, 1)$, correspondiente al caso $g = 0$.

Para cada valor simulado de y_i , se construyó el intervalo de confianza asociado al parámetro g mediante la expresión analítica deducida anteriormente:

$$g_i \in \left[\frac{y_i - 1}{a}, \frac{y_i + 1}{a} \right]$$

Dado que el valor verdadero inyectado de g fue cero, se contabilizó en cuántos casos el intervalo calculado incluía efectivamente el valor $g = 0$. El porcentaje de éxito sobre el total de simulaciones define el coverage observado del método. En este caso, se obtuvo una cobertura empírica muy cercana al 68% , en excelente acuerdo con el nivel de confianza nominal esperado. Este resultado confirma que, en un contexto lineal y gaussianamente bien comportado, el uso del criterio de Wilks produce intervalos con las propiedades frecuentistas deseadas.

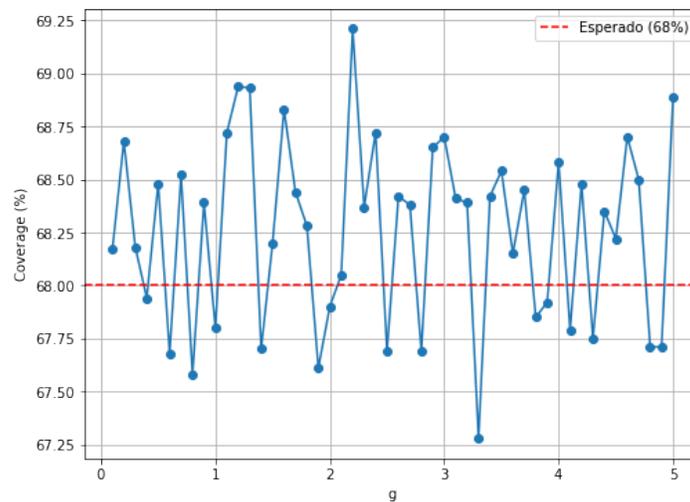


Figura 4.2: Resultado del coverage para el método de Wilks del caso a·g

Seguidamente, se construyó una figura que representa los intervalos de confianza analíticos y numéricos en función de y , superponiéndolos. Esta representación permitió confirmar de forma gráfica que ambos métodos proporcionan exactamente el mismo resultado en este caso sencillo.

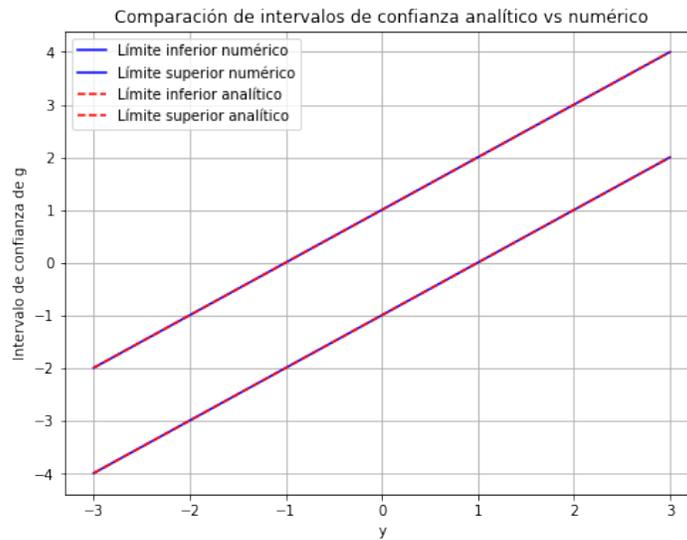


Figura 4.3: Comparación de intervalos de confianza por el método analítico y numérico

A continuación, se implementó una versión simplificada del método de Neyman para construir bandas de confianza sobre el estimador óptimo \hat{g} en un modelo lineal como el anterior. Para cada valor ideal de g , se generaron múltiples simulaciones de datos y , se calcularon los correspondientes estimadores \hat{g} mediante la solución analítica $\hat{g} = y/a$, y se extrajeron los percentiles 16% y 84% que delimitan el 68% central de los resultados.

Al repetir este proceso para un rango de valores de g , se obtuvieron las bandas de confianza que muestran cómo varía la dispersión esperada del estimador óptimo en función del valor verdadero. Estas bandas permiten visualizar y cuantificar el coverage del método bajo diferentes condiciones, ofreciendo una perspectiva clara sobre la fiabilidad de los intervalos de confianza generados.

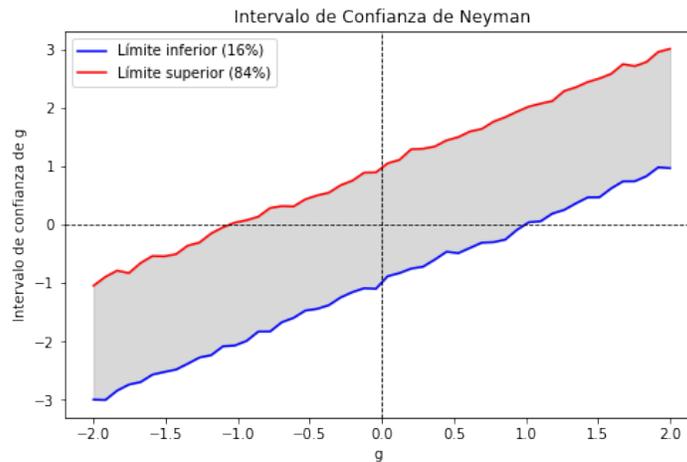


Figura 4.4: Intervalo de confianza por el método de Neyman para el caso $a \cdot g$

Al representar gráficamente los intervalos de confianza construidos mediante el método de Neyman para diferentes valores del parámetro g , se observa que los límites superior e inferior del intervalo no forman líneas rectas perfectas, sino curvas suaves con pequeñas irregularidades. Esta característica puede parecer contradictoria si se tiene en cuenta que el modelo subyacente es lineal y la distribución de los datos es normal.

Sin embargo, esta discrepancia se explica por el hecho de que los intervalos han sido obtenidos de forma empírica, a través de simulaciones Monte Carlo. Como resultado, los límites del intervalo dependen de las fluctuaciones estadísticas propias del proceso de simulación.

En contraste, si se empleara el resultado analítico exacto, se definirían bandas perfectamente rectas y paralelas respecto del eje g como bien muestra la figura 4.3.

Por tanto, las pequeñas desviaciones respecto a la linealidad observadas en la figura son una consecuencia directa de la estimación numérica basada en un número finito de simulaciones y de la resolución limitada en el muestreo de valores de g . A medida que se incrementa el número de réplicas Monte Carlo, estas curvas tienden a suavizarse y aproximarse a las líneas rectas correspondientes al resultado analítico. Además, una ventaja de este método es que se podrá aplicar a casos más complejos, como se irá viendo.

4.2. Dependencia cuadrática sin término lineal $b \cdot g^2$

Aplicando el procedimiento anterior de Feldman-Cousins a un caso más complejo como el caso bg^2 , se obtiene el resultado que se muestra en la figura 4.5.

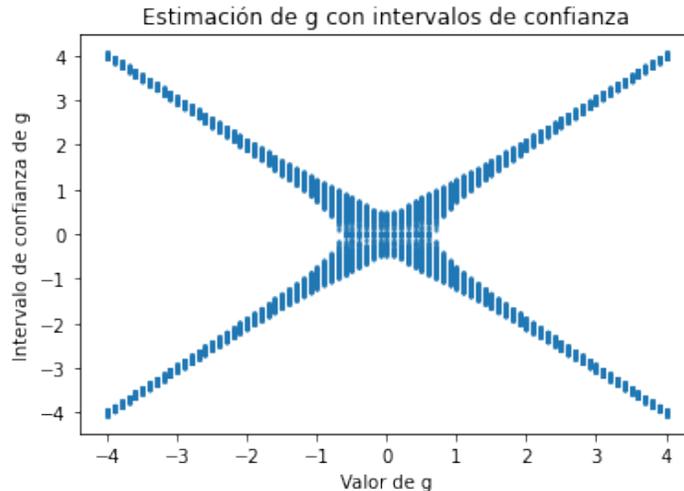


Figura 4.5: Intervalo de confianza por el método de Feldman-Cousins para el caso bg^2

El resultado presenta una característica forma en "X", reflejo de la simetría intrínseca del modelo: dado que g^2 es positivo, los mismos valores de y pueden corresponder tanto a $+g$ como a $-g$, produciendo dos ramas simétricas en la estimación. Cerca de $g = 0$, la sensibilidad del modelo disminuye debido a que el término cuadrático es pequeño y el ruido relativo domina,

ampliando los intervalos y generando una región central más ancha. En cambio, para valores grandes de $|g|$, el término bg^2 crece rápidamente y el efecto del ruido disminuye, estrechando los intervalos y mejorando la precisión de la estimación. Este comportamiento ilustra cómo la estructura funcional del modelo influye directamente en la forma y fiabilidad de los intervalos de confianza.

Cabe destacar que lo que realmente nos interesa de esta representación no es la nube completa de puntos individuales, sino las curvas que delimitan su contorno: es decir, las envolventes superior e inferior. Estas curvas definen, para cada valor dado de g , el rango dentro del cual se espera que caiga el dato y con una cierta probabilidad (en este caso al 68% de confianza). En otras palabras, los límites de esta gráfica constituyen la banda de aceptación en el espacio (g, y) , y son los que se utilizan posteriormente para determinar el intervalo de confianza asociado a un valor observado y_{obs} .

Por otro lado, era interesante también desarrollar un método para obtener el intervalo de confianza de un g dado. Para ello se desarrollaron diferentes métodos explicados a continuación.

Inicialmente se intentó aproximar la envolvente ajustándola a polinomios, con el objetivo de obtener una representación funcional de los intervalos. De esta manera, se obtendría el intervalo de confianza de un g dado obteniendo los puntos de corte con dichos polinomios. Sin embargo, se detectó que este método presentaba limitaciones en términos de exactitud.

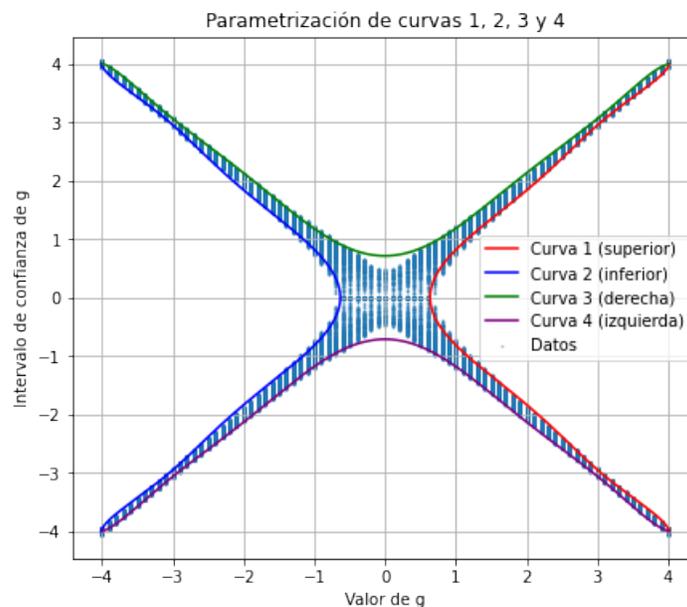


Figura 4.6: Parametrización con curvas de la envolvente para el caso bg^2

Como alternativa, se implementó una interpolación lineal entre los diferentes puntos de la envolvente utilizando funciones provistas por Python, logrando así una mejora sustancial en la precisión y fiabilidad de los intervalos calculados.

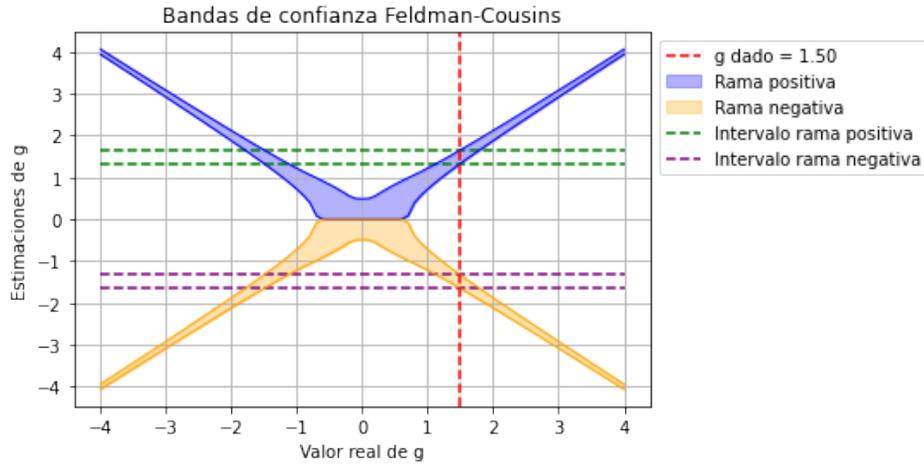


Figura 4.7: Interpolación de los puntos de la envolvente para el caso bg^2

Una vez hemos obtenido resultados satisfactorios para este método, calcular el intervalo de confianza para un g dado es muy simple, pues solo es necesario encontrar el punto de corte con las ramas construidas. Se puede ver en la figura 4.7 que para valores de g pequeños solo hay un intervalo, mientras que para valores algo más grandes conviene destacar que hay dos intervalos de confianza, pues esta recta corta cuatro veces las bandas, dos en la rama negativa y otras dos en la positiva, tal y como ocurre en el ejemplo ilustrado. Esta figura permite visualizar de forma clara cómo la simetría del modelo $y = bg^2$ se refleja en la estructura de los intervalos de confianza. La aparición de soluciones múltiples no es un artefacto numérico, sino una consecuencia directa de la degeneración funcional del modelo, en el que distintos valores de g pueden dar lugar al mismo valor observado y . Este comportamiento es fundamental en física de partículas, ya que en análisis EFT puede conducir a interpretaciones ambiguas si no se utilizan métodos capaces de capturar estas regiones disjuntas. La figura 4.7, por tanto, no solo ilustra un caso particular, sino que ejemplifica una situación representativa donde el uso de métodos exactos como Feldman-Cousins resulta esencial para garantizar la validez estadística del intervalo construido.

4.3. Caso $ag + bg^2$

En esta etapa, se abordó el caso más complejo correspondiente al modelo $ag + bg^2$. Los intervalos de confianza fueron obtenidos utilizando tanto el método Wilks como el método de Feldman-Cousins.

La obtención de los intervalos de confianza por el método de Feldman-Cousins se realiza de manera similar a la expuesta en las secciones anteriores, obteniéndose una figura similar a 4.7 pero con ligeras diferencias en la forma, pues esta depende de los parámetros a y b .

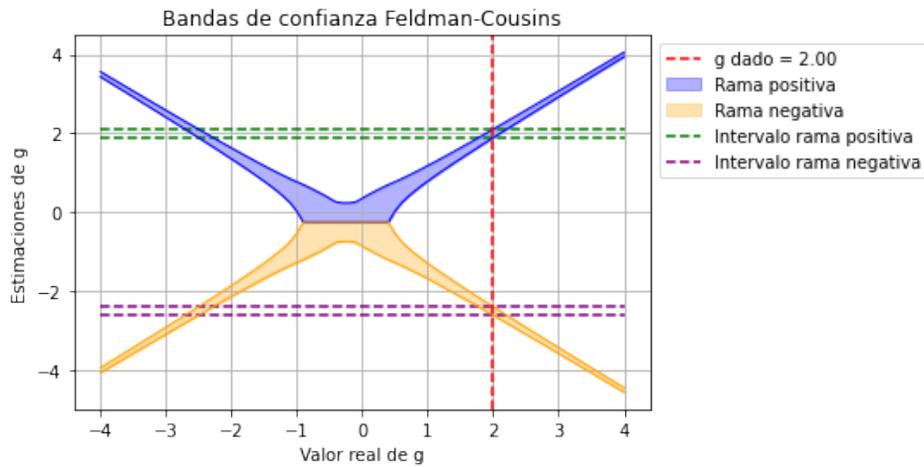


Figura 4.8: Intervalos de confianza por el método de Feldman-Cousins para el caso $ag + bg^2$

Tras esto, para el mismo caso, pero esta vez por el método de Wilks, se obtuvieron los intervalos de confianza, tal y como se ve en la siguiente figura.

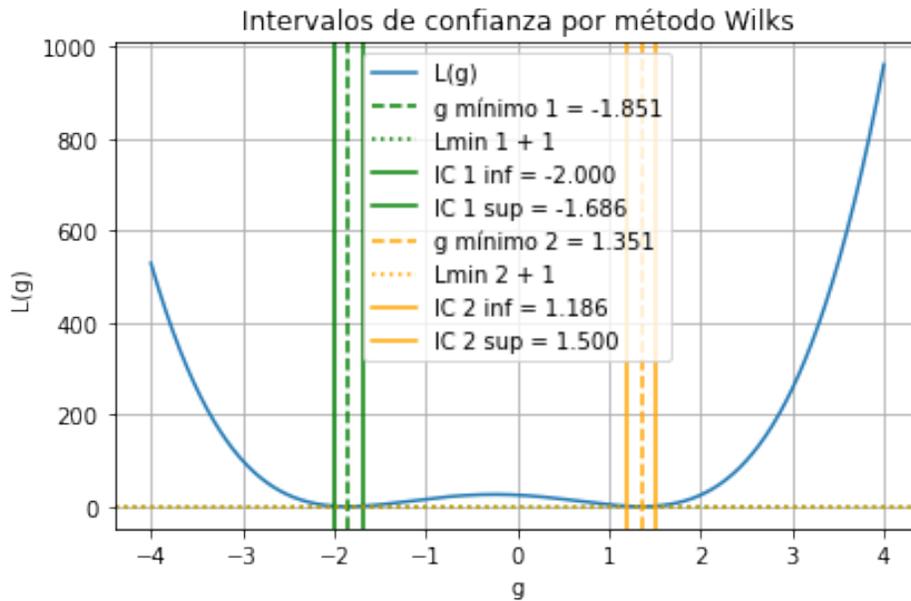


Figura 4.9: Intervalos de confianza por el método de Wilks para el caso $ag + bg^2$

Es interesante ahora comparar para distintos valores del y observado como varía el intervalo de confianza de g por ambos métodos. Para ello se representaron a la vez, viendo así en qué puntos ambos métodos concuerdan y cuándo hay discrepancias entre ellos.

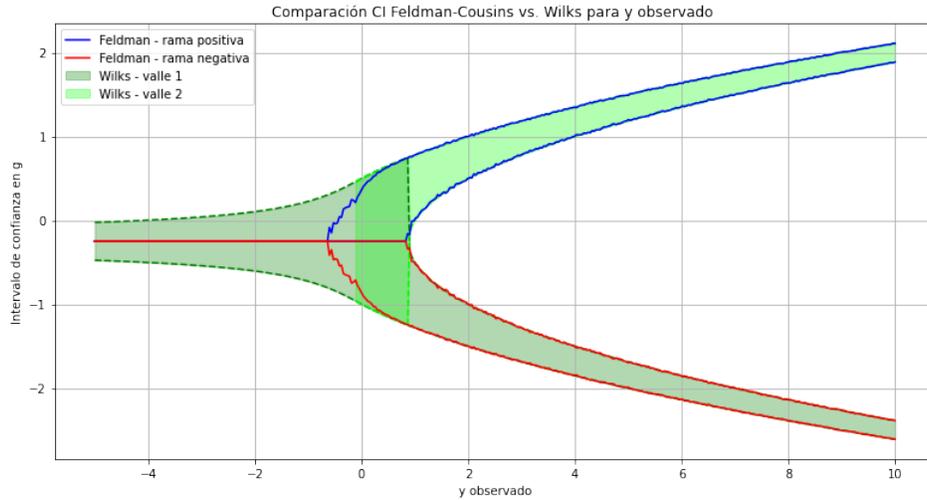


Figura 4.10: Comparación de los intervalos de confianza por el método de Wilks y el de Feldman-Cousins con parametrización en g_{exp}

En este contexto, la variable y representa la observación simulada, que sigue una distribución normal centrada en el valor predicho por el modelo: $y \sim \mathcal{N}(ag + bg^2, 1)$. Esta simplificación permite estudiar la influencia de los parámetros a y b sobre la forma funcional del observable, conservando una estructura suficientemente rica como para reflejar las dificultades estadísticas que aparecen en EFT. Así, los "valores pequeños de y " mencionados en esta sección se refieren a observaciones cercanas a cero, que pueden ser compatibles con múltiples valores de g debido a la degeneración del modelo cuando $g \approx 0$.

En esta última figura 4.10 se puede apreciar bien que ambos métodos concuerdan y tienen sentido para valores de y algo grandes, mientras que para valores pequeños aparecen una serie de problemas. Se puede ver que para estos valores pequeños de y el método de Feldman-Cousins no está bien definido, ya que no debería cortarse así. En los casos anteriores hemos parametrizado el método de Feldman-Cousins usando g_{teo} y g_{exp} porque suele ser exacto casi siempre, pero en este caso se ve que no. El problema de esto reside en que suponiendo que tenemos un modelo del tipo $y \rightarrow g^2$, para el cual se define el estimador $g_{exp} = \sqrt{y}$ si $y > 0$, y $g_{exp} = 0$ en caso contrario. Este estimador g_{exp} se obtiene a partir de la minimización de la función de pérdida asociada al modelo combinado $y = ag + bg^2$, es decir:

$$\mathcal{L}(g) = (y - ag - bg^2)^2$$

En este caso, la función a minimizar es una cuadrática en g con un término lineal y uno cuadrático, lo que impide obtener una expresión cerrada simple para el estimador. Por tanto, la minimización se realiza numéricamente. Además, dependiendo de los valores de a , b y y , puede haber una o varias soluciones reales, lo que introduce ambigüedad en regiones como $y \approx 0$, donde la función objetivo se vuelve poco sensible a variaciones en g .

Este comportamiento afecta de forma importante a la calidad estadística del procedimiento cuando se utiliza g_{exp} como variable de ordenación en el método de Feldman-Cousins, espe-

cialmente en las regiones donde la relación entre g e y es degenerada.

Esta construcción implica que todos los valores negativos de y son proyectados al mismo valor $g_{\text{exp}} = 0$. Al realizar la construcción de las bandas parametrizadas en función de g_{exp} , se pierde información esencial: por ejemplo, no se distingue si un $g_{\text{exp}} = 0$ proviene de $y = -1$ o de $y = -10$, lo cual introduce un error conceptual. Si, además, en la representación gráfica correspondiente se ignoran los casos con $y < 0$ (como se ha hecho), o se extiende artificialmente el valor $y = 0$ hacia valores negativos, se incurre en una descripción incorrecta del comportamiento del sistema.

La solución para esto sería tan simple como aplicar el mismo método pero en función de y , y parametrizar sobre ello, ya que en ese caso se preserva toda la información estadística de las simulaciones, sin proyecciones degeneradas. Por ello, en este tipo de modelos es preferible usar y como variable de ordenación en el método de Feldman-Cousins.

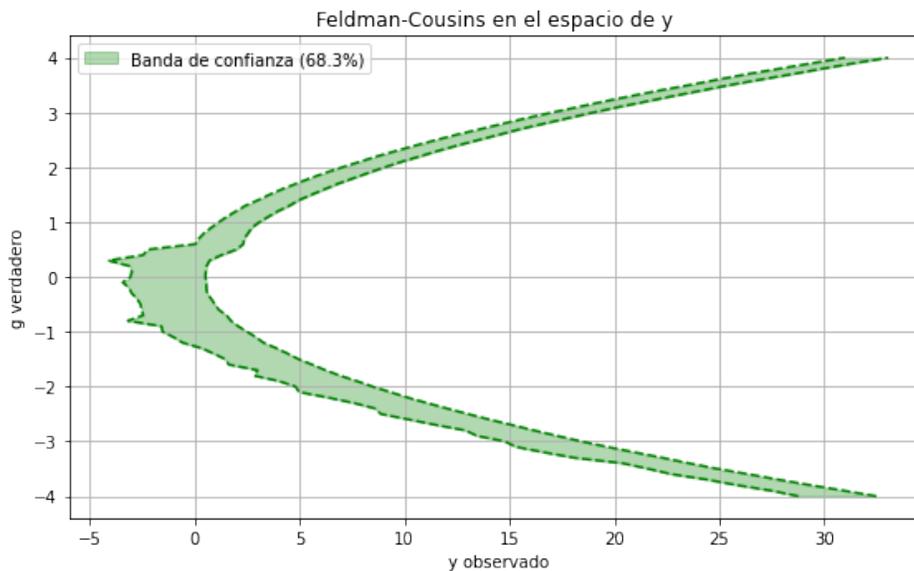


Figura 4.11: Representación del método de Feldman-Cousins basado en y

De este resultado pueden llamar la atención las fluctuaciones que hay para valores muy negativos, pero esto es normal, ya que al hacer la simulación aparecen pocos casos por tener probabilidad muy baja.

Una vez hecha esta corrección, la comparación de los intervalos de confianza para valores de y grandes nos devolvería los mismos resultados que antes, mientras que para valores de y pequeños, donde se presentaban problemas, obtendríamos lo siguiente.

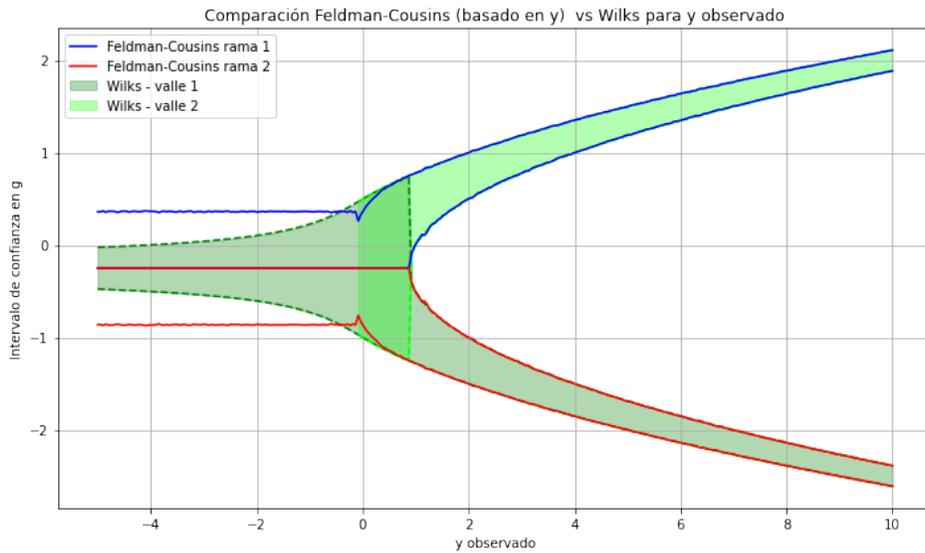


Figura 4.12: Comparación de los intervalos de confianza por el método de Wilks y el de Feldman-Cousins con parametrización en y

Como se puede apreciar, al hacer el cambio de parametrización a y , se solucionan ya los cortes extraños que hacían los intervalos de confianza del método de Feldman-Cousins para valores de $y < 0$.

Una vez resuelto el problema anterior, ya se podría evaluar el rendimiento estadístico de ambos métodos comparando el coverage de estos.

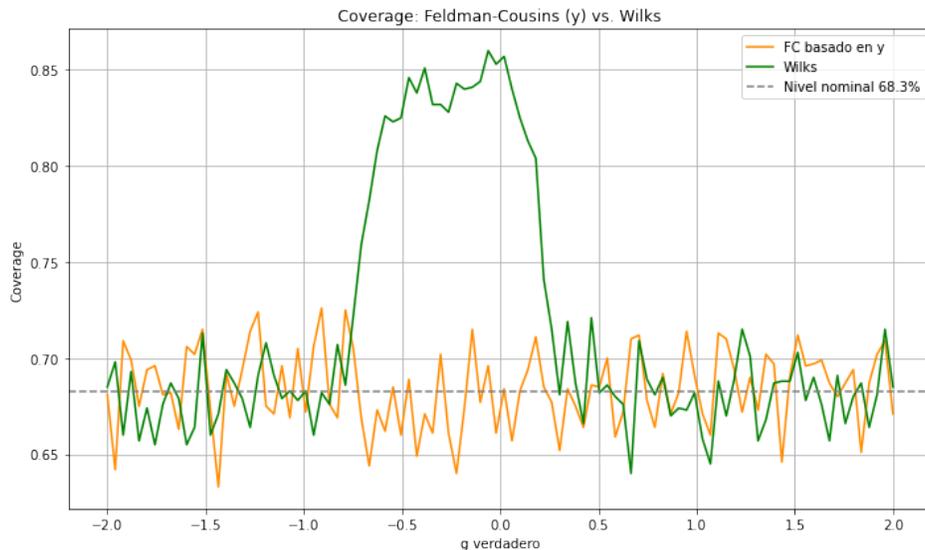


Figura 4.13: Comparación del coverage para el método de Wilks y el de Feldman-Cousins basado en y para el caso $ag + bg^2$

Como se puede apreciar, el método de Feldman-Cousins basado en y da un coverage esperado cercano al 68% mientras que en el método de Wilks se puede apreciar como en una zona hay una sobrecompensación del mismo. Esto se debe a que en la zona cercana al mínimo los intervalos de confianza calculados son exageradamente grandes.

Siguiendo en la misma línea de trabajo, se comparó también el coverage por el método de Wilks con el de Feldman-Cousins basado en g_{exp} , pero no sin antes hacer una aproximación. Esta consiste en dar el intervalo de confianza de $y=0$ a todos aquellos puntos donde $y < 0$. Esta aproximación en principio no nos garantizaría coverage.

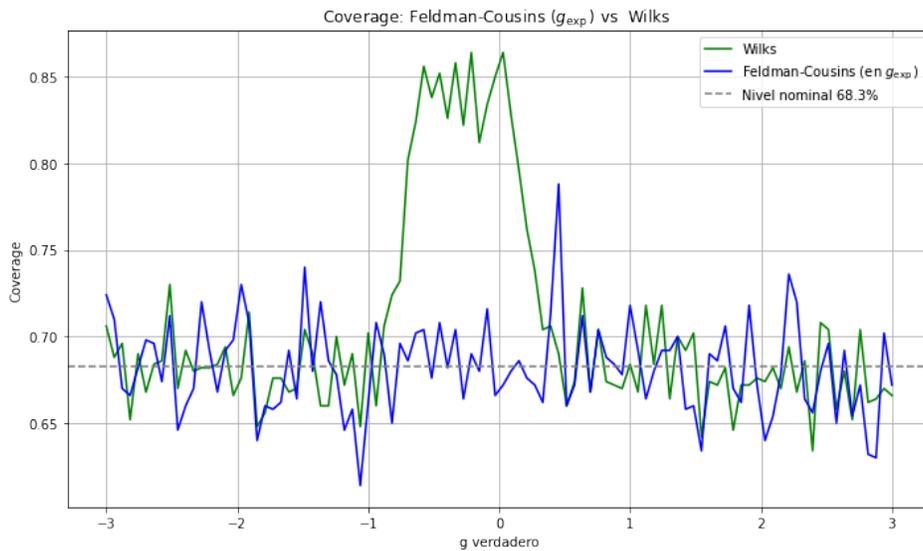


Figura 4.14: Comparación del coverage para el método de Wilks y el de Feldman-Cousins basado en g_{exp} para el caso $ag + bg^2$

A pesar de no garantizar este último método buenos resultados de coverage, como se puede ver en esta figura 4.14 el resultado es aparentemente correcto, salvo algún pico pronunciado debido probablemente a las fluctuaciones estadísticas.

CAPÍTULO 5

CÁLCULO DEL INTERVALO DE CONFIANZA PARA EL CASO DE HISTOGRAMAS

En este capítulo se aborda el caso en que los datos se organizan como un histograma, representando, por ejemplo, el valor de una variable cinemática como el momento de una partícula. Cada canal del histograma se modela con una dependencia cuadrática respecto al parámetro de interés, g :

$$y_i = a_i g + b_i g^2, \quad (5.1)$$

donde los coeficientes a_i y b_i pueden variar entre canales.

El likelihood total toma entonces la forma de una suma de pérdidas cuadráticas:

$$L(g) = \sum_{i=1}^m (y_i - a_i g - b_i g^2)^2, \quad (5.2)$$

Representando, por ejemplo, el likelihood frente a g para un caso de dos dimensiones con coeficientes a y b fijos, se obtiene la siguiente figura.

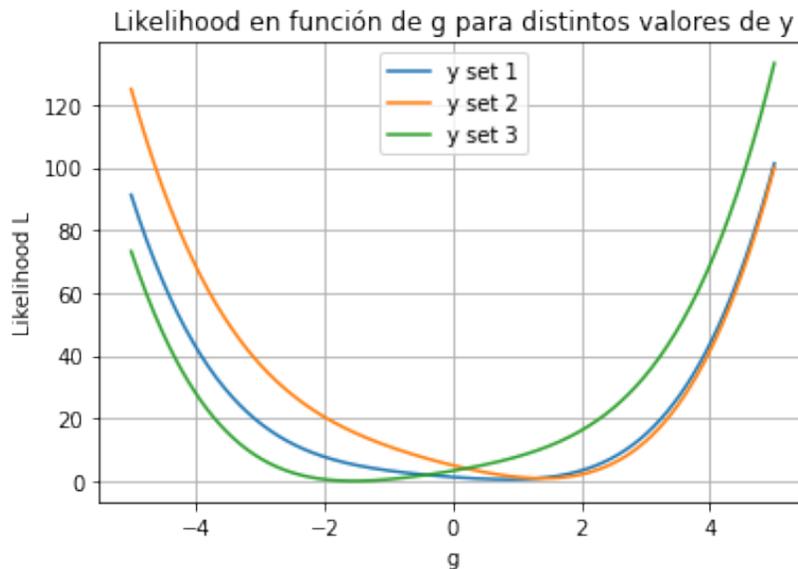


Figura 5.1: Representación del likelihood frente a g para diferentes valores de y

Esta forma permite aplicar directamente el método de Wilks, resolviendo las raíces del polinomio resultante con métodos numéricos. Dependiendo de la forma del polinomio, se pueden obtener uno o varios intervalos de confianza.

Aunque en este caso no es viable representar los intervalos frente a y debido a la dimensionalidad creciente, el cálculo de coverage se puede llevar a cabo de manera análoga a casos anteriores. Para ello, se generan valores aleatorios de $\vec{y} = (y_1, y_2, \dots, y_m)$ según distribuciones normales centradas en las predicciones teóricas $a_i g + b_i g^2$, y se verifica si el valor de g inyectado está contenido en el intervalo calculado.

Para el caso de dos dimensiones con coeficientes $a=[1, -0.5]$ y $b=[0.2, 0.3]$ se obtiene el siguiente coverage.

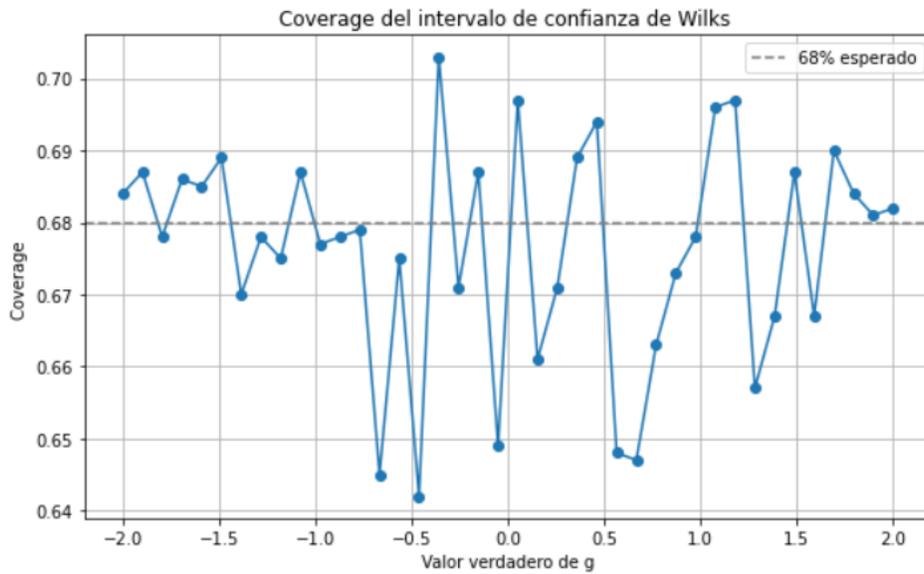


Figura 5.2: Coverage por el método de Wilks para el caso $a=[1, -0.5]$ y $b=[0.2, 0.3]$

Como se puede ver en esta figura 5.2 el método de Wilks nos da un coverage esperado cercano al 68 %. No obstante, se puede apreciar un ligero undercoverage en la zona central, en torno al 3 % cual no es alarmante y está dentro del error estadístico, teniendo en cuenta las simulaciones utilizadas.

Por otro lado, en cuanto al método de Feldman-Cousins, se consideraron dos estrategias. La primera consiste en aplicar el procedimiento estándar utilizando la variable g_{exp} , obtenida mediante la minimización de $L(g)$. Esta implementación es conceptualmente similar al caso unidimensional, aunque requiere evaluar una verosimilitud más compleja. Para el mismo caso analizado por el metodo de Wilks anteriormente, el coverage por este método es el representado a continuación.

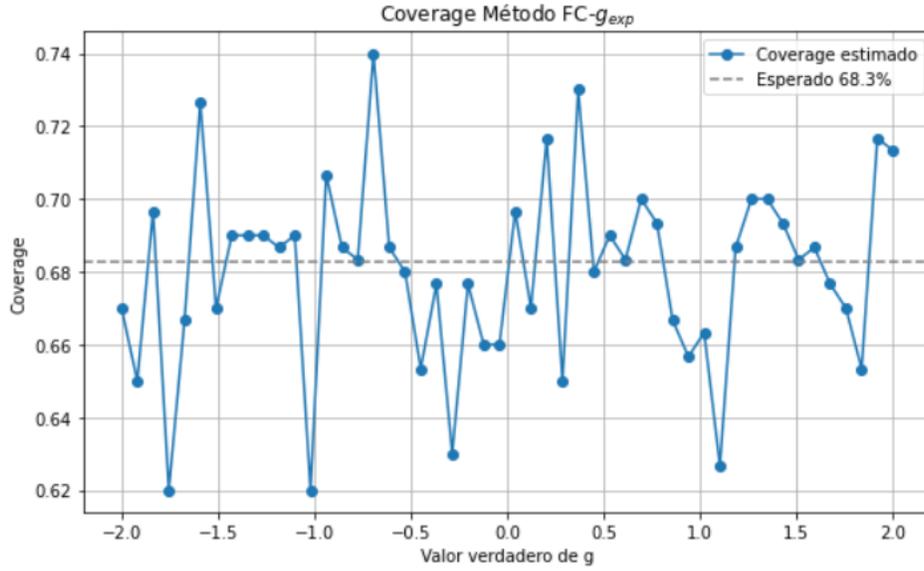


Figura 5.3: Coverage por el método de Feldman-Cousins basado en g_{exp} para el caso $a=[1, -0.5]$ y $b=[0.2, 0.3]$.

La segunda estrategia se basa en construir una función $R(g, \vec{y})$, definida como:

$$R(g, \vec{y}) = L(g; \vec{y}) - L_{\min}(\vec{y}), \quad (5.3)$$

donde $L_{\min}(\vec{y})$ representa el valor mínimo de la verosimilitud para una realización concreta de datos.

La cantidad $R(g, \vec{y})$ definida en la ecuación anterior se corresponde con una estadística de razón de verosimilitud (likelihood ratio), expresada como diferencia entre el likelihood evaluado en un valor genérico g y su mínimo global para los datos observados \vec{y} . Dado que esta construcción se basa en una minimización, $R(g, \vec{y}) \geq 0$ por definición, y se anula únicamente en el punto óptimo $g = \hat{g}$ que minimiza la pérdida total.

Este valor $R(g, \vec{y})$ cumple la misma función que la regla de ordenación en el método de Feldman-Cousins: permite ordenar los pseudo-experimentos simulados en función de su compatibilidad relativa con el valor de g considerado. Es decir, ordenar por $R(g, \vec{y})$ es estadísticamente equivalente a ordenar por la razón de verosimilitud entre g y \hat{g} , tal como se plantea en el enfoque original de Feldman y Cousins. Esta elección asegura que los intervalos construidos mantengan la cobertura deseada de forma coherente con los principios frecuentistas.

De esta manera, se generan múltiples simulaciones para cada valor de g , y se calcula $R_{\max}(g)$ como el valor crítico que delimita la región de aceptación con el nivel de confianza deseado. Dado un conjunto de observaciones, el intervalo de confianza para g se obtiene evaluando para qué valores se cumple $R(g, \vec{y}) < R_{\max}(g)$, lo cual se resuelve numéricamente mediante barrido e interpolación. Esta última condición refleja el hecho de que $R_{\max}(g)$ ha sido construido de forma que, para cada valor fijo de g , un $1 - \alpha$ de los pseudo-experimentos simulados satisfacen $R < R_{\max}(g)$. Por tanto, si los datos observados también cumplen $R(g, \vec{y}) < R_{\max}(g)$, se

consideran compatibles con ese valor de g al nivel de confianza especificado. El conjunto de todos esos g define el intervalo de confianza buscado.

Este enfoque, aunque más costoso computacionalmente, permite tener en cuenta la dependencia explícita de la verosimilitud con todos los datos observables y ofrece intervalos con propiedades de cobertura exactas, salvo problemas numéricos, en presencia de correlaciones entre canales.

El caso con coeficientes $a = [1, 1]$ y $b = [1, -1]$ representa el ejemplo más complejo e ilustrativo dentro de este trabajo, ya que combina todos los elementos estadísticos críticos que hemos venido analizando. En particular, define dos canales con dependencias funcionales distintas respecto al parámetro g :

$$y_1 = g + g^2, \quad y_2 = g - g^2$$

La primera función es estrictamente creciente en el dominio positivo (convexa), mientras que la segunda es creciente hasta $g = 0.5$ y decreciente después (cóncava). Esta diferencia de curvatura introduce una fuerte asimetría en la verosimilitud total, lo que da lugar a un likelihood con múltiples mínimos locales. Como resultado, la construcción del intervalo de confianza ya no puede considerarse unidimensional o monomodal: para ciertos valores de la observación \vec{y} , pueden aparecer múltiples regiones desconectadas del parámetro g compatibles con los datos al nivel de confianza deseado.

Desde el punto de vista estadístico, este tipo de estructura funcional exige métodos de construcción de intervalos que sean capaces de capturar regiones disjuntas del espacio de parámetros. En particular, los métodos basados en aproximaciones asintóticas, como Wilks, tienden a fallar en estos casos, ya que asumen una única región de máxima verosimilitud bien definida. En contraste, el método de Feldman-Cousins, al utilizar reglas de ordenación basadas en la razón de verosimilitud, permite identificar correctamente estos intervalos múltiples, siempre que se implemente adecuadamente en el espacio multicanal.

Este escenario es, por tanto, el banco de pruebas ideal para poner a prueba la robustez de los distintos métodos de construcción de intervalos desarrollados en este trabajo, así como la necesidad de validar empíricamente el coverage cuando se trabaja con modelos no lineales y multivariantes. De esta manera, el intervalo de confianza por este método se obtendría calculando los puntos de corte de la línea formada por los $R_{\text{máx}}(g)$ con la función, tal y como se ve a continuación.

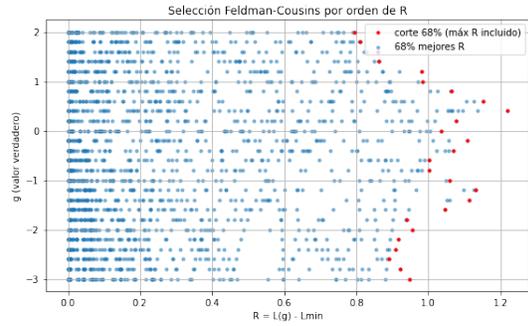


Figura 5.4: 68 % de los mejores valores de R para 100 simulaciones y 26 puntos.

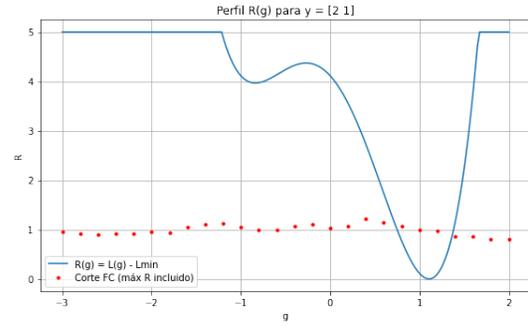


Figura 5.5: Intervalo de confianza para 100 simulaciones y 26 puntos de R_{max} .

Se puede observar que, si la línea de puntos coincidiera exactamente con el valor 1, el resultado sería consistente con la predicción asintótica de Wilks, sin embargo, se aprecian ligeras desviaciones respecto a ese valor. Este comportamiento es esperable y coherente con lo observado en otros estudios recientes, como en el caso de [3].

Procediendo como en el método anterior y sirviendo de comparativa entre métodos, se puede calcular aquí también el coverage, obteniendo resultados similares en ambos métodos.

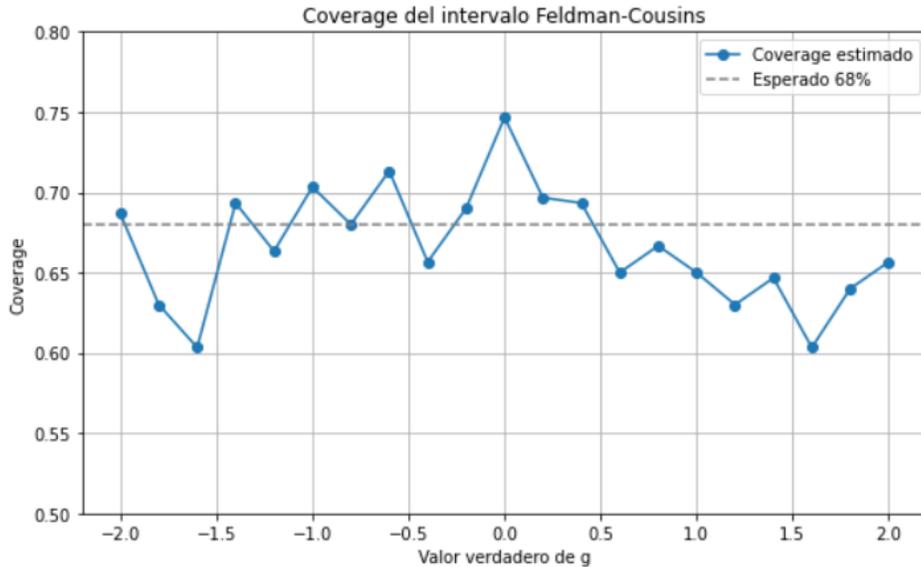


Figura 5.6: Coverage por el método de Feldman-Cousins basado en R_{max} para el caso $a=[1, 1]$ y $b=[1, -1]$

Aunque anteriormente se comentó que no era viable representar los intervalos de confianza en espacios de dimensión alta, en el caso bidimensional sí es posible visualizar la banda de aceptación en función de $\vec{y} = (y_1, y_2)$. Por ello, se ha incluido esta representación 3D del espacio (y_1, y_2, g) , donde se muestran los puntos que satisfacen la condición $R(g, \vec{y}) < R_{máx}(g)$ para un nivel de confianza dado. Aunque esta visualización no se generaliza fácilmente a espacios de

mayor dimensión, sí permite en este caso apreciar la estructura de la banda, que presenta una forma característica de "Y", y entender mejor la distribución de los valores de g compatibles con los datos observados.

Intervalos de confianza (68% mejores R) en espacio (y_1, y_2, g)

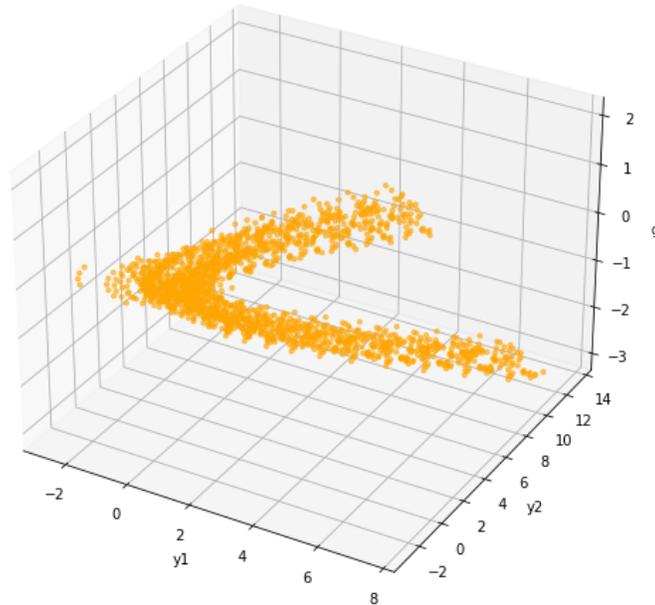


Figura 5.7: Representación 3D de los intervalos de confianza para el caso bidimensional $a=[1, 1]$ y $b=[1, -1]$.

5.1. Estudio del coverage para diferentes casos

Una vez se ha visto cómo funciona el método de Wilks y esta última versión del método de Felman-Cousins para obtener intervalos de confianza de una medida, es interesante ver cuál de todos nos garantiza mejores resultados. Una de las comparaciones que se pueden hacer para lograr este objetivo es el estudio del coverage. En este apartado, para un caso 2D y para diferentes valores de los parámetros a y b se realiza este estudio.

Para comenzar, es interesante estudiar el caso lineal, es decir $a=[1, 1]$ y $b=[0, 0]$.

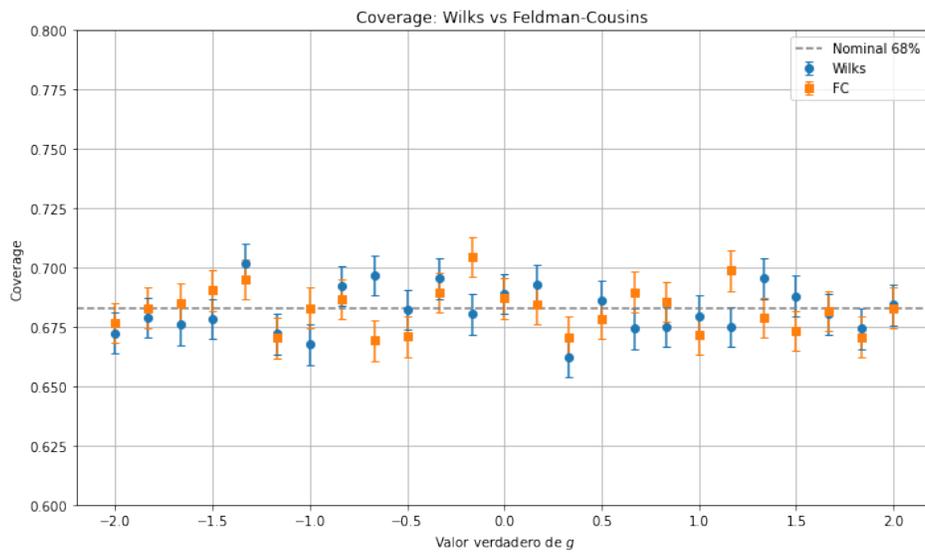


Figura 5.8: Comparación del coverage mediante el método de Wilks y Feldman-Cousins basado en y para el caso $a=[1, 1]$ y $b=[0, 0]$

Como se puede apreciar en esta figura 5.10, ambos métodos se ajustan bien al valor esperado de 68% tal y como era de esperar. Cabe destacar que las barras de error observadas son las correspondientes al error estadístico el cual se puede calcular como $\Delta = \sqrt{(c * (1 - c)/N)}$ donde c es el coverage y N el número de simulaciones ($N=3000$).

Otro caso de interés es el caso cuadrático, es decir $a=[0, 0]$ y $b=[1, 1]$.

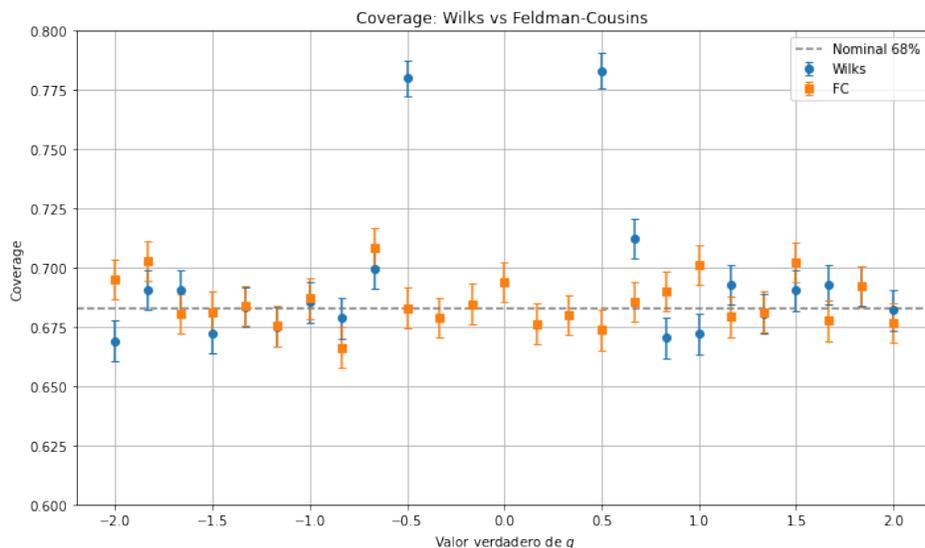


Figura 5.9: Comparación del coverage mediante el método de Wilks y Feldman-Cousins basado en y para el caso $a=[0, 0]$ y $b=[1, 1]$

Este es un caso verdaderamente interesante, ya que, como se puede ver, los métodos de Feldman-Cousins proporcionan el coverage esperado del 68 % en todo momento, mientras que el método de Wilks en la zona de valores centrales de g da una sobrestimación del mismo. Esto es debido a una exageración en la anchura del intervalo de confianza en la zona del mínimo, como se comentó anteriormente en la sección 4.3.

Por último, es interesante estudiar un caso que mezcle los dos anteriores como el caso $a=[1, 1]$ y $b=[1, -1]$. Para estos valores de los parámetros a y b la figura del coverage tendría la siguiente forma.

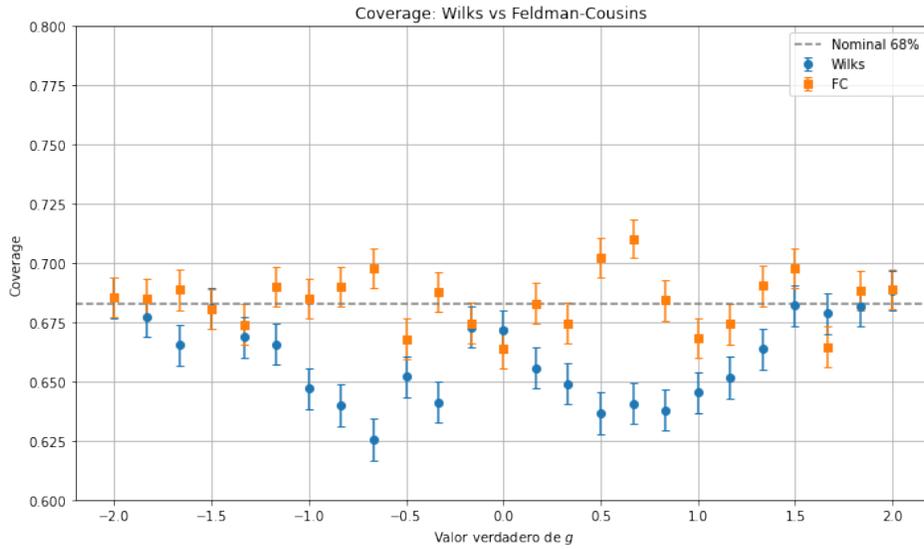


Figura 5.10: Comparación del coverage mediante el método de Wilks y Feldman-Cousins basado en y para el caso $a=[1, 1]$ y $b=[1, -1]$

En este último caso se puede ver que el método de Feldman Cousins sigue ofreciendo buenos resultados aunque se puede ver algunas ligeras fluctuaciones las cuales pueden deberse a un error de muestreo del R_{max} el cual es demasiado complicado de estimar analíticamente pero sabemos que existe. Los resultados que ofrecen el método de Wilks también presentan estas fluctuaciones, pero bastante más pronunciadas, además se puede apreciar una subestimación del coverage por parte de este método.

5.2. Estudio de la eficiencia computacional

Anteriormente se ha demostrado que métodos como el de Feldman-Cousins son más adecuados para garantizar coverage en modelos no lineales típicos de EFT, especialmente frente a las limitaciones del enfoque asintótico de Wilks. Sin embargo, esta mejora en robustez estadística viene acompañada de un coste computacional elevado, ya que requiere la simulación de un número considerable de pseudo-experimentos para construir la banda de aceptación y evaluar la cobertura empírica. En el contexto de EFT, donde la dimensionalidad del espacio

de parámetros puede ser alta y la evaluación del ajuste suele implicar modelos complejos, esta carga computacional puede convertirse en un cuello de botella real. Por ello, resulta fundamental estudiar si es posible reducir significativamente los parámetros de simulación, como el número de puntos en el barrido (M) o el número de réplicas (N), sin comprometer la validez estadística del procedimiento. Este capítulo analiza precisamente esa posibilidad, evaluando la estabilidad del coverage bajo diferentes configuraciones de simulación, con el objetivo de establecer criterios prácticos que permitan aplicar estos métodos de forma eficiente en análisis más realistas.

La parte central del algoritmo consiste en la generación de la banda de aceptación utilizada para construir los intervalos de confianza, en función del valor del parámetro teórico g . El procedimiento puede descomponerse en los siguientes pasos:

1. Se realiza un barrido uniforme en el espacio del parámetro teórico g , generando un conjunto de M valores distintos, $\{g_i^{\text{teo}}\}_{i=1}^M$.
2. Para cada valor g_i^{teo} , se simulan N experimentos independientes (pseudo-experimentos), obteniendo para cada uno un valor de la observación asociada (por ejemplo, un máximo local R_{obs}).
3. A partir del conjunto de N simulaciones para cada g_i^{teo} , se determina un valor representativo $R_{\text{max}}(g_i^{\text{teo}})$, típicamente el cuantil correspondiente al límite superior al $1 - \alpha$ de confianza. Dado que estos valores se calculan solo en los puntos discretos del barrido, se construye una función continua mediante interpolación. Este paso permite evaluar posteriormente $R_{\text{max}}(g)$ para valores de g intermedios, lo cual es esencial para garantizar que el procedimiento produzca intervalos continuos y precisos.

Aunque esta interpolación se ha utilizado sistemáticamente a lo largo del trabajo, su importancia se hace especialmente evidente en este capítulo, donde se analiza la eficiencia computacional al reducir el número de puntos del barrido (M). Con menos puntos simulados, el riesgo de obtener una banda de aceptación discontinua o poco ajustada aumenta considerablemente, por lo que contar con una interpolación robusta resulta crucial para preservar la calidad estadística del método.

4. Estos valores $R_{\text{max}}(g_i^{\text{teo}})$ se almacenan como una tabla de M puntos, que permite construir la banda de aceptación en todo el dominio del parámetro, utilizando una interpolación continua en g .
5. Esta banda se considera fija a partir de este punto, y se utiliza posteriormente para evaluar si nuevas observaciones caen dentro o fuera de la región de aceptación.
6. Finalmente, se realiza una segunda ronda de simulaciones (barrido y pseudo-experimentos) para calcular el coverage empírico del procedimiento, es decir, la fracción de intervalos que contienen el valor verdadero del parámetro.

El coste computacional de este algoritmo depende directamente de los parámetros M (número de puntos en el barrido teórico) y N (número de pseudo-experimentos por punto). Ambos pasos, la generación de la banda y la evaluación del coverage, implican simulaciones extensivas,

por lo que reducir M y N podría suponer un ahorro significativo de tiempo y recursos.

El objetivo del estudio de eficiencia es comprobar si el valor del coverage se mantiene estable, dentro de la incertidumbre estadística esperada, al disminuir considerablemente los valores de M y N . En otras palabras, se quiere verificar si el algoritmo produce intervalos con coverage correcto incluso cuando se calcula la banda con una resolución mucho más baja. Esta validación es fundamental para poder aplicar el método de forma eficiente en análisis más complejos o en exploraciones extensas del espacio de parámetros, donde el coste computacional se vuelve crítico.

Para conseguir esto estudiamos un ejemplo conocido que ya se ha tratado anteriormente, $a=[1,1]$ y $b=[1, -1]$.

Se puede comenzar con los resultados contenidos en las figuras 5.4 y 5.5. En este caso se utilizaron 26 puntos de R_{max} realizando 100 simulaciones para obtener cada uno.

Bajando el número de puntos de R_{max} drásticamente a 6 pero manteniendo el número de simulaciones para obtener cada uno obtenemos lo siguientes resultados.

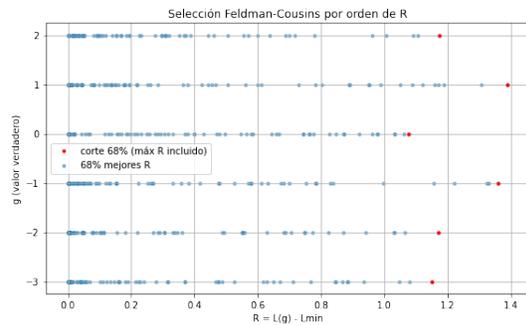


Figura 5.11: 68 % de los mejores valores de R para 100 simulaciones y 6 puntos.

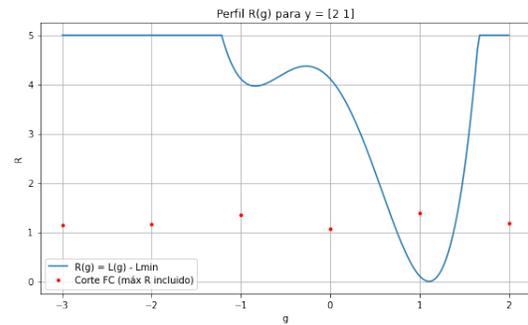


Figura 5.12: Intervalo de confianza para 100 simulaciones y 6 puntos de R_{max} .

Estos resultados a simple vista no parecen estar muy alejados a los obtenidos con el número de simulaciones originales aunque luego se comprobará su fiabilidad estudiando el coverage. Para el último caso mantenemos el número de puntos de R_{max} en 6 pero ahora bajamos a 10 las simulaciones para obtener cada uno, dando como resultado las figuras mostradas a continuación, las cuales sí que parecen ofrecer resultados peores.

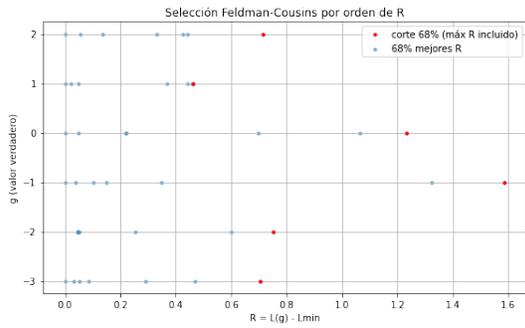


Figura 5.13: 68% de los mejores valores de R para 10 simulaciones y 6 puntos.

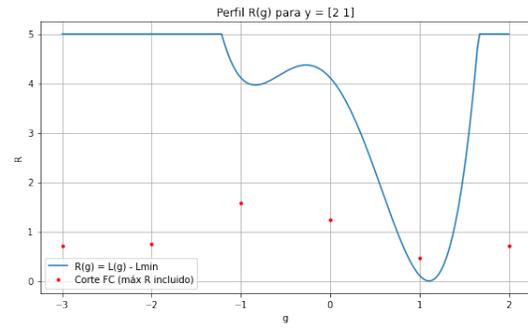


Figura 5.14: Intervalo de confianza para 10 simulaciones y 6 puntos de R_{max} .

Una vez tenemos identificados los tres casos de estudio, es interesante graficarlos en una misma figura para observar mejor las diferencias.

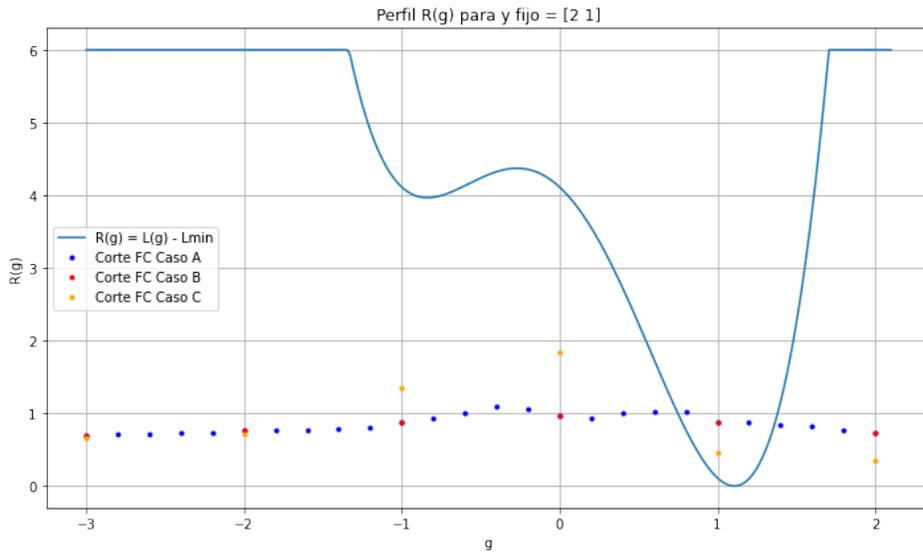


Figura 5.15: Comparativa de los intervalos de confianza para diferentes casos de número de simulaciones y puntos.

Se puede apreciar cómo el caso A (100 simulaciones, 26 puntos) y B (100 simulaciones, 6 puntos) ofrecen resultados similares a simple vista, mientras que el caso C (10 simulaciones, 6 puntos) difiere bastante de estos dos últimos.

No obstante, para comparar bien la calidad de los resultados al variar el número de simulaciones y puntos, se debe comparar el coverage que proporciona cada uno, tal y como se ve en la siguiente figura.

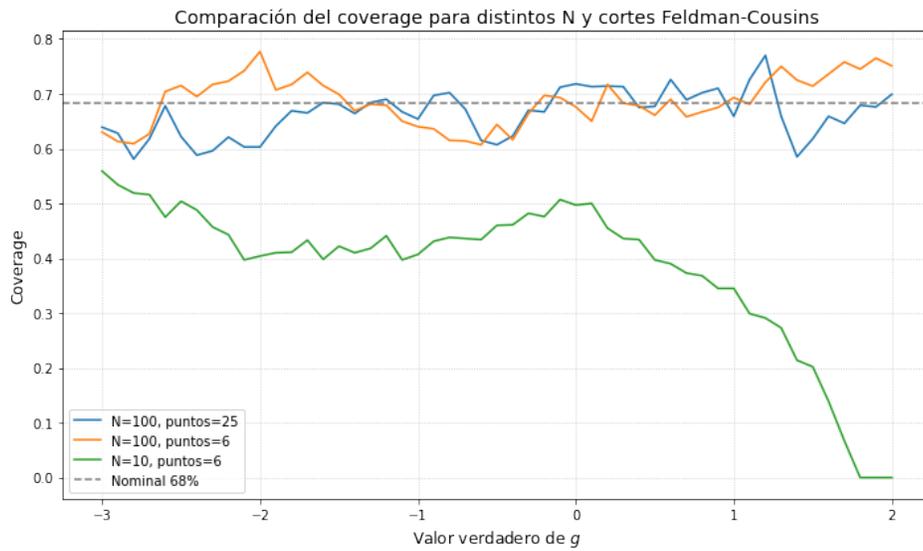


Figura 5.16: Comparativa de los coverages para diferentes casos de número de simulaciones y puntos

En conjunto, estos resultados permiten extraer algunas recomendaciones prácticas. En primer lugar, el número de pseudo-experimentos por punto (N) tiene un impacto directo sobre la estabilidad del procedimiento: valores muy bajos, como $N = 10$, generan fluctuaciones estadísticas inaceptables en la estimación de la banda de aceptación. En cambio, con $N = 100$ ya se obtiene una buena estabilidad, a pesar de tratarse de un valor relativamente bajo, lo que demuestra que el método es robusto incluso con recursos limitados.

En cuanto al número de puntos del barrido (M), se observa que una discretización con tan solo 6 valores de g distribuidos uniformemente es suficiente para recuperar correctamente el intervalo de confianza, siempre que se emplee una interpolación adecuada. Este resultado es especialmente relevante si se considera que, en aplicaciones reales, cada simulación puede implicar una evaluación compleja y costosa, donde el tiempo de cómputo total escala con $M \times N$. Por tanto, reducir ese producto de miles a cientos de simulaciones puede hacer viable la aplicación del método en contextos físicos realistas, sin comprometer su validez estadística.

CAPÍTULO 6

CONCLUSIONES

El presente trabajo ha abordado la construcción de intervalos de confianza en el contexto de teorías de campo efectivo (EFT), prestando especial atención a las implicaciones estadísticas derivadas de la estructura funcional no lineal que caracteriza a estos modelos. A lo largo del documento se han estudiado distintos enfoques para la estimación de parámetros, desde métodos asintóticos ampliamente utilizados, como el de Wilks, hasta construcciones más robustas y exactas, como las de Neyman y Feldman-Cousins.

Los resultados obtenidos ponen de manifiesto las limitaciones del uso directo del teorema de Wilks en contextos donde las condiciones asintóticas o de linealidad no se cumplen. En particular, se ha observado que, para modelos con dependencia cuadrática o combinada del parámetro, este método presenta desviaciones significativas en coverage, con tendencia a la sobreestimación en regiones cercanas al mínimo. Aunque su aplicabilidad es razonable en modelos lineales bien comportados, su fiabilidad decrece cuando se introducen términos no lineales o degeneraciones funcionales.

Por otro lado, el método de Neyman ha demostrado ser un método válido para asegurar el coverage nominal, especialmente cuando se dispone de herramientas numéricas que permiten su implementación empírica mediante simulaciones Monte Carlo, aunque no siempre es del todo práctico. Por otro lado, su flexibilidad puede volverse una desventaja si la regla de ordenación no se elige adecuadamente, lo que comprometería la calidad de los intervalos resultantes.

En este sentido, el método de Feldman-Cousins se presenta como una solución particularmente atractiva. Su regla de ordenación basada en la razón de verosimilitud permite una construcción objetiva y consistente de los intervalos, que se adapta automáticamente a la simetría del problema y evita decisiones arbitrarias sobre la unilateralidad o bilateralidad del intervalo. Esta característica se ha revelado especialmente relevante en los modelos con degeneraciones, como el caso $y \sim g^2$, o en aquellos con múltiples canales y curvaturas opuestas, como el ejemplo con coeficientes $a = [1, 1]$ y $b = [1, -1]$ discutido en el capítulo 5.

Este último caso ha resultado ser el más representativo de las dificultades estadísticas en EFT, combinando no linealidades, degeneraciones y estructuras multicanal que generan múltiples soluciones locales. La implementación del método de Feldman-Cousins en este entorno ha requerido un tratamiento cuidadoso del procedimiento, incluyendo la elección de la variable de ordenación (donde se ha demostrado que el uso de y es preferible frente a g_{exp}), así como un análisis detallado del comportamiento del coverage. El éxito de este enfoque en capturar correctamente regiones disjuntas del parámetro valida su aplicabilidad en contextos físicos reales donde la interpretación correcta de los intervalos es crítica.

Además, se ha estudiado la viabilidad práctica de estos métodos en términos computacionales, evaluando el impacto de reducir parámetros clave como el número de puntos del barrido o el número de pseudo-experimentos por valor de g . Los resultados obtenidos sugieren que es posible alcanzar un compromiso razonable entre eficiencia computacional y fidelidad estadística, siempre que se controlen adecuadamente las fuentes de error numérico.

En conjunto, este trabajo ofrece una visión completa de las herramientas disponibles para la construcción de intervalos de confianza en EFT, resaltando la necesidad de validar empíricamente el coverage cuando se trabaja fuera del régimen asintótico. Esta conclusión es coherente con estudios recientes realizados por la colaboración CMS en el análisis SMP-24-003 [3], así como con el estudio teórico presentado en [1] donde se señala que el teorema de Wilks no es aplicable en presencia de términos cuadráticos dominantes en EFT.

Como perspectiva futura, sería interesante extender este análisis a modelos multivariantes de mayor dimensión, así como estudiar el impacto de parámetros no deseados (nuisance parameters) mediante técnicas de perfilado o marginalización. Además, la incorporación de incertidumbres sistemáticas y la evaluación de su efecto sobre el coverage constituye un paso necesario hacia aplicaciones más realistas en el ámbito de la física de partículas.

Referencias

- [1] Florian U. Bernlochner et al. “Cover Your Bases: Asymptotic Distributions of the Profile Likelihood Ratio When Constraining Effective Field Theories in High-Energy Physics”. En: *Phys. Rev. D* 107 (2023), pág. 056008. DOI: [10.1103/PhysRevD.107.056008](https://doi.org/10.1103/PhysRevD.107.056008). arXiv: [2207.01350](https://arxiv.org/abs/2207.01350) [[hep-ph](#)].
- [2] I. Brivio et al. “SMEFTsim: a practical guide to SMEFT at the LHC”. En: *JHEP* 02 (2021), pág. 095.
- [3] CMS Collaboration. “Combined effective field theory interpretation of Higgs boson, electroweak vector boson, top quark, and multi-jet measurements”. En: *arXiv preprint* (2025). Submitted to EPJC. arXiv: [2504.02958](https://arxiv.org/abs/2504.02958) [[hep-ph](#)]. URL: <https://arxiv.org/abs/2504.02958>.
- [4] CMS Collaboration. “Search for physics beyond the standard model in top quark production with additional leptons in the context of effective field theory”. En: *JHEP* 2023.12 (2023), pág. 68. DOI: [10.1007/JHEP12\(2023\)068](https://doi.org/10.1007/JHEP12(2023)068).
- [5] Glen Cowan. *Statistical Data Analysis*. Oxford University Press, 1998.
- [6] K. Cranmer. *Practical Statistics for the LHC*. <https://arxiv.org/abs/physics/0511028>. 2008.
- [7] Kyle Cranmer. *Confidence Intervals Lecture*. 2008. URL: https://indico.cern.ch/event/35523/contributions/1768686/attachments/704225/966801/Cranmer_CI_lecture.pdf.
- [8] G.J. Feldman y R.D. Cousins. “Unified approach to the classical statistical analysis of small signals”. En: *Phys. Rev. D* 57 (1998), págs. 3873-3889. DOI: [10.1103/PhysRevD.57.3873](https://doi.org/10.1103/PhysRevD.57.3873).
- [9] B. Grzadkowski et al. “Dimension-six terms in the Standard Model Lagrangian”. En: *JHEP* 2010 (2010), pág. 085.
- [10] Joaquín Pardo. *Códigos del Trabajo de Fin de Grado*. https://github.com/wronjol/Codigos_Tfg. Accedido el 26 de junio de 2025. 2025.
- [11] S. S. Wilks. “The large-sample distribution of the likelihood ratio for testing composite hypotheses”. En: *Ann. Math. Stat.* 9.1 (1938), págs. 60-62.