



CAPITULO VI

INCERTIDUMBRES SISTEMATICAS



Incertidumbres sistemáticas (I)

En el MLE, la matriz de covariancia es el estimador de las incertidumbres estadísticas.

Pero otras fuentes de incertidumbre contribuyen también a diluir la precisión de una medida. En lenguaje de física, éstas a menudo se llaman “errores sistemáticos”. Para discutir de estas incertidumbres en el contexto MLE, modificamos ligeramente la notación, y reescribimos la función de verosimilitud como sigue:

$$\mathcal{L}(\mu_1, \dots, \mu_p, \theta_1, \dots, \theta_k; \vec{x}) ,$$

donde hemos explícitamente separado el conjunto de parámetros de \mathcal{L} en dos subconjuntos:

- ▶ los *parámetros de interés* μ_1, \dots, μ_p , (POI en inglés), son las cantidades que deseamos estimar;
- ▶ los *parámetros de molestia* $\theta_1, \dots, \theta_k$ (NP en inglés, por nuisance parameters) que representan fuentes potenciales de sesgos sistemáticos;

La lógica es la siguiente: si le asignásemos valores imprecisos o errados a algunos NPs, la formas de las PDFs resultantes pueden distorsionarse, y los estimadores de los POIs resultarán sesgados.

El error sistemático cuantifica la magnitud de esos sesgos.

- ▶ A menudo, la estimación de los sistemáticos es parte la dominante del trabajo de análisis de datos
- ▶ La calidad de un resultado científico releva en gran parte de la calidad de la estimación de la componente sistemática de su incertidumbre
- ▶ Las medidas experimentales son a menudo clasificados en dos categorías:
 - ▶ “dominadas por la estadística”, $\sigma_{\text{STAT}} > \sigma_{\text{SYST}}$
 - ▶ “dominadas por los sistemáticos”, $\sigma_{\text{STAT}} < \sigma_{\text{SYST}}$



Incertidumbres sistemáticas (II)

Las incertidumbres sistemáticas debidas a los NPs se clasifican a menudo en dos categorías:

- ▶ Errores de “Tipo-I” : nuestra muestra (u otras muestras de control) pueden en principio proveer información sobre el NP considerado, y la incertidumbre proveniente de este NP debe en principio disminuir con el tamaño de las muestras utilizadas.
- ▶ Errores de “Tipo-II”, que provienen de suposiciones incorrectas del modelo (p.e. el uso de funciones inadecuadas para las PDFs), u aspectos mal controlados en los datos, como cambios en las condiciones de adquisición de las muestras, o la presencia de especies no tomadas en cuenta.
 - ▶ Algunas personas lo separan en dos categorías: los Tipo-II “de modelo” y los Tipo-III “de teoría” ... lo cual da “The Good, the Bad, the Ugly”.

Algunos comentarios:

- ▶ Ciertos sistemáticos son una mezcla de ambos tipos: por ejemplo, si las muestras de control para controlar un NP no son completamente representativas de las propiedades de ese NP.
- ▶ Un NP enteramente de Tipo-I es una variable aleatoria, y por tanto se pueden estimar intervalos de confianza usando los métodos antes descritos.
- ▶ Los errores de Tipo-II son difíciles de manejar correctamente, y no siempre hay un procedimiento bien definido y consensual para estimar el impacto de esos errores. En los peores casos, no siempre se puede saber de qué manera un error de Tipo-II impacta los intervalos de confianza sobre los POIs.
- ▶ Frente a casos ambiguos, hay cierto consenso (parcial) en que es mejor tener estimaciones “conservadoras” de los sistemáticos, en particular cuando contribuyen de manera subdominante al “error budget”. Pero en esas situaciones, las discusiones suelen ser controversiales, incluso amargas...



El método de *profile-likelihood*, (¿verosimilitud de perfil?) permite manejar de manera elegante los sistemáticos de Tipo-I.

Consiste en asignar una verosimilitud específica a los NPs “perfileables”, de tal manera que la \mathcal{L} original se modifica en dos componentes:

$$\mathcal{L}(\mu, \theta) = \mathcal{L}_\mu(\mu, \theta) \mathcal{L}_\theta(\theta) .$$

Entonces, para un valor fijo de μ , \mathcal{L} es maximizada con respecto al NP θ . Si se recorre una secuencia de valores de μ , \mathcal{L} es una función que solamente depende de μ ; si dice que la molestia ha sido “profiled-out”.

Supongamos que el valor del NP es conocido: $\theta = (\theta_0 \pm \Delta_0)$.

- ▶ en ese caso la penalidad $\mathcal{L}_\theta(\theta)$ corresponde a una Gaussiana,

$$\mathcal{L}_\theta(\theta) = P_{\text{Gauss}(\theta; \theta_0, \Delta_0)} = \frac{1}{\sqrt{2\pi}\Delta_0} e^{-(\theta - \theta_0)^2 / 2\Delta_0^2} ,$$

- ▶ la maximización de $\mathcal{L}(\mu, \theta)$ “empujará” el estimador $\hat{\theta}$ del NP hacia su valor nominal θ_0
- ▶ el “pull” $(\hat{\theta} - \theta_0) / \Delta_0$ es un indicador de la coherencia del modelo; grandes valores (positivos o negativos) del *pull* indican una “tensión” proveniente de ese NP.



Incertidumbres sistemáticas (IV)

Como ejemplo en física, consideremos la medida de una sección eficaz de un proceso σ (inicial \rightarrow final).

- ▶ El número de eventos observados de ese tipo permite acceder a la sección eficaz:

$$N = \sigma \int \mathcal{L} dt .$$

- ▶ Si solamente una fracción de los procesos de ese tipo son detectados (p.e. debido a efectos de *aceptancia geométrica* del detector, u otras fuentes de ineficiencia), se requiere conocer la eficiencia de reconstrucción ε para convertir el número observado de procesos \hat{N} en una medida de $\hat{\sigma}$:

$$\frac{\hat{N}}{\varepsilon} = \hat{\sigma} \int \mathcal{L} dt .$$

- ▶ La eficiencia ε es claramente un NP: independientemente de la precisión con la cual se haya medido \hat{N} ,
 - ▶ un valor incorrecto de ε sesga directamente la medida de $\hat{\sigma}$
 - ▶ la incertidumbre sobre ε se propaga directamente sobre la incertidumbre sobre $\hat{\sigma}$
- ▶ Si $\hat{\varepsilon}$ puede estimarse sobre una muestra de control de calidad (por ejemplo, una simulación con gran estadística, o una muestra de control de alta pureza), el impacto de la molestia será controlado de manera coherente y su propagará correctamente sobre la medida del PIO.
- ▶ Un análisis *elegante* efectuará un *ajuste simultáneo a las muestras de señal y control*, de tal manera que
 - ▶ los valores y las incertidumbres de los NPs se extraen directamente de la covariancia del ajuste ML
 - ▶ lo cual asegura que las correlaciones con los POIs se propagan correctamente
 - ▶ y que los intervalos de confianza heredan de todas esas relaciones entre POIs y NPs.



Otro ejemplo de “profile-likelihood” : la búsqueda de una resonancia (un “bump”) sobre un fondo uniforme.

- ▶ Si la fracción de señal (pureza) es muy pequeña, la anchura Γ del bump no puede estimarse directamente sobre la muestra
- el valor de anchura a utilizar en la PDF de señal debe provenir de fuentes externas (p.e. una simulación detallada de la función de resolución del detector).
- ▶ Esa anchura es claramente un parámetro de molestia:
- si su valor se sobreestima, eso se traduciría en una subestimación del cociente señal/fondo, y por tanto en un aumento de la varianza de los POIs de la señal, así como tal vez posibles sesgos en sus valores centrales (p.e. si el fondo es asimétrico, el impacto será diferente a ambos lados del pico de la señal)
- ▶ Si las muestras de control permiten acceder a una estimación independiente de la anchura $\hat{\Gamma} \pm \hat{\sigma}_{\Gamma}$ de la anchura del pico de señal, esta información puede implementarse utilizando una PDF Gaussiana, de media $\hat{\Gamma}$ y de anchura $\hat{\sigma}_{\Gamma}$, en la componente \mathcal{L}_{Γ} de la verosimilitud.
- Este término adicional cumple el rol de una penalidad en el MLE, y por lo tanto “empuja” los valores de Γ adentro del intervalo dado por $\pm \hat{\sigma}_{\Gamma}$



Incertidumbres sistemáticas (VI)

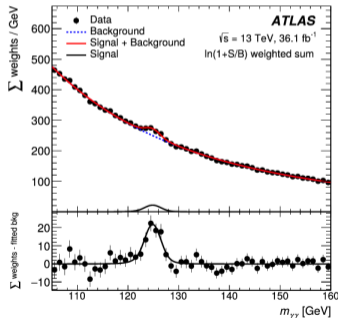
Un ejemplo práctico del “profile-likelihood” : el análisis del bosón de Higgs en el canal difotón en el LHC. Este canal se caracteriza por un cociente S/B extremadamente bajo.

Los parámetros de interés son :

- ▶ la masa del Higgs,
- ▶ el producto $\sigma(pp \rightarrow H + X) \times \mathcal{BR}(H \rightarrow \gamma\gamma)$ de la sección eficaz de producción del Higgs en colisiones protón-protón, multiplicada por la tasa de decaimiento del Higgs en el canal difotón.

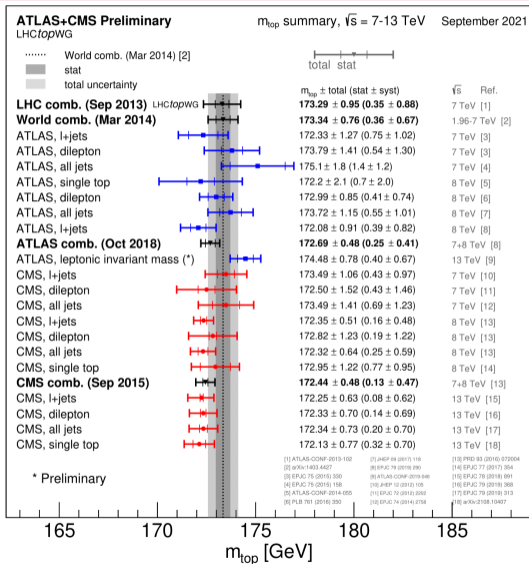
Mientras que la (larga) lista de parámetros de molestia incluye, entre otros:

- ▶ La eficiencia $\varepsilon(H \rightarrow \gamma\gamma)$ de reconstrucción de la señal, es decir el cociente entre el número de eventos $H(\rightarrow \gamma\gamma)_{\text{RECO}}$ registrados, reconstruidos, identificados y seleccionados, y el número de eventos $(H \rightarrow \gamma\gamma)_{\text{TRUE}}$ realmente producidos,
- ▶ Los parámetros necesarios para caracterizar la resolución en masa invariante difotón, es decir la distribución de la diferencia entre la masa reconstruida $m(\gamma\gamma)_{\text{RECO}}$ y la verdadera masa $m(\gamma\gamma)_{\text{TRUE}}$,
- ▶ Los parámetros necesarios para caracterizar la forma del fondo. Para simplificar, supondremos que una función uni-paramétrica es suficiente (p.e. una función exponencial decreciente)





Reportar las incertidumbres sistemáticas: ejemplos



Source	Uncertainty [%]
Statistical uncertainty	7.5
Systematic uncertainties	6.4
Background modelling (spurious signal)	3.8
Photon energy scale & resolution	3.6
Photon selection efficiency	2.6
Luminosity	1.8
Pile-up modelling	1.4
Trigger efficiency	1.0
Theoretical modelling	0.4
Total	9.8

Objetivo :
 Dar tanta información detallada como sea posible...