

Física Experimental / Laboratorio I

Clase #2

Análisis estadístico de incertezas aleatorias

Algunas definiciones importantes:

Valor medio, o "media" de un conjunto de mediciones x_1, \dots, x_N de la *misma* cantidad física:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{N} (x_1 + \dots + x_N)$$

Análisis estadístico de incertezas aleatorias

Algunas definiciones importantes:

Valor medio, o "media" de un conjunto de mediciones x_1, \dots, x_N de la *misma* cantidad física:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{N} (x_1 + \dots + x_N)$$

Desviación standard de un conjunto de mediciones x_1, \dots, x_N de la *misma* cantidad física:

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (d_i)^2} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Donde $d_i = x_i - \bar{x}$ es la desviación de los datos respecto al promedio. El promedio de d_i no hubiera sido una buena medida de la incerteza, ya que por definición

$$\bar{d} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) = \frac{1}{N} \sum_{i=1}^N x_i - \frac{1}{N} N \bar{x} = \bar{x} - \bar{x} = 0$$

Histogramas: muestra de cantidades enteras

Supongamos que hacemos un experimento consistente en N determinaciones de la misma cantidad física x_1, \dots, x_N , la cual **solamente puede tomar valores enteros**. Entonces:

$$\bar{x} = \frac{\sum_k x_k n_k}{N}$$

n_k : cantidad de veces que se ha encontrado el valor x_k . Notar que $\sum_k n_k = N$.

Definimos la fracción de veces en que ocurre un cierto valor de x_k , como

$$F_k = \frac{n_k}{N}$$

Histogramas: muestra de cantidades enteras

Supongamos que hacemos un experimento consistente en N determinaciones de la misma cantidad física x_1, \dots, x_N , la cual **solamente puede tomar valores enteros**. Entonces:

$$\bar{x} = \frac{\sum_k x_k n_k}{N}$$

n_k : cantidad de veces que se ha encontrado el valor x_k . Notar que $\sum_k n_k = N$.

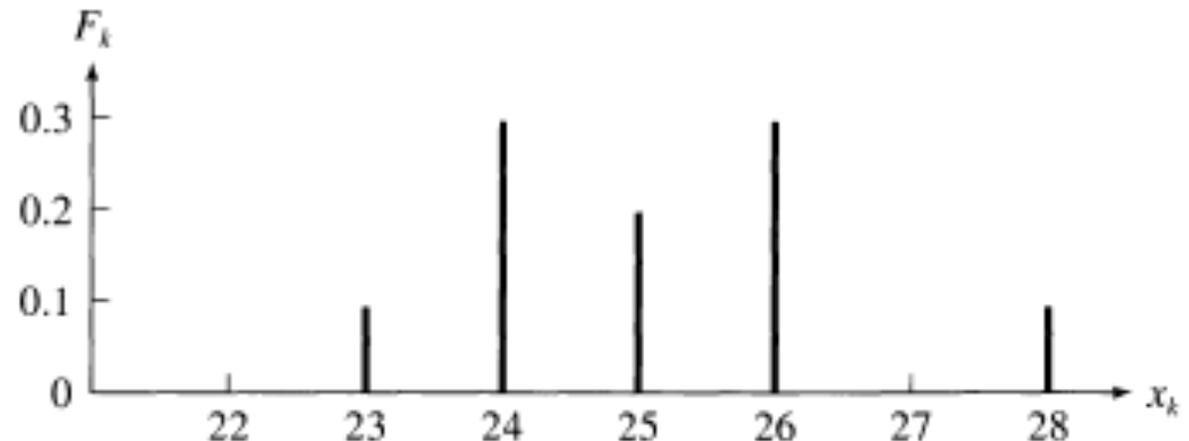
Definimos la fracción de veces en que ocurre un cierto valor de x_k , como

$$F_k = \frac{n_k}{N}$$

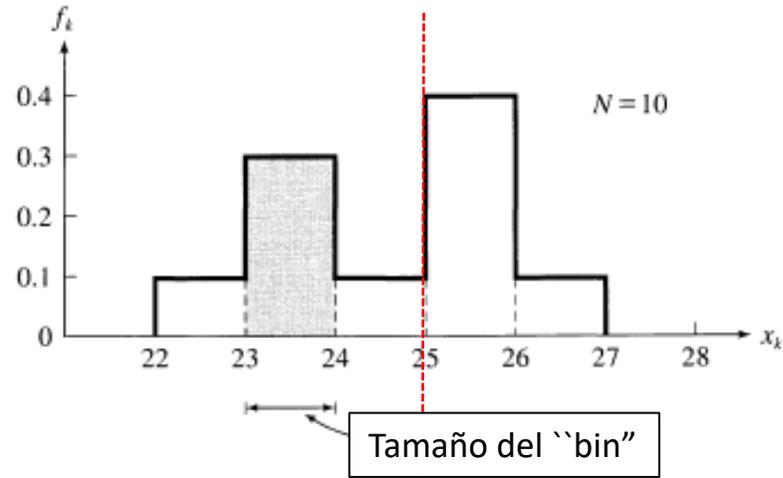
F_k describe la distribución de los resultados obtenidos:

$$\bar{x} = \sum_k x_k F_k$$

$$\sum_k F_k = 1$$



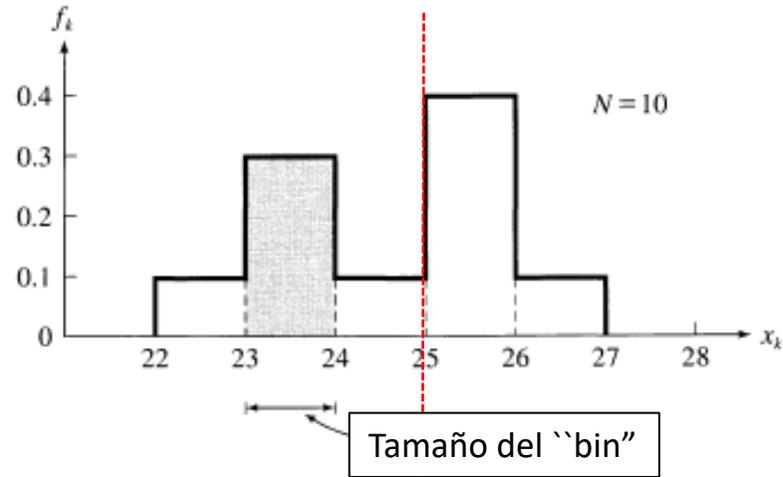
Histogramas: muestra de cantidades reales



Los x_k puede tomar valores reales

“bin”: agrupamiento en entornos limitados de valores.

Histogramas: muestra de cantidades reales



Los x_k puede tomar valores reales

“bin”: agrupamiento en entornos limitados de valores.

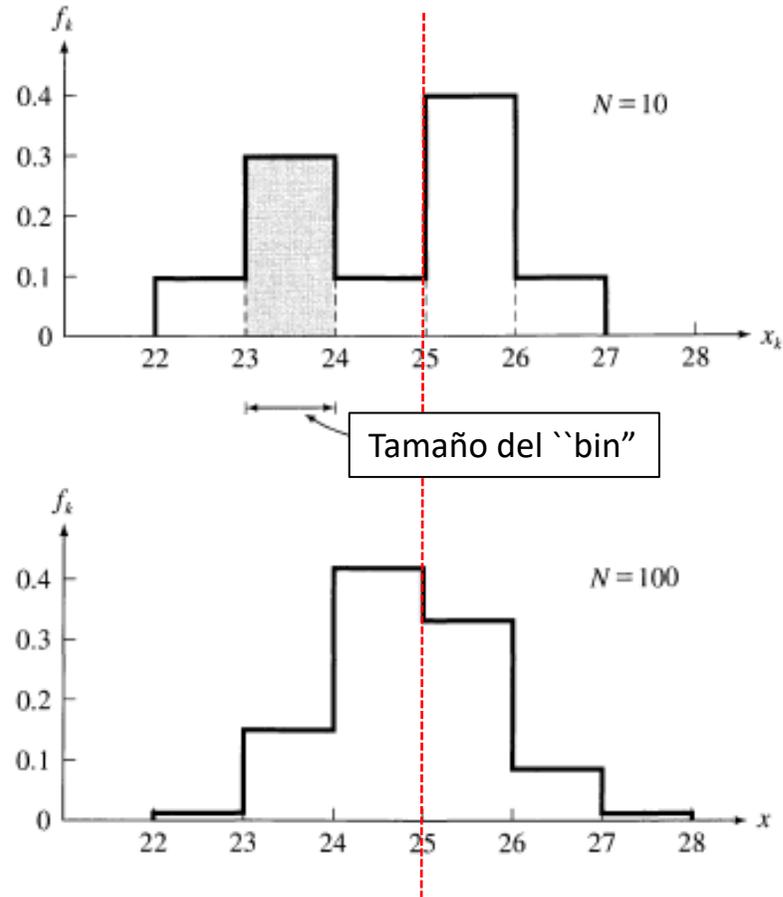
Δk : ancho del bin k-ésimo

f_k : altura del rectángulo que se dibuja en este bin

$f_k \Delta k$: fracción de medidas observadas en el bin k.

Esta área tiene el mismo significado que la altura F_k en el gráfico de barras.

Distribución normal o de Gauss: muestra de cantidades reales



Los x_k puede tomar valores reales

“bin”: agrupamiento en entornos limitados de valores.

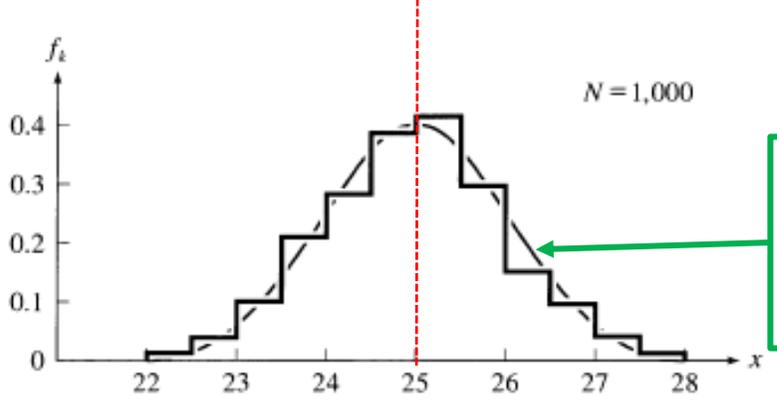
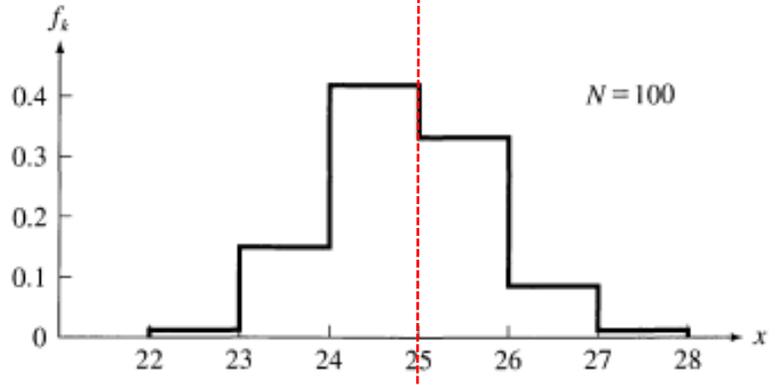
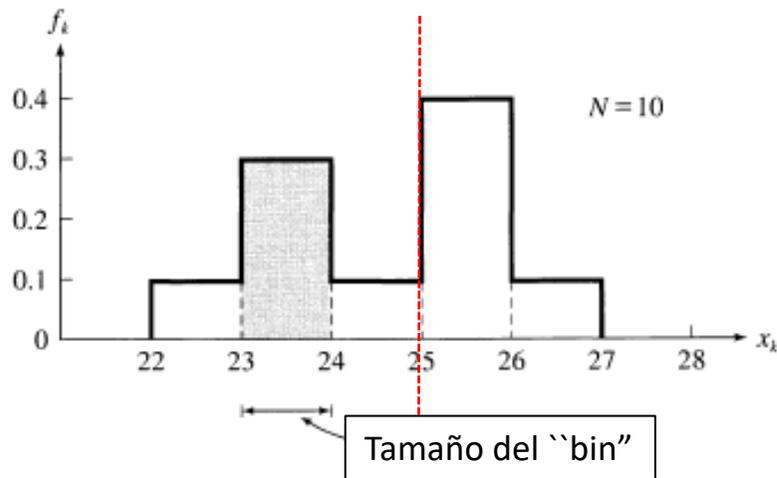
Δk : ancho del bin k-ésimo

f_k : altura del rectángulo que se dibuja en este bin

$f_k \Delta k$: fracción de medidas observadas en el bin k.

Esta área tiene el mismo significado que la altura F_k en el gráfico de barras.

Distribución normal o de Gauss: muestra de cantidades reales



Los x_k puede tomar valores reales

“bin”: agrupamiento en entornos limitados de valores.

Δk : ancho del bin k-ésimo

f_k : altura del rectángulo que se dibuja en este bin

$f_k \Delta k$: fracción de medidas observadas en el bin k.

Esta área tiene el mismo significado que la altura F_k en el gráfico de barras.

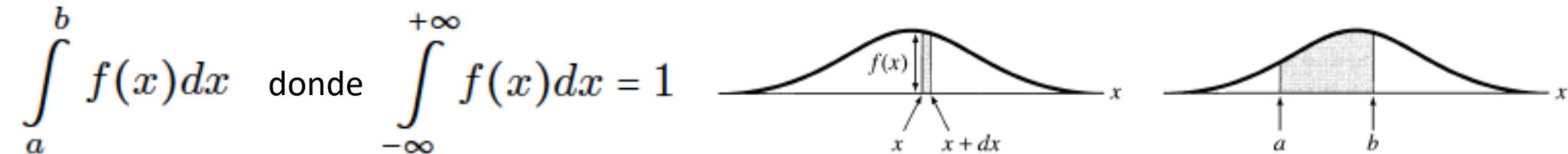
Tiende a una curva continua $f(x)$ para $N \rightarrow$ infinito: la distribución “normal” o de Gauss (*teorema central del límite*)

Distribuciones “límites”

Área $f(x)dx$: fracción de mediciones que se encuentran entre x y $x + dx$.

Cumple el mismo rol de F_k : es la “probabilidad” de que una medición cualquiera resulte en un valor entre x y $x + dx$.

De manera más general, la fracción de mediciones que caen entre los valores $x = a$ y $x = b$ es



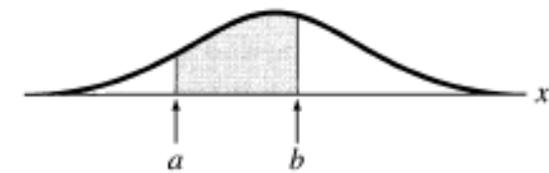
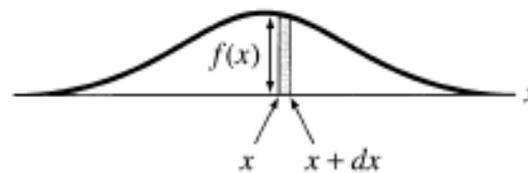
Distribuciones “límites”

Área $f(x)dx$: fracción de mediciones que se encuentran entre x y $x + dx$.

Cumple el mismo rol de F_k : es la “probabilidad” de que una medición cualquiera resulte en un valor entre x y $x + dx$.

De manera más general, la fracción de mediciones que caen entre los valores $x = a$ y $x = b$ es

$$\int_a^b f(x)dx \quad \text{donde} \quad \int_{-\infty}^{+\infty} f(x)dx = 1$$



En este caso entonces resulta:

$$\bar{x} = \int_{-\infty}^{+\infty} x f(x) dx$$

Valor medio

$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x - \bar{x})^2 f(x) dx$$

Desviación standard (σ_x)

Distribución normal o de Gauss

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}$$

Distribución normal o de Gauss

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}$$

Con esta definición, el valor medio resulta ser: $\bar{x} = \int_{-\infty}^{+\infty} xG_{X,\sigma}(x)dx$

Haciendo el cambio de variables $y = x - X$, de donde $dx = dy$, y $x = y + X$, resulta

$$\bar{x} = \frac{1}{\sigma\sqrt{2\pi}} \left(\int_{-\infty}^{+\infty} ye^{-y^2/2\sigma^2} dy + X \int_{-\infty}^{+\infty} e^{-y^2/2\sigma^2} dy \right)$$

(Note: A red diagonal line is drawn through the first integral, with a red "= 0" written above it. A blue "= σ√2π" is written above the second integral.)



$$\bar{x} = X$$



Distribución normal o de Gauss

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}$$

Con esta definición, el valor medio resulta ser: $\bar{x} = \int_{-\infty}^{+\infty} xG_{X,\sigma}(x)dx$

Haciendo el cambio de variables $y = x - X$, de donde $dx = dy$, y $x = y + X$, resulta

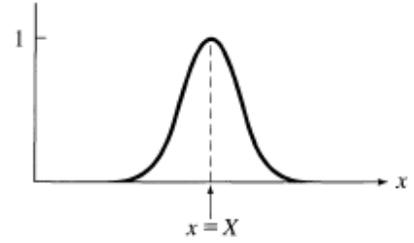
$$\bar{x} = \frac{1}{\sigma\sqrt{2\pi}} \left(\int_{-\infty}^{+\infty} ye^{-y^2/2\sigma^2} dy + X \int_{-\infty}^{+\infty} e^{-y^2/2\sigma^2} dy \right)$$

= 0

= $\sigma\sqrt{2\pi}$



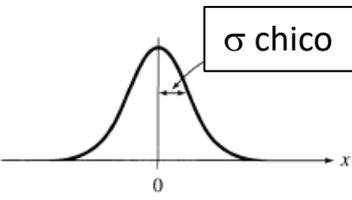
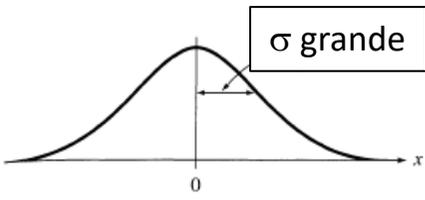
$$\bar{x} = X$$



$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x - \bar{x})^2 G_{X,\sigma}(x) dx$$



$$\sigma_x^2 = \sigma^2$$

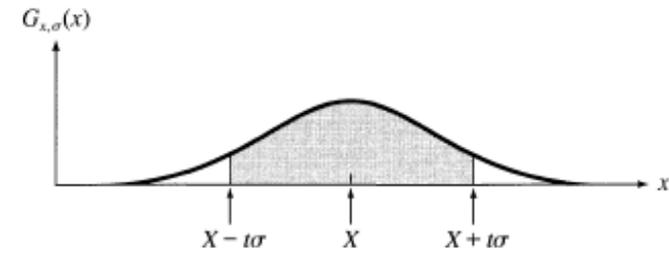


$\bar{x} = X, y = x - X, y z = y/\sigma$

Distribución normal o de Gauss: σ como límite de confianza del 68%

Si nos preguntamos cuál es la probabilidad de encontrar una medición dada en el entorno $\pm t\sigma$ de X tenemos que

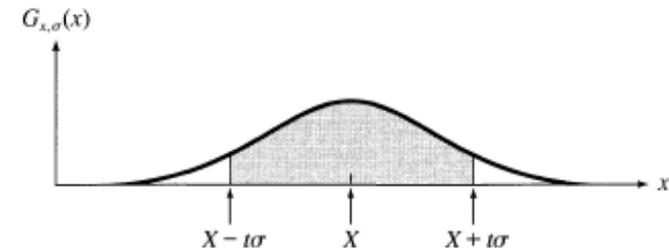
$$Prob(X \pm t\sigma) = \int_{X-t\sigma}^{X+t\sigma} G_{X,\sigma}(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{X-t\sigma}^{X+t\sigma} e^{-(x-X)^2/2\sigma^2} dx$$



Distribución normal o de Gauss: σ como límite de confianza del 68%

Si nos preguntamos cuál es la probabilidad de encontrar una medición dada en el entorno $\pm t\sigma$ de X tenemos que

$$Prob(X \pm t\sigma) = \int_{X-t\sigma}^{X+t\sigma} G_{X,\sigma}(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{X-t\sigma}^{X+t\sigma} e^{-(x-X)^2/2\sigma^2} dx$$



Sustituyendo $z = (x - X)/\sigma$, de donde $dx = \sigma dz$ y los límites de integración $z = \pm t$:

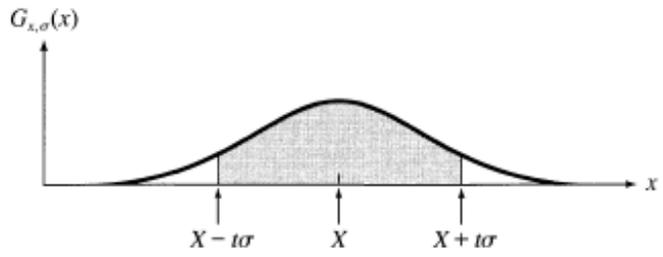
$$Prob(X \pm t\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-z^2/2} dz$$

: $erf(t)$ “error function”

Distribución normal o de Gauss: σ como límite de confianza del 68%

Si nos preguntamos cuál es la probabilidad de encontrar una medición dada en el entorno $\pm t\sigma$ de X tenemos que

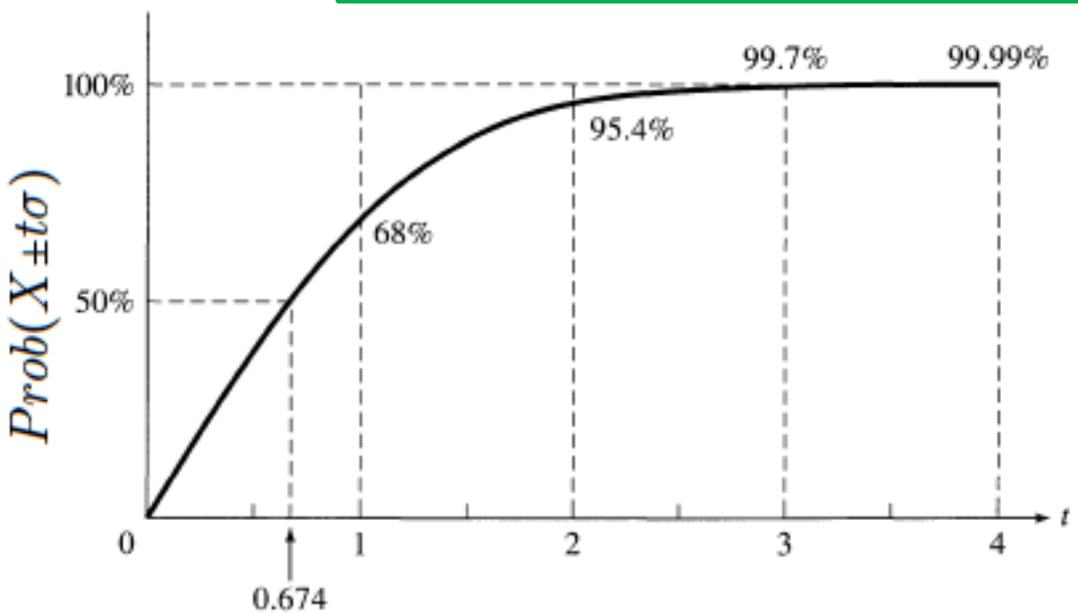
$$Prob(X \pm t\sigma) = \int_{X-t\sigma}^{X+t\sigma} G_{X,\sigma}(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{X-t\sigma}^{X+t\sigma} e^{-(x-X)^2/2\sigma^2} dx$$



Sustituyendo $z = (x - X)/\sigma$, de donde $dx = \sigma dz$ y los límites de integración $z = \pm t$:

$$Prob(X \pm t\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-z^2/2} dz$$

: $erf(t)$ “error function”



$$5\sigma \sim 0.00003\%$$

Distribución normal o de Gauss: justificación de \bar{x} como mejor estimador de X

Si $f(x)$ fuera conocida, podríamos obtener el valor “verdadero” X , y su desviación standard σ .

Sin embargo, **nunca conocemos a la distribución límite**. En la práctica podemos obtener un número limitado de valores medidos x_1, x_2, \dots, x_N : el problema es obtener al mejor estimador de X y σ basado en estos N valores medidos.

Distribución normal o de Gauss: justificación de \bar{x} como mejor estimador de X

Si $f(x)$ fuera conocida, podríamos obtener el valor “verdadero” X , y su desviación standard σ .

Sin embargo, **nunca conocemos a la distribución límite**. En la práctica podemos obtener un número limitado de valores medidos x_1, x_2, \dots, x_N : el problema es obtener al mejor estimador de X y σ basado en estos N valores medidos.

Si las mediciones siguen una distribución gaussiana $G_{X,\sigma}(x)$, entonces:

$$Prob(x \text{ entre } x_1 \text{ y } x_1 + dx_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_1-X)^2/2\sigma^2} dx_1$$

Si por otro lado las mediciones son independientes una de la otra,

$$\begin{aligned} Prob_{X,\sigma}(x_1, \dots, x_N) &= Prob(x_1) \times \dots \times Prob(x_N) \\ &\propto \frac{1}{\sigma^N} e^{-\sum (x_i-X)^2/2\sigma^2} \end{aligned}$$

Distribución normal o de Gauss: justificación de \bar{x} como mejor estimador de X

Si $f(x)$ fuera conocida, podríamos obtener el valor “verdadero” X , y su desviación standard σ .

Sin embargo, **nunca conocemos a la distribución límite**. En la práctica podemos obtener un número limitado de valores medidos x_1, x_2, \dots, x_N : el problema es obtener al mejor estimador de X y σ basado en estos N valores medidos.

Si las mediciones siguen una distribución gaussiana $G_{X,\sigma}(x)$, entonces:

$$Prob(x \text{ entre } x_1 \text{ y } x_1 + dx_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_1-X)^2/2\sigma^2} dx_1$$

Si por otro lado las mediciones son independientes una de la otra,

$$\begin{aligned} Prob_{X,\sigma}(x_1, \dots, x_N) &= Prob(x_1) \times \dots \times Prob(x_N) \\ &\propto \frac{1}{\sigma^N} e^{-\sum (x_i-X)^2/2\sigma^2} \end{aligned}$$

Encontrar los mejores “estimadores” de X y σ equivale a preguntarse cuáles son aquellos valores que *maximizan* la probabilidad de obtener las mediciones x_1, x_2, \dots, x_N (“principio de máxima verosimilitud”).

Distribución normal o de Gauss: justificación de \bar{x} como mejor estimador de X

Maximizar a $\frac{1}{\sigma^N} e^{-\sum (x_i - X)^2 / 2\sigma^2}$ es equivalente a *minimizar* $\sum_{i=1}^N (x_i - X)^2 / \sigma^2$

Diferenciando respecto a X , e igualando a cero, resulta

$$\sum_{i=1}^N (x_i - X) = 0$$

De donde,

$$\text{(mejor estimador de } X) = \frac{\sum x_i}{N}$$

El mejor estimador de X es el promedio de las N mediciones $\bar{x} = \sum x / N$.

Distribución normal o de Gauss: justificación de σ_x como mejor estimador de σ

Encontrar el mejor estimador de σ sigue el mismo procedimiento. Es decir, hay que buscar el valor de σ que maximiza la probabilidad de obtener el conjunto de mediciones x_1, x_2, \dots, x_N . Esto resulta en (hacer el ejercicio):

$$(\text{mejor estimador de } \sigma) = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - X)^2}$$

El mejor estimador de σ es la desviación estándar de las N mediciones.

Distribución normal o de Gauss: justificación de σ_x como mejor estimador de σ

Encontrar el mejor estimador de σ sigue el mismo procedimiento. Es decir, hay que buscar el valor de σ que maximiza la probabilidad de obtener el conjunto de mediciones x_1, x_2, \dots, x_N . Esto resulta en (hacer el ejercicio):

$$(\text{mejor estimador de } \sigma) = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - X)^2}$$

El mejor estimador de σ es la desviación estándar de las N mediciones.

Pero ojo que **no conocemos** a X , sino a su mejor estimador \bar{x} . Como vimos que \bar{x} es el valor que *minimiza* a la suma involucrada en este estimador de σ , es de esperar que al reemplazar a X por \bar{x} estemos *subestimando* a σ . Se puede demostrar que en este caso es:

$$(\text{mejor estimador de } \sigma) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Distribución normal o de Gauss: justificación de σ_x como mejor estimador de σ

Encontrar el mejor estimador de σ sigue el mismo procedimiento. Es decir, hay que buscar el valor de σ que maximiza la probabilidad de obtener el conjunto de mediciones x_1, x_2, \dots, x_N . Esto resulta en (hacer el ejercicio):

$$(\text{mejor estimador de } \sigma) = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - X)^2}$$

El mejor estimador de σ es la desviación estándar de las N mediciones.

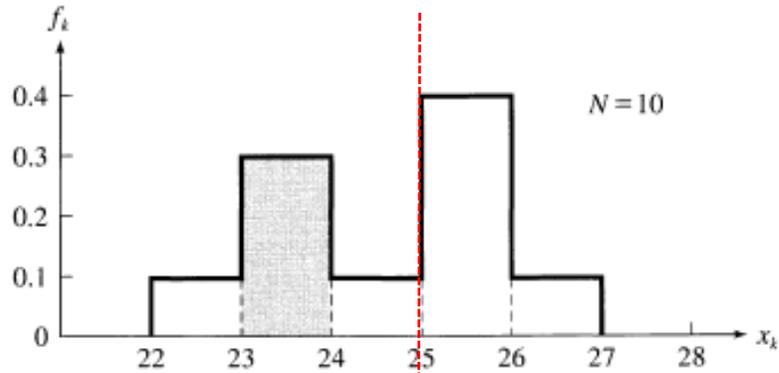
Pero ojo que **no conocemos** a X , sino a su mejor estimador \bar{x} . Como vimos que \bar{x} es el valor que *minimiza* a la suma involucrada en este estimador de σ , es de esperar que al reemplazar a X por \bar{x} estemos *subestimando* a σ . Se puede demostrar que en este caso es:

$$(\text{mejor estimador de } \sigma) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

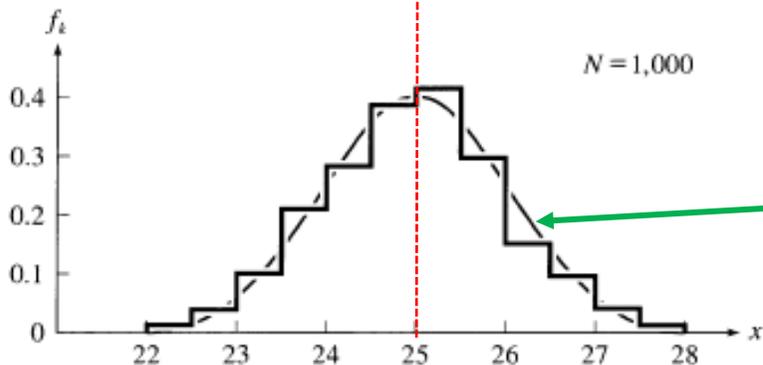
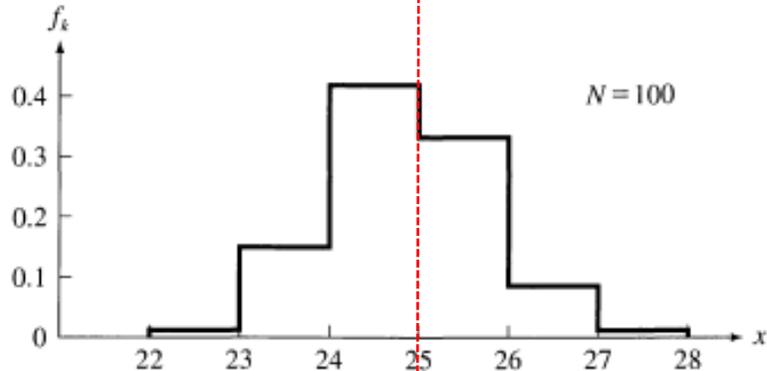
Finalmente, también se puede demostrar que

$$\frac{\delta \sigma_x}{\sigma_x} = \frac{1}{\sqrt{2(N-1)}}$$

Cómo usamos entonces a esta información?

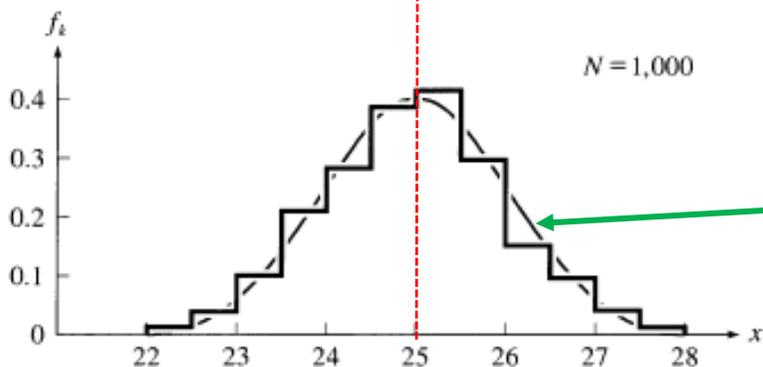
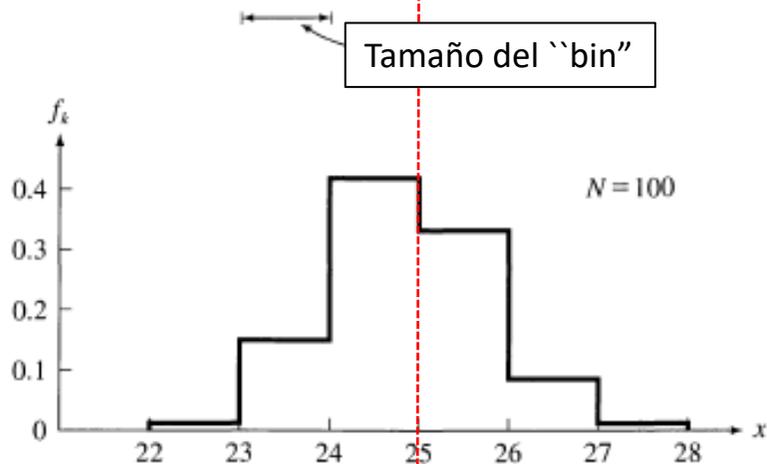
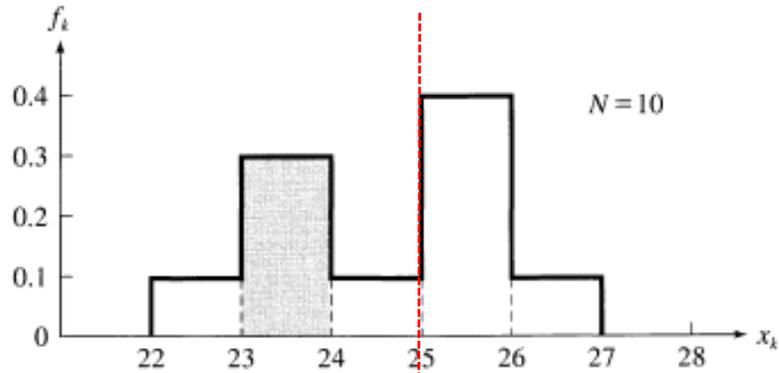


Tamaño del "bin"



Tiende a una curva continua para $N \rightarrow$ infinito:
la distribución "normal" o de Gauss (*teorema central del límite*)

Cómo usamos entonces a esta información?



La distribución caracteriza al método

σ_x es aproximadamente independiente de N

σ_x es la incerteza para una medición individual

Tiende a una curva continua para $N \rightarrow$ infinito: la distribución "normal" o de Gauss (*teorema central del límite*)

Propagación de errores: justificación de la suma en cuadratura

Si consideramos una función general $q(x,y)$ donde x e y se han determinado con incertezas σ_x y σ_y . La pregunta es cómo determinar la incerteza de q .

Cantidad más una constante: $q = x + A$

$$(\text{Probabilidad de obtener } x) \propto e^{-(x-X)^2/2\sigma_x^2}$$

Como $x = q - A$

$$\begin{aligned}(\text{Probabilidad de obtener } q) &\propto e^{-[(q-A)-X]^2/2\sigma_x^2} \\ &= e^{-[q-(A+X)]^2/2\sigma_x^2}\end{aligned}$$

Es decir, los valores de q seguirán una distribución normal, centrada en el valor desplazado $X + A$, y con una incerteza igual a la de x (σ_x).

Propagación de errores: justificación de la suma en cuadratura

Si consideramos una función general $q(x,y)$ donde x e y se han determinado con incertezas σ_x y σ_y . La pregunta es cómo determinar la incerteza de q .

Cantidad **más** una constante: $q = x + A$

$$(\text{Probabilidad de obtener } x) \propto e^{-(x-X)^2/2\sigma_x^2}$$

Como $x = q - A$

$$\begin{aligned}(\text{Probabilidad de obtener } q) &\propto e^{-[(q-A)-X]^2/2\sigma_x^2} \\ &= e^{-[q-(A+X)]^2/2\sigma_x^2}\end{aligned}$$

Es decir, los valores de q seguirán una distribución normal, centrada en el valor desplazado $X + A$, y **con una incerteza igual a la de x** (σ_x).

Cantidad **por** una constante: $q = Bx$, es decir $x=q/B$

$$\begin{aligned}(\text{Probabilidad de obtener } q) &\propto e^{-[(\frac{q}{B}-X)]^2/2\sigma_x^2} \\ &= e^{-[q-(BX)]^2/2B^2\sigma_x^2}\end{aligned}$$

Los valores de q seguirán una distribución normal, centrada en BX , y **con una incerteza $B\sigma_x$** .

Propagación de errores: justificación de la suma en cuadratura

Suma de dos cantidades: $q = x + y$

Para dos valores arbitrarios de x e y : $Prob(x, y) \propto \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right]$

Propagación de errores: justificación de la suma en cuadratura

Suma de dos cantidades: $q = x + y$

Para dos valores arbitrarios de x e y : $Prob(x, y) \propto \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right]$

Teniendo en cuenta que $\frac{x^2}{A} + \frac{y^2}{B} = \frac{(x+y)^2}{A+B} + \frac{(Bx - Ay)^2}{AB(A+B)} = \frac{(x+y)^2}{A+B} + z^2$

Propagación de errores: justificación de la suma en cuadratura

Suma de dos cantidades: $q = x + y$

Para dos valores arbitrarios de x e y : $Prob(x, y) \propto \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right]$

Teniendo en cuenta que $\frac{x^2}{A} + \frac{y^2}{B} = \frac{(x+y)^2}{A+B} + \frac{(Bx - Ay)^2}{AB(A+B)} = \frac{(x+y)^2}{A+B} + z^2$

Entonces $Prob(x + y, z) \propto \exp\left[-\frac{(x+y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right] \exp\left[-\frac{z^2}{2}\right]$

Integrando para todo z , $Prob(x + y) \propto \exp\left[-\frac{(x+y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right]$

Los valores de $x + y$ siguen una distribución normal, con ancho $\sqrt{\sigma_x^2 + \sigma_y^2}$.

Propagación de errores: justificación de la suma en cuadratura

Por una constante, multiplica σ

Caso más general: $q(x, y) \approx q(X, Y) + \left(\frac{\partial q}{\partial x}\right)(x - X) + \left(\frac{\partial q}{\partial y}\right)(y - Y)$

Más una constante, no modifica σ

Entonces:

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x} \sigma_x\right)^2 + \left(\frac{\partial q}{\partial y} \sigma_y\right)^2}$$

Regla general para funciones de varias variables

Entonces en general, si los errores son aleatorios y las variables son independientes:

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x} \sigma_x\right)^2 + \dots + \left(\frac{\partial q}{\partial z} \sigma_z\right)^2}$$

Y, en cualquier caso, es siempre menor a la suma ordinaria

$$\delta q \leq \left|\frac{\partial q}{\partial x}\right| \sigma_x + \dots + \left|\frac{\partial q}{\partial z}\right| \sigma_z$$

Propagación de errores: casos usuales derivados de la fórmula general

Sumas y restas

Supongamos que x, \dots, w son cantidades medidas con incertezas $\delta x, \dots, \delta w$, y estos valores medidos se utilizan para computar la cantidad

$$q = x + \dots + z - (u + \dots + w)$$

Entonces, si las incertezas en x, \dots, w se sabe que son independientes y aleatorias,

$$\sigma_q = \sqrt{(\sigma_x)^2 + \dots + (\sigma_z)^2 + (\sigma_u)^2 + \dots + (\sigma_w)^2}$$

Y, en cualquier caso, σ_q nunca es mayor a la suma directa de todas las incertezas,

$$\sigma_q \leq \sigma_x + \dots + \sigma_z + \sigma_u + \dots + \sigma_w$$

Propagación de errores: casos usuales derivados de la fórmula general

Productos y cocientes

Supongamos ahora que x, \dots, w son cantidades medidas con incertezas $\delta x, \dots, \delta w$, y estos valores medidos se utilizan para computar la cantidad

$$q = \frac{x \times \dots \times z}{u \times \dots \times w}$$

Entonces, si las incertezas en x, \dots, w se sabe que son independientes y aleatorias,

$$\frac{\sigma_q}{|q|} = \sqrt{\left(\frac{\sigma_x}{|x|}\right)^2 + \dots + \left(\frac{\sigma_z}{|z|}\right)^2 + \left(\frac{\sigma_u}{|u|}\right)^2 + \dots + \left(\frac{\sigma_w}{|w|}\right)^2}$$

Y, en cualquier caso, σ_q nunca es mayor a la suma directa de todas las incertezas “relativas”,

$$\frac{\sigma_q}{|q|} \leq \frac{\sigma_x}{|x|} + \dots + \frac{\sigma_z}{|z|} + \frac{\sigma_u}{|u|} + \dots + \frac{\sigma_w}{|w|}$$

Propagación de errores: casos usuales derivados de la fórmula general

Potencias

Si $q(x) = x^n$, entonces

$$\frac{\sigma_q}{|q|} = |n| \frac{\sigma_x}{|x|}$$

Producto por una constante

Si $q = Bx$, entonces

$$\sigma_q = |B| \sigma_x$$

Propagación de errores: incerteza del valor medio

La desviación standard se puede entender como la incerteza asociada *a una medición única*.

Cuando se mide muchas veces, y es posible de esa manera determinar a \bar{x} y a σ_x , entonces se reporta a \bar{x} como el mejor estimador para la cantidad buscada (suponiendo que no hay incertezas sistemáticas).

El valor medio es la suma de las variables independientes x_i , todas con la misma desviación standard σ_x . Utilizando la regla general para propagación de errores:

$$\sigma_{\bar{x}} = \sqrt{\sum_{i=1}^N \left(\frac{\partial \bar{x}}{\partial x_i} \sigma_x \right)^2} = \sqrt{\sum_{i=1}^N \left(\frac{1}{N} \sigma_x \right)^2} = \sqrt{\frac{1}{N^2} \sum_{i=1}^N \sigma_x^2}$$

Propagación de errores: incerteza del valor medio

La desviación standard se puede entender como la incerteza asociada *a una medición única*.

Cuando se mide muchas veces, y es posible de esa manera determinar a \bar{x} y a σ_x , entonces se reporta a \bar{x} como el mejor estimador para la cantidad buscada (suponiendo que no hay incertezas sistemáticas).

El valor medio es la suma de las variables independientes x_i , todas con la misma desviación standard σ_x . Utilizando la regla general para propagación de errores:

$$\sigma_{\bar{x}} = \sqrt{\sum_{i=1}^N \left(\frac{\partial \bar{x}}{\partial x_i} \sigma_x \right)^2} = \sqrt{\sum_{i=1}^N \left(\frac{1}{N} \sigma_x \right)^2} = \sqrt{\frac{1}{N^2} \sum_{i=1}^N \sigma_x^2}$$

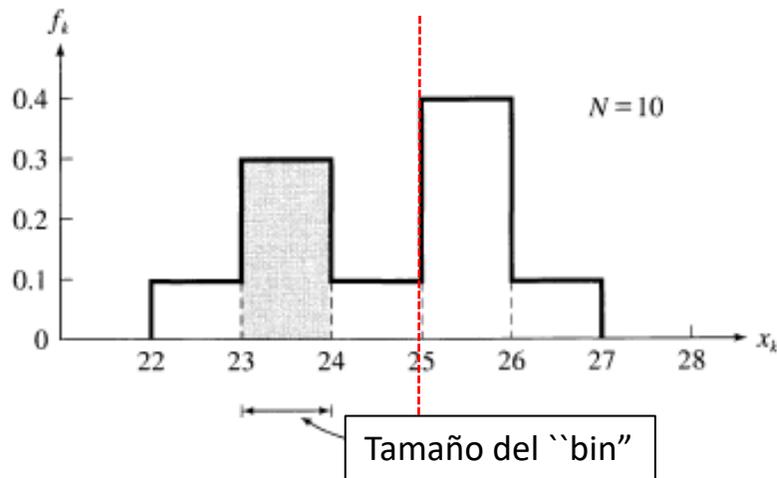
De donde

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}$$

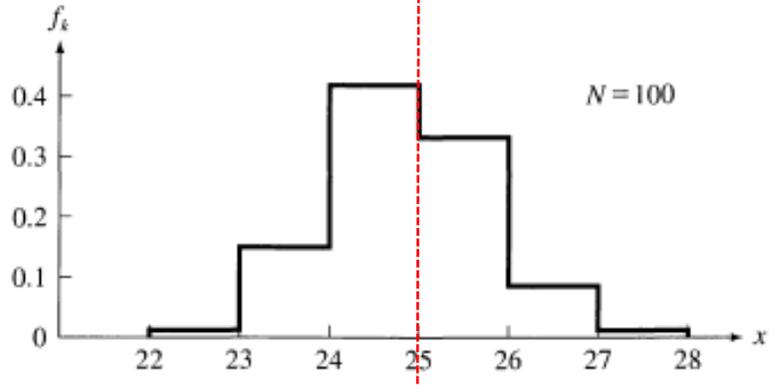
Reportaremos el resultado de un experimento consistente en N determinaciones de la misma cantidad física x_1, \dots, x_N de la siguiente manera,

$$x = \bar{x} \pm \sigma_{\bar{x}}$$

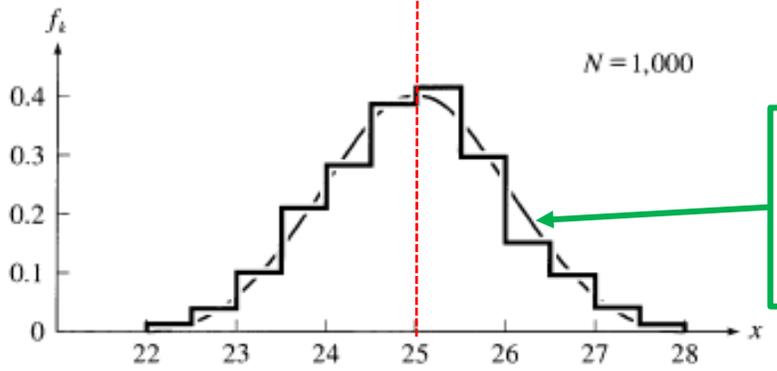
Para recordar el efecto de hacer una muestra más grande:



σ_x es aproximadamente independiente de N



$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}} : \text{decrece con N al definirse mejor la distribución}$$



Tiende a una curva continua para $N \rightarrow$ infinito: la distribución "normal" o de Gauss (*teorema central del límite*)

Cómo incluir los errores sistemáticos en la incerteza total

Finalmente, y aunque no tiene una demostración formal, si existieran incertezas sistemáticas cuya magnitud pueda determinarse, se puede utilizar argumentos de razonabilidad para concluir que

$$\delta x_{tot} = \sqrt{(\delta x_{ale})^2 + (\delta x_{sis})^2}$$

Y ante la duda,

$$\delta x_{tot} \leq |\delta x_{ale}| + |\delta x_{sis}|$$

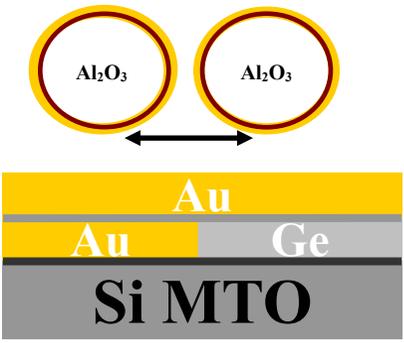
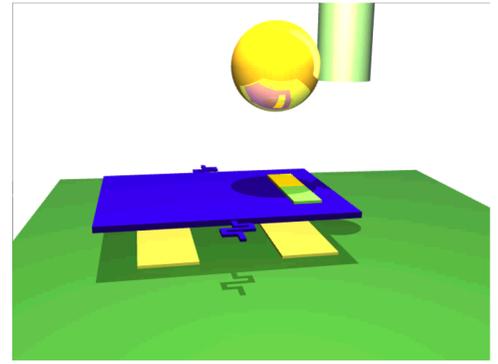
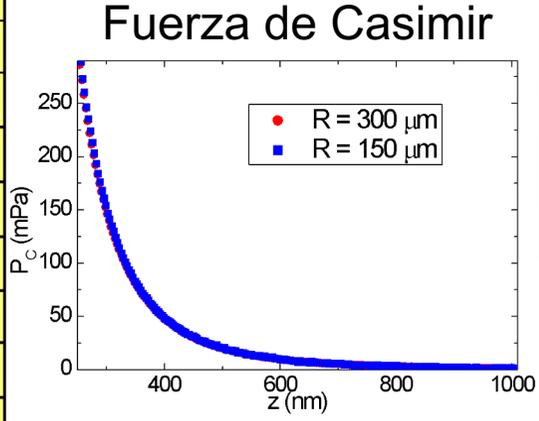
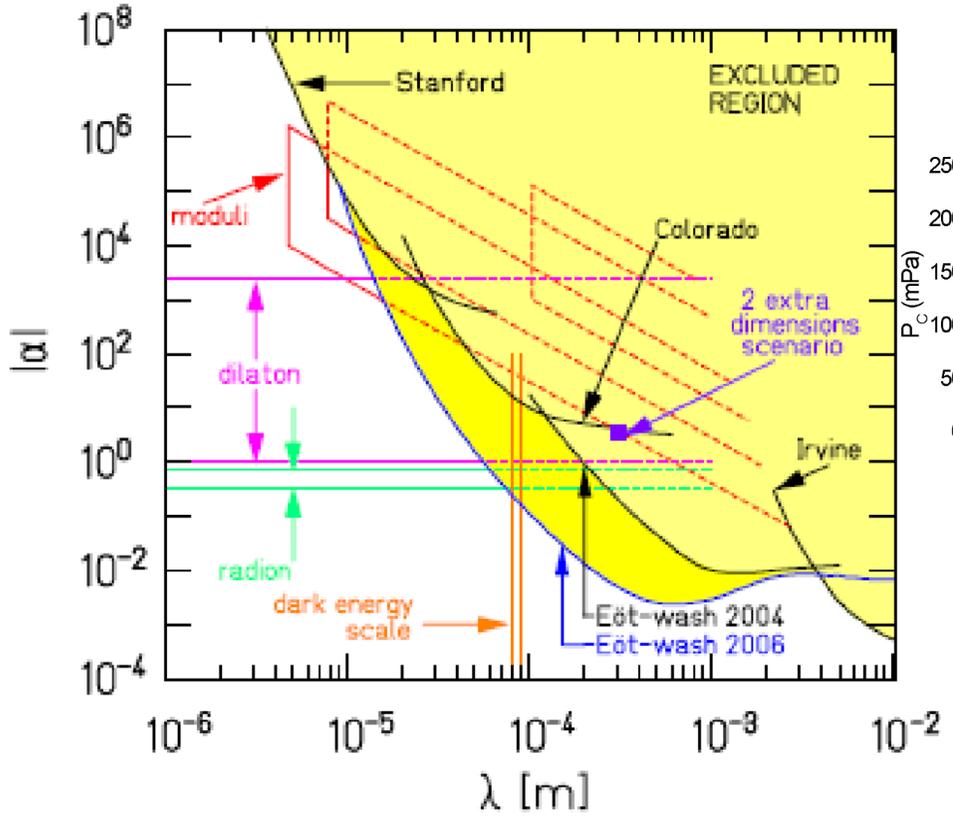
Atracción gravitatoria: test de modelos

PHYSICAL REVIEW D 75, 077101 (2007)

Tests of new physics from precise measurements of the Casimir pressure between two gold-coated plates

R. S. Decca,¹ D. López,² E. Fischbach,³ G. L. Klimchitskaya,⁴ D. E. Krause,^{5,3} and V. M. Mostepanenko⁶

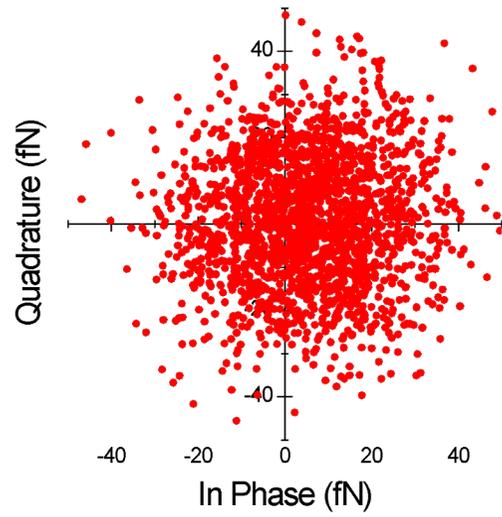
$$V(r) = -G \frac{m_1 m_2}{r} \left(1 + \alpha \cdot e^{-r/\lambda} \right) \quad \text{Potencial tipo "Yukawa"}$$



Atracción gravitatoria: test de modelos

$$z = 500 \text{ nm}$$

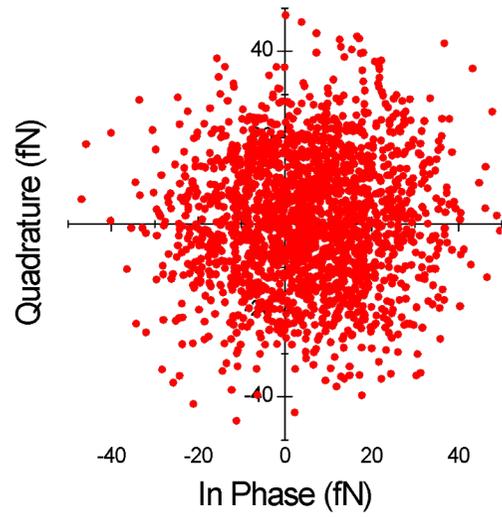
1 seg



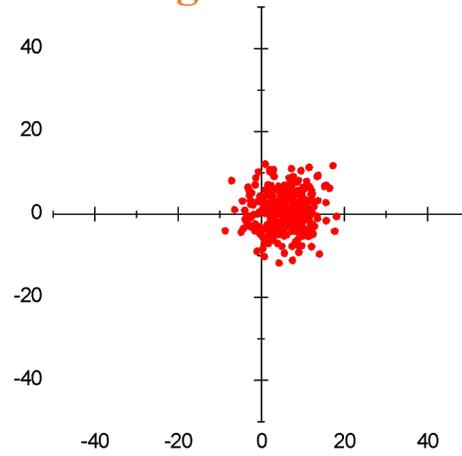
Atracción gravitatoria: test de modelos

$z = 500 \text{ nm}$

1 seg

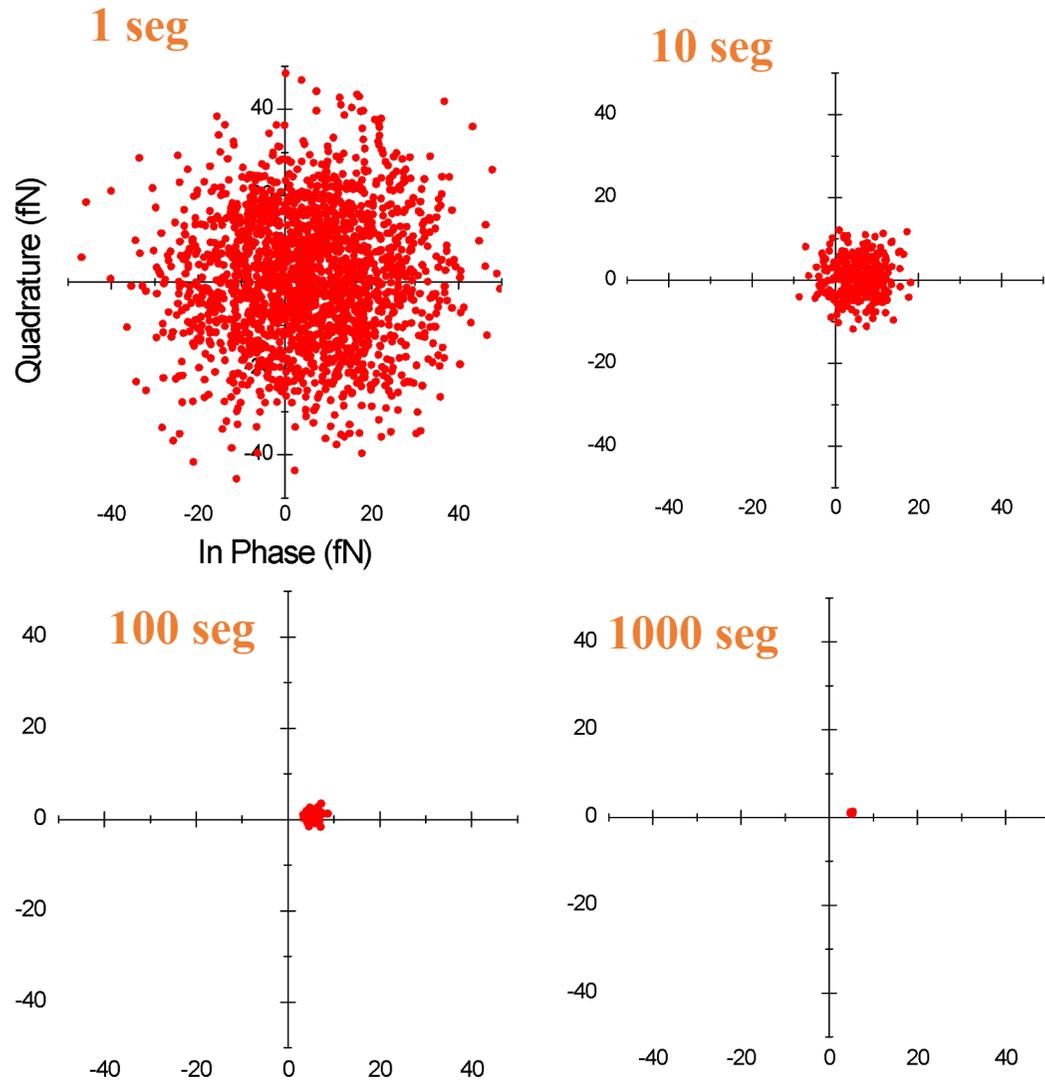


10 seg



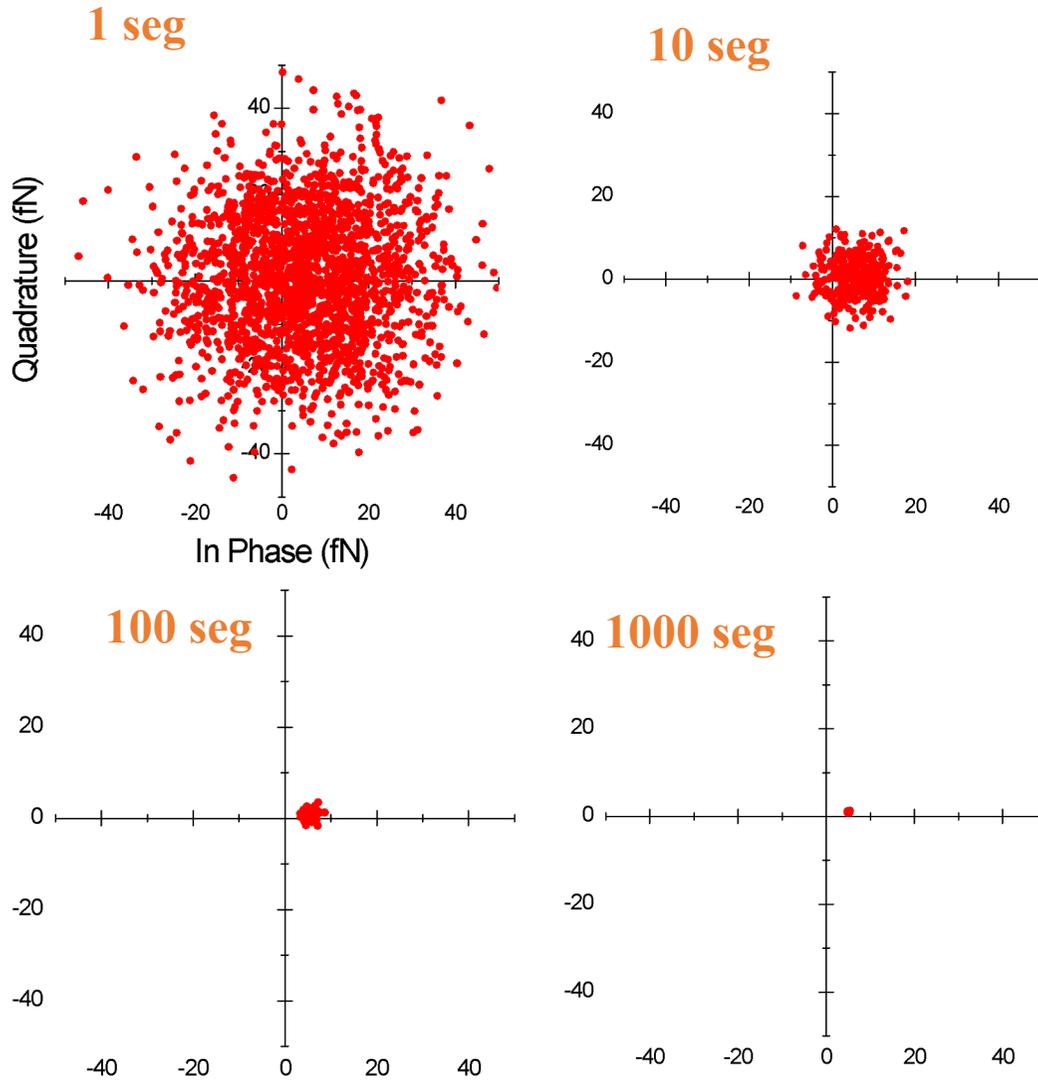
Atracción gravitatoria: test de modelos

$$z = 500 \text{ nm}$$

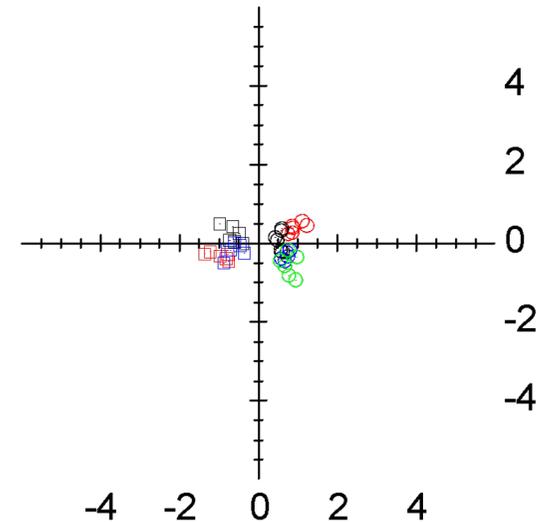


Atracción gravitatoria: test de modelos

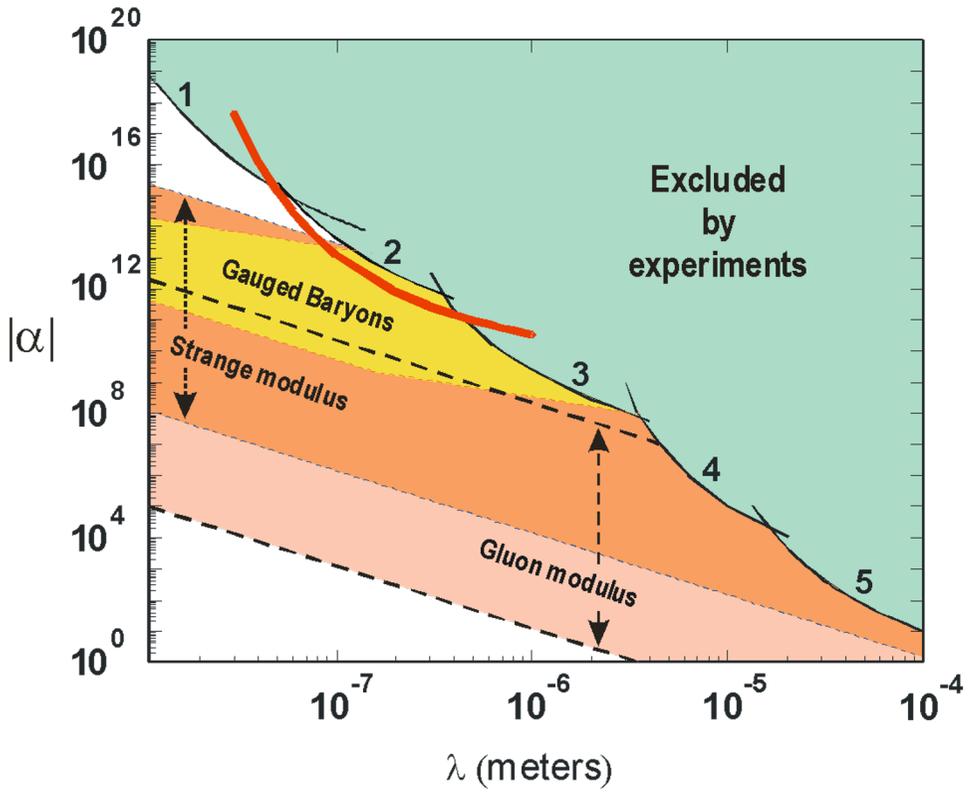
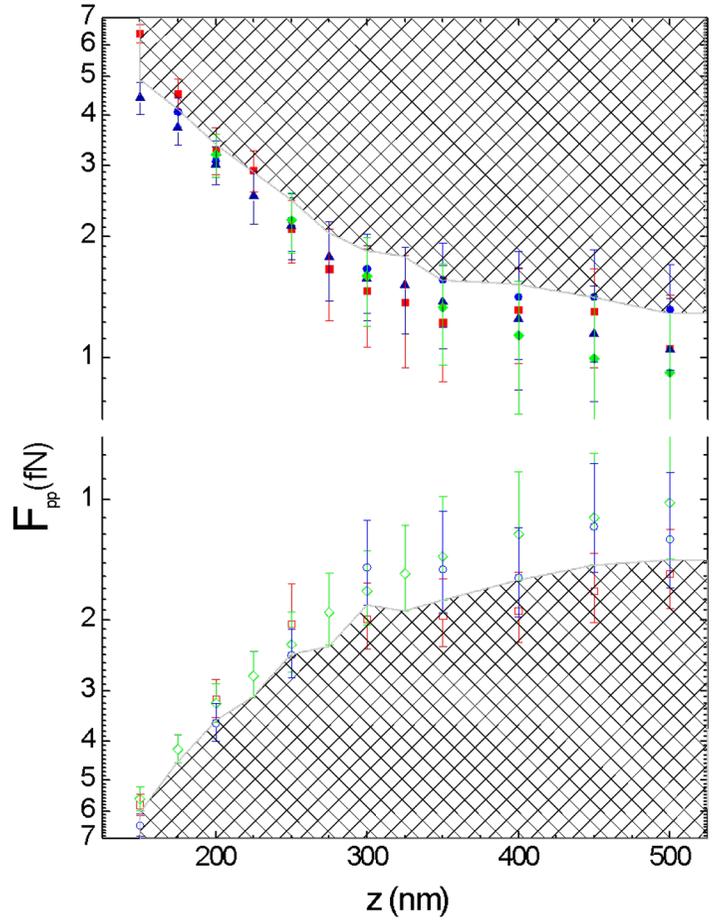
$$z = 500 \text{ nm}$$



**Test de consistencia:
más muestras!**



Atracción gravitatoria: test de modelos



**La incertidumbre en el experimento...
y en la teoría**

Una receta a veces útil: propagación de errores “paso a paso”

Supongamos por ejemplo tener que calcular la cantidad

$$q = x (y - z \sin u)$$

a partir de las cantidades medidas x , y , z y u , y obtener la incerteza δq a partir de las correspondientes incertezas de x , y , z y u . Esto se puede hacer con los siguientes pasos:

- i) Calcular la *función* $\sin u$ y su incerteza,
- ii) Luego evaluar el *producto* de z con $\sin u$, y obtener su incerteza
- iii) Seguidamente la *diferencia* entre y y $z \sin u$, con su incerteza, y
- iv) Finalmente terminar con el *producto* de x por $(y - z \sin u)$

Una receta a veces útil: propagación de errores “paso a paso”

Supongamos por ejemplo tener que calcular la cantidad

$$q = x (y - z \sin u)$$

a partir de las cantidades medidas x , y , z y u , y obtener la incerteza δq a partir de las correspondientes incertezas de x , y , z y u . Esto se puede hacer con los siguientes pasos:

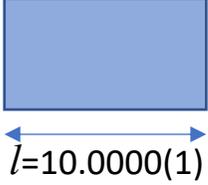
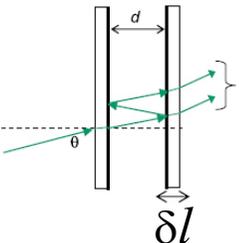
- i) Calcular la *función* $\sin u$ y su incerteza,
- ii) Luego evaluar el *producto* de z con $\sin u$, y obtener su incerteza
- iii) Seguidamente la *diferencia* entre y y $z \sin u$, con su incerteza, y
- iv) Finalmente terminar con el *producto* de x por $(y - z \sin u)$

Pero ojo que esto no se puede hacer si los distintos “pasos” no corresponden a incertezas independientes, por ejemplo en funciones como:

$$q = y - x \sin y \quad \text{ó} \quad q = \frac{y - x \sin z}{y + x \sin z}$$

En este caso usar la fórmula general

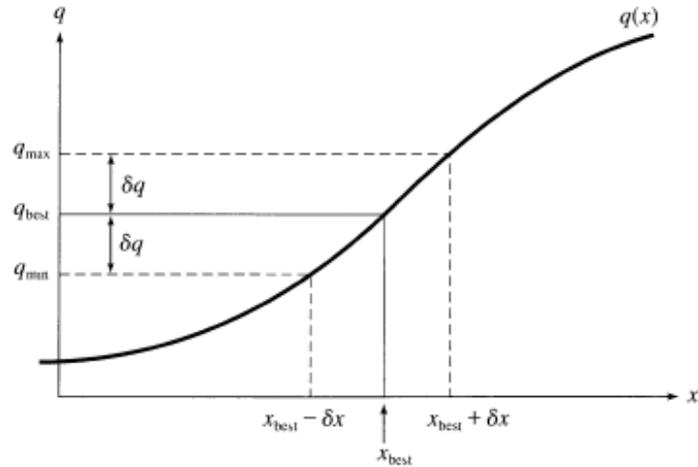
El "quiz" de la incerteza: aproximadamente cuál es?

1		Centímetro costurero Medir una barra de $\sim 12.345(1)$ cm	1cm	1mm	0.1mm
2	 $l = 10.0000(1)$ cm	Medir una mesa de $1.2345(1)$ m	10cm	5cm	Menos pero no sé cuánto
3		Microscopio óptico $\lambda = 500$ nm Ancho de una raya	1cm	500nm	Mucho menos pero no sé cuánto
4		Microscopio electrónico 10KeV - $\lambda_B = 0.01$ nm Ancho de una raya	$1\mu\text{m}$	1nm	0.01nm
5		Interferómetro óptico $\lambda = 500$ nm Determinar δl	1cm	500nm	Mucho menos pero no sé cuánto

Propagación de errores: casos usuales derivados de la fórmula general

Función arbitraria de una variable

En el caso más general, si q es una función arbitraria de una variable, $q(x)$, entonces



$$\sigma_q = \left| \frac{dq}{dx} \right| \sigma_x$$

El porqué la varianza va dividida por (N-1)

Vimos que el mejor estimador de la desviación standard era $\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - X)^2}$

donde X es el valor "real" que es desconocido. Podemos reescribir esto como

$$\begin{aligned} N\sigma_x^2 &= \sum_{i=1}^N (x_i + \bar{x} - \bar{x} - X)^2 \\ &= \sum_{i=1}^N (x_i - \bar{x})^2 + \sum_{i=1}^N (\bar{x} - X)^2 + 2(\bar{x} - X) \sum_{i=1}^N (x_i - \bar{x}) \end{aligned}$$

= 0

El segundo término es N veces la varianza del valor medio $N\sigma_x^2/N$. Y el último término es cero. Tenemos entonces que

$$N\sigma_x^2 = \sum_{i=1}^N (x_i - \bar{x})^2 + \sigma_x^2$$

de donde

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$