Introducción a la estadística de las mediciones

Dpto. de Física, FCEyN, UBA
Mónica B. Agüero

Los **errores casuales** o **aleatorios** generan distinguibilidad entre dato y dato (obtenemos resultados distintos), cuando repetimos un experimento en idénticas condiciones. Sin embargo, mediante la teoría estadística podemos determinar el valor de la magnitud medida y estimar su error estadístico.

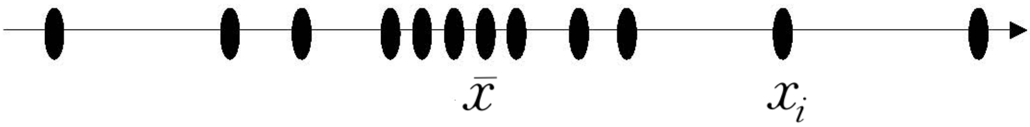
Supongamos que realizamos una serie de N mediciones de una misma magnitud obteniendo los resultados $x_1, ..., x_N$. Nos podemos preguntar:

¿Hay alguna regularidad en los resultados?

¿Alguno de ellos aparece con más frecuencia que los demás?

¿Qué resultado representa mejor al grupo de observaciones en su totalidad?

Si consideremos el caso particular de una serie de resultados distribuidos alrededor del promedio $\bar{x}$, observamos que hay valores que están más cerca del promedio y otros, menores en número, se encuentran más lejos. Si hacemos una nueva medición x_{N+1} no sabemos de antemano



el resultado que va a salir pero sí podremos decir que con cierta probabilidad estará cerca del promedio. No podemos predecir el valor de una medición dada pero sí podemos decir algo sobre la probabilidad de que su valor caiga en un determinado intervalo de valores posibles.

Una herramienta útil que vamos a usar en este curso para analizar este conjunto de datos son los histogramas.

1. Histograma

Si dividimos el eje x en pequeños intervalos iguales Δx, podemos contar el número de observaciones ΔN que caen en cada intervalo y representarlo gráficamente. Este gráfico de barras verticales se llama histograma (ver Fig. 1). Los histogramas son útiles para describir distribuciones con un gran número de datos $(x_1, x_2, ..., x_N)$ obtenidos experimentalmente. Para realizar un histograma:

- marcamos intervalos regulares sobre un eje horizontal (en el rango donde están los valores de las mediciones).
- sobre cada intervalo dibujamos un rectángulo cuya altura es proporcional a la cantidad de mediciones que caen dentro de dicho intervalo (*número de ocurrencia*).
- Cuanto más grande sea la estadística, más pequeños podemos hacer los intervalos Δx y tener un número suficientemente grande de datos ΔN en cada intervalo.

- En vez del *número de ocurrencia* podemos graficar, en el eje de las ordenadas, la *frecuencia de ocurrencia* que se obtiene dividiendo el *número de observación* de cada intervalo por el *número total de observaciones*: $\frac{\Delta N}{N}$.

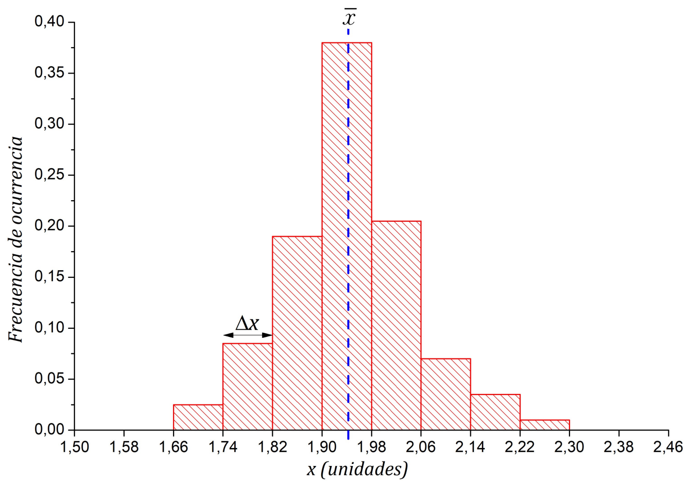


Figura 1: Modelo de histograma. El ancho Δx de las barras del histograma se lo conoce como *factor de clase* o *bin size*.

Ahora podemos apreciar de un solo vistazo cómo se distribuyen los valores a lo largo de la escala. Esta distribución es clave para una interpretación satisfactoria de las mediciones.

2. Tratamiento estadístico de datos

Pero queremos ir más lejos. Entonces, para analizar la serie de N mediciones de una misma magnitud obtenida en igualdad de condiciones vamos a emplear la Teoría Estadística. Para ello veamos algunas definiciones a tener en cuenta.

2.1. Medidas de Centralización

Hay diferentes posibilidades para definir el resultado que representa mejor al grupo de observaciones en un conjunto de datos [1]. Se definen (Fig. 2):

- **Mediana:** si ordeno las mediciones de menor a mayor, la mediana es aquel valor que separa por la mitad las observaciones de tal forma que el 50 % de estas son menores que la *mediana* y el otro 50 % son mayores.

- **Moda:** valor de la variable que más veces se repite, es decir, aquella cuya frecuencia absoluta es mayor. No tiene por qué ser única (ver Fig. 3). Si queremos destacar aquel valor que sale más veces, entonces mencionamos el *valor modal.*

- **Media o promedio:** Dada una serie de mediciones $x_1, x_2, ..., x_N$, la media es el promedio aritmético de todas las mediciones de esa cantidad efectuadas en las mismas condiciones:

$$\bar{x} = \frac{1}{N} \sum x_i \tag{1}$$

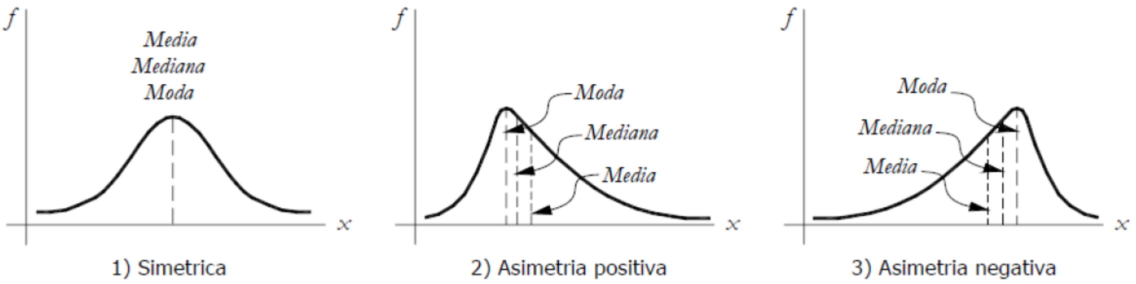


Figura 2: Distribuciones donde se ejemplifica la relación entre moda, mediana y media. Notar que para una distribución simétrica, la media, la mediana y la moda coinciden todas en el centro de la distribución.

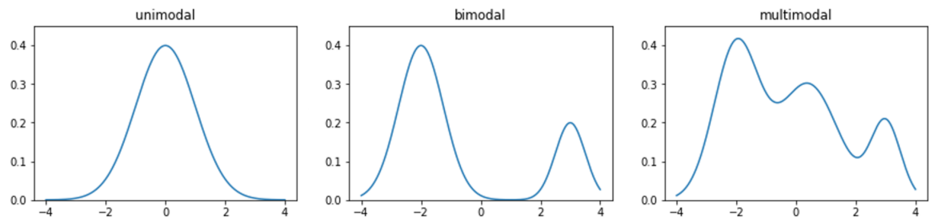


Figura 3: Ejemplos de distribuciones unimodal, bimodal y multimodal.

2.2. Amplitud de las distribuciones

Otra parámetro que se define es el **desvío estándar** σ [1,2]. Esta cantidad es una medida de la distribución de las mediciones alrededor del valor más probable. Nos da una idea de cuán dispersos están los datos alrededor del valor promedio.

$$\sigma = \sqrt{\frac{1}{N-1}\sum(\bar{x} - x_i)^2} \tag{2}$$

La Ec. (2) corresponde al desvío estándar de una muestra. El desvío estándar de una muestra estima la desviación estándar de una población basada en una muestra aleatoria (porque no se dispone de datos sobre la población completa). Sin embargo, cuando la muestra es grande, $N-1$ tiende a N. Por este motivo el desvío estándar de la muestra tiende al poblacional para muestras grandes. En la Sec. 2.3 se ejemplifica el rol de σ para una distribución Gaussiana.

2.3. Distribución de Gauss

Si la forma obtenida para el histograma es una barra central rodeada por barras decrecientes distribuidas más o menos simétricamente a su alrededor, se dice entonces que dicho histograma presenta una típica *distribución normal* o *Gaussiana.*

La experiencia muestra que, para todos los casos de errores casuales, el histograma correspondiente a la serie de N mediciones $(x_1, x_2, ..., x_N)$ puede ser aproximado por una función continua bien definida conocida como curva de **distribución de Gauss**. Los datos están distribuidos alrededor del promedio $\bar{x}$, donde algunos valores que estarán cerca del promedio y, otros menos en número, estarán lejos.

Sea ΔN el número de mediciones que caen en el intervalo (entre x y $x+\Delta x$). Se comprueba experimentalmente que ese número depende del valor de x y de la longitud del intervalo Δx en forma aproximada [3]

$$\Delta N \approx \frac{N}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(\bar{x}-x)^2}{2\sigma^2}\right] \Delta x \tag{3}$$

La relación se transforma en igualdad para diferenciales dN y dx. Y la expresión

$$\lim_{\Delta x \to 0} \frac{\Delta N}{\Delta x} = \frac{dN}{dx} = \frac{N}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(\bar{x}-x)^2}{2\sigma^2}\right] = G(x) \tag{4}$$

se llama *curva de distribución* o *curva de Gauss*.

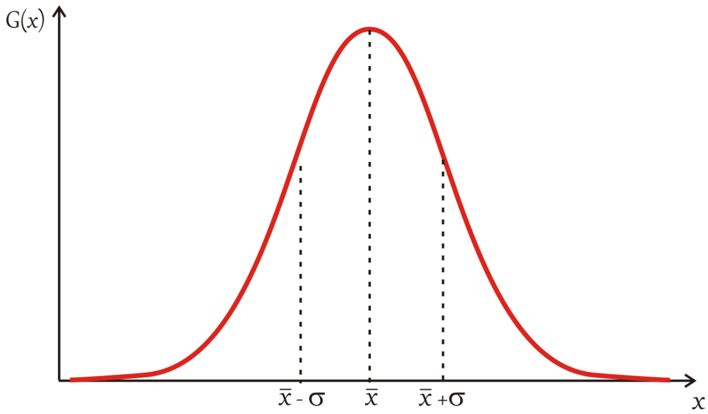


Figura 4: Curva de Gauss. En el eje de las abscisas se indican el promedio $\bar{x}$ y los puntos de inflexión de la curva ($\bar{x}-\sigma$ y $\bar{x}+\sigma$).

Podemos notar lo siguiente:

1. Presenta un máximo en $x = \bar{x}$.
2. Es simétrica respecto de ese valor medio $\bar{x}$.
3. Tiene forma de campana.
4. Sus puntos de inflexión están en $\bar{x} \pm \sigma$.
5. Tiende a cero a medida que nos alejamos de $\bar{x}$.

Cuanto **menor sea** σ, **más aguda** resultará la **curva resultante** y los errores mayores tendrán una menor probabilidad de ocurrencia, lo que significa que la **precisión del experimento es alta**. En cambio, si σ **es grande**, la **curva** será **achatada** (o dispersa) y esto indica una **baja precisión del experimento** (habrá un número considerable de mediciones con grandes desviaciones). En la Fig. 5 se muestran ejemplos de estos casos.

El área total bajo la curva de Gauss es

$$\int_{-\infty}^{\infty} \frac{N}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(\bar{x}-x)^2}{2\sigma^2}\right] dx = \frac{N}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{-(\bar{x}-x)^2}{2\sigma^2}\right] dx = \frac{N}{\sqrt{2\pi}}\sqrt{2\pi} = N \tag{5}$$

donde recordemos que N representa el número total de observaciones.

Si bien es imposible predecir el valor exacto que saldrá de una medición dada, sí se puede decir algo sobre la probabilidad de que ese valor este comprendido en un intervalo dado. El número de observaciones cuyo valor está comprendido entre x_1 y x_2 será

$$\frac{N}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} \exp\left[\frac{-(\bar{x}-x)^2}{2\sigma^2}\right] dx = \Delta N \tag{6}$$

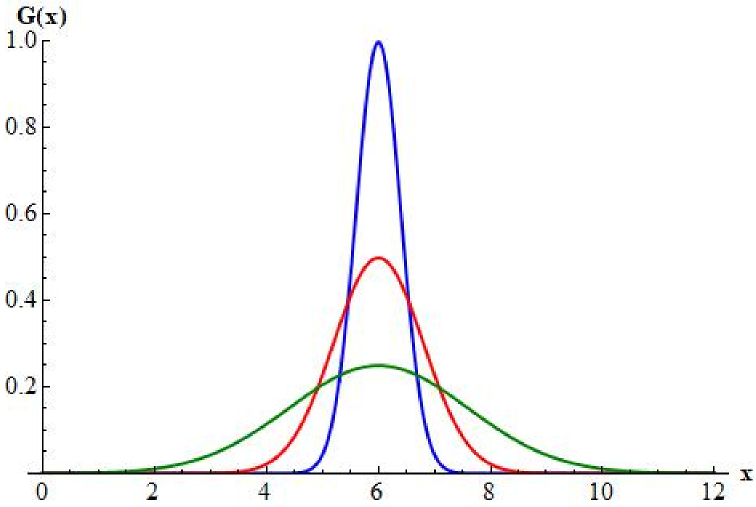


Figura 5: Efecto de σ sobre el ancho de la campana de Gauss.Curva azul: $\sigma = 0,4$. Curva roja: $\sigma = 0,8$. Curva verde: $\sigma = 1{,}6$. Notar que las tres curvas están centradas en $\bar{x} = 6$.

Si ahora dividimos ΔN por N (número total de datos)

$$\frac{\Delta N}{N} = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} \exp\left[\frac{-(\bar{x}-x)^2}{2\sigma^2}\right] dx, \tag{7}$$

obtenemos lo que se llama la **probabilidad de que una observación dada esté comprendida en este intervalo**. Efectivamente, la probabilidad es por definición el cociente entre el número de casos "favorables" (o sea, en este caso, los que están en ese intervalo $(x_1,\, x_2)$) y el número de casos totales (N). El número $100\frac{\Delta N}{N}$ representa la probabilidad expresada en porcentaje. La probabilidad de encontrar un dato entre $-\infty$ y $+\infty$ es 1 (o sea 100 %). La probabilidad de encontrar un dato fuera del intervalo $(x_1,\, x_2)$ será $1 - \frac{\Delta N}{N}$.

En resumen: de esto deducimos que, si bien es imposible predecir el valor exacto que saldrá en una medición dada, sí podemos decir algo sobre la probabilidad de que ese valor esté comprendido en un intervalo dado. Y para la predicción de esas probabilidades hemos utilizado la función de Gauss.

A partir de la Ec. (7) se obtiene que la probabilidad de que una medición dada caiga entre $\bar{x} - \sigma$ y $\bar{x} + \sigma$ es del 68 % y entre $\bar{x} - 2\sigma$ y $\bar{x} + 2\sigma$ es del 95.4 % (ver Fig. 6).

Supongamos haber hecho 100 mediciones de una magnitud, con un valor medio $\bar{x}$ y una dispersión estándar σ de cada dato [3]. Supongamos que entre los 100 datos haya tres que difieren del valor medio en más de 3σ, por ejemplo. De acuerdo con la función de Gauss, la probabilidad de que un dato caiga fuera del intervalo $\bar{x} \pm 3\sigma$ es

$$1 - \frac{1}{\sigma\sqrt{2\pi}} \int_{\bar{x}-3\sigma}^{\bar{x}+3\sigma} \exp\left[\frac{-(\bar{x}-x)^2}{2\sigma^2}\right] dx = 0,0028 \quad (0,28\,\%) \tag{8}$$

O sea que sólo entre 1000 datos podría esperarse que haya 3 fuera de ese intervalo. Esto indica que esas tres mediciones particulares padecen de un defecto "extra-gaussiano": deben rechazarse. De esta manera podemos fijar para cada serie de mediciones un "límite de confidencia". Cualquier dato cuyo valor caiga fuera del intervalo dado por el límite de confidencia

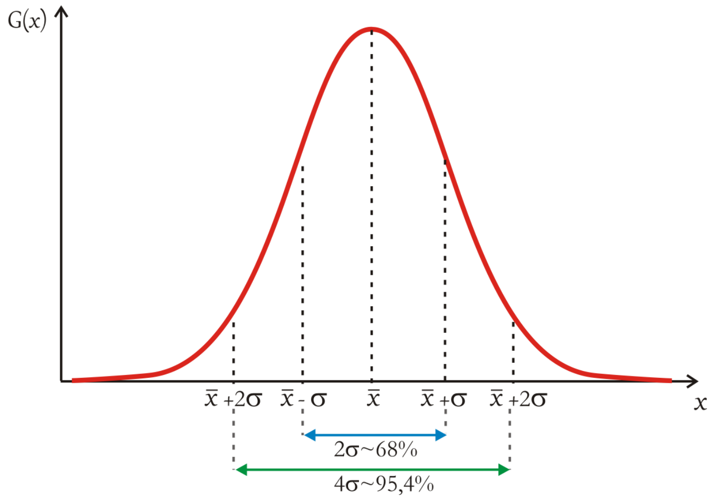


Figura 6: Distribución de probabilidades.

debe ser rechazado. Por supuesto que para determinar el límite de confidencia no puede darse un criterio unívoco; ese límite dependerá además del número total de mediciones N. En este curso no profundizaremos sobre ese tema.

2.4. Distribución de los promedios

Cuando tenemos varias (M) series de N mediciones, con sus promedios parciales $\bar{x}^k$, se comprueba experimentalmente que estos M promedios también se distribuyen "gaussianamente" alrededor del promedio total $\bar{X}$. O sea, su distribución es a su vez una curva de Gauss en la que la dispersión cuadrática del promedio ξ es el parámetro que fija sus puntos de inflexión. La interpretación de ξ, como uno de los parámetros de esa curva de Gauss, nos dice que el 68 % de los promedios parciales estarán entre $\bar{X} - \xi$ y $\bar{X} + \xi$.

Cuando hacemos una sola serie de N mediciones, obteniendo un promedio $\bar{x}$, sabemos a priori que ese promedio, y todos los promedios de otras series de N mediciones, pertenecerán a una distribución de Gauss alrededor del "valor verdadero". Y podemos interpretar el error estándar del promedio ξ como aquel valor que determina el intervalo alrededor del promedio, $\bar{x} \pm \xi$, dentro del cual el "verdadero valor" de la magnitud estará comprendido con una probabilidad del 68 %.

Cuando hago estadística reporto un promedio, entonces el error que tengo que reportar es el de los promedios

$$\epsilon_{est} = \xi = \frac{\sigma}{\sqrt{N}} \tag{9}$$

No en todos los procesos de medición los datos se distribuyen de acuerdo con una curva de Gauss. Además, también pueden aparecer errores sistemáticos u otras condiciones físicas que "distorsionan" la distribución de Gauss. La forma cualitativa más simple, para verificar si una distribución dada de datos es Gaussiana, consiste en comparar el histograma obtenido con la curva de Gauss teórica correspondiente.

2.5. Comentarios finales

- Dada una serie de mediciones, σ (desvío estándar) me indica que hay un 68 % de probabilidad de que una nueva medición caiga en el intervalo $(\bar{x} - \sigma, \bar{x} + \sigma)$.

- Dada una serie de promedios, ξ (desvío estándar del promedio, error estándar) me indica que, si realizo una nueva serie de mediciones, hay un 68 % de probabilidad de que el nuevo promedio caiga en el intervalo $(\bar{X} - \xi, \bar{X} + \xi)$, donde $\bar{X}$ es el promedio de los promedios.

- Cuando realizo estadística reporto el resultado del promedio $\bar{x}$ con una incerteza $\epsilon_{est} = \frac{\sigma}{\sqrt{N}}$.

- El error estadístico puede reducirse aumentando el número de mediciones pero su límite está dado por otras fuentes de error (por ejemplo, la incertidumbre instrumental). Se puede utilizar este concepto para predecir el número de mediciones que se deben realizar a fin de obtener una cota de incertidumbre similar a la incertidumbre instrumental.

- **Incerteza final:** recordar sumar todas las contribuciones. Por ejemplo, en una serie de mediciones donde solo haya error estadístico ϵ_{est} y error instrumental ϵ_{inst}, el error final será $\epsilon = \sqrt{\epsilon_{inst}^2 + \epsilon_{est}^2}$.

Referencias

[1] D. C. Baird, *Experimentación: Una introducción a la teoría de mediciones y al diseño de experimentos*, Prentice-Hall Hispanoaméricana S.A, México (1991).

[2] J. R. Taylor, *An introduction to error analysis: the study of uncertainties in physical measurements*, University Science Book, California (1997).

[3] J. G. Roederer, *Mecánica Elemental*, Eudeba, Universidad de Buenos Aires, Bs. As. (2005).