

A Estadística aplicada a las medidas nucleares

En un estudio estadístico se considera un conjunto, llamado *población*, de *elementos*, que se denominan *individuos*, de los cuales se estudia un aspecto determinado, que se denomina *carácter* y también *variable estadística* o simplemente *variable*.

Estas observaciones pueden incluir el total de la población o limitarse, por razones de economía, a una parte de los individuos que la constituyen, es decir a una *muestra*. En este último caso, a partir de los datos obtenidos para la muestra elegida, se infieren resultados para toda la población.

Los caracteres que se estudian se agrupan en: cuantitativos y cualitativos. Entre los cuantitativos se pueden distinguir los discretos y los continuos. Los caracteres discretos son aquellos que solamente presentan valores discretos (p.e. 0,1,2,3,... o también 12,3;13;14,5;16...). En cambio los caracteres continuos son aquellos que, a priori, pueden tomar cualquier valor real en un intervalo de valores usuales, sabiendo que, al hacer la medida, se dará una aproximación.

Las técnicas estadísticas son, en general, una herramienta de especial ayuda para el proceso de conocimiento de un fenómeno físico. Su aportación engloba varias etapas de este proceso:

- El diseño de experimentos: frecuentemente las conclusiones de un experimento bien diseñado se pueden extraer fácilmente, aun con métodos estadísticos elementales. Por el contrario, el análisis estadístico más sofisticado no puede compensar un experimento mal diseñado.
- Reducción de datos: en ocasiones los numerosos resultados obtenidos en la adquisición de medidas pueden ser representados mediante funciones de distribución que facilitan su comprensión y manejo.
- El proceso de inferencia: donde a partir de la información extraída de una muestra se estiman propiedades de la población.
- Construcción de modelos.
- Predicción de resultados.

A pesar del amplio alcance de las técnicas estadísticas, este anexo se va a centrar en aquellos conceptos que más usualmente se utilizan en las medidas nucleares.

Asimismo toda la exposición se va a focalizar básicamente en el análisis de variables discretas que adoptan valores enteros, como usualmente sucede en las aplicaciones nucleares.

A.1 Distribución estadística. Parámetros que caracterizan una distribución estadística

Supóngase que, al estudiar un carácter discreto de una población de N individuos, se ha obtenido el conjunto de datos: x_1, x_2, \dots, x_N que se considerarán ordenados en forma creciente, y entre los que pueden haber valores repetidos.

Para facilitar el análisis de colecciones de datos, especialmente cuando estos son numerosos, se acostumbra a organizarlos en forma de tablas. Indiquemos por x_1, x_2, \dots, x_m los diferentes valores que toma la variable.

A cada valor x_i se le asigna su *frecuencia absoluta* n_i , que es el número de veces que dicho dato aparece en la colección (ver tabla A.1). Está claro que la suma de todas las frecuencias absolutas coincide con el número total N de individuos de la población

$$N = \sum_{i=1}^m n_i,$$

donde el sumatorio se extiende al número de valores distintos que ha tomado la variable.

La interpretación de los datos queda reforzada si se calculan las proporciones del número de veces que aparece un dato x_i en relación al número total de individuos N de la población estudiada. Esta proporción se llama *frecuencia relativa*, y por tanto es igual al cociente entre su frecuencia absoluta y N , se suele indicar por $F(x_i)$, así pues

$$F(x_i) = \frac{n_i}{N}.$$

Estas informaciones suelen disponerse de la manera siguiente:

Tabla A.1 Distribución estadística

| Valores de la variable x_i | Frecuencia absoluta n_i | Frecuencia relativa $F(x_i)$ |
|---------------------------------|------------------------------|---------------------------------|
| x_1 | n_1 | n_1/N |
| x_2 | n_2 | n_2/N |
| . | . | . |
| . | . | . |
| . | . | . |
| x_m | n_m | n_m/N |
| | $\sum_{i=1}^m n_i = N$ | $\sum_{i=1}^m F(x_i) = 1$ |

Las frecuencias relativas facilitan también la comparación de diferentes series de datos.

El conjunto de datos, junto con sus frecuencias relativas, se denomina *distribución experimental de frecuencias*, o *distribución estadística*.

Ejemplo

Sea una colección de N medidas experimentales independientes de la misma variable física, que corresponden a los valores tomados por un determinado carácter en los N individuos de una cierta población,

8; 14; 5; 8; 12; 8; 10; 3; 13; 9; 7; 12; 9; 6; 10; 10; 6; 8; 11; 7; ($N = 20$) .

Estos datos se pueden estructurar tal como muestra la tabla A.2.

Esta tabla define una distribución estadística, y aporta toda la información contenida en la colección de datos inicial.

Para presentar la información que aportan los datos gráficamente, se representan los datos, x_i , en el eje de abscisas, y sobre cada uno de estos valores se eleva un segmento o barra vertical de longitud su frecuencia relativa.

Este gráfico recibe el nombre de *diagrama de barras*.

Tabla A.2 Ejemplo de distribución estadística

| Valores de la variable x_i | Frecuencia absoluta n_i | Frecuencia relativa $F(x_i)$ |
|---------------------------------|------------------------------|---------------------------------|
| 3 | 1 | $F(3) = 1/20 = 0,05$ |
| 5 | 1 | $F(5) = 1/20 = 0,05$ |
| 6 | 2 | $F(6) = 2/20 = 0,10$ |
| 7 | 2 | $F(7) = 2/20 = 0,10$ |
| 8 | 4 | $F(8) = 4/20 = 0,20$ |
| 9 | 2 | $F(9) = 2/20 = 0,10$ |
| 10 | 3 | $F(10) = 3/20 = 0,15$ |
| 11 | 1 | $F(11) = 1/20 = 0,05$ |
| 12 | 2 | $F(12) = 2/20 = 0,10$ |
| 13 | 1 | $F(13) = 1/20 = 0,05$ |
| 14 | 1 | $F(14) = 1/20 = 0,05$ |
| | ----- 20 | ----- 1 |

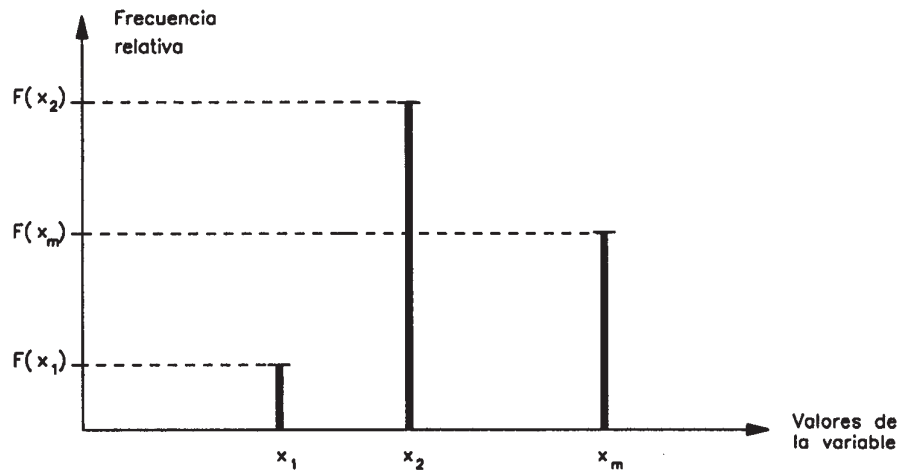


Fig. A.1 Diagrama de barras

Ejemplo

Considerando la colección de las 20 medidas experimentales del ejemplo anterior, su representación gráfica vendrá dada por la figura A.2.

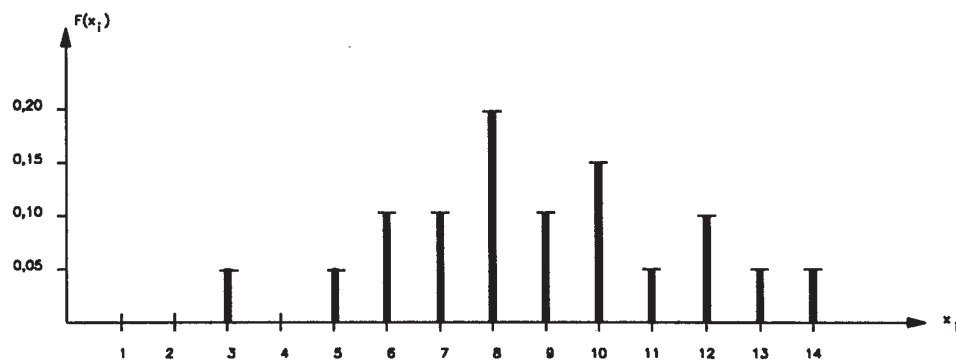


Fig. A.2 Diagrama de barras de la distribución estadística de la tabla A.2

Cuando se trabaja con distribuciones, conviene poderlas describir mediante un pequeño número de valores, denominados *características* o *parámetros de la distribución*, que representan algunos rasgos esenciales de la misma. Ejemplos de propiedades esenciales son la *posición* de la distribución, (normalmente caracterizada por medio de la abscisa de algún punto sobre el cual está centrada la distribución), el grado de *dispersión* de los datos alrededor del valor central, y el grado de *asimetría* de la distribución.

La medida de posición más importante es la *media* que no es más que el promedio ordinario de la colección de números que constituyen todos los datos.

Entre las características de dispersión destaca la *varianza*, que mide lo alejados o aproximados que están los datos respecto de la media.

En primer lugar, se considerará el caso en que el conjunto de individuos analizados constituye toda la población.

Si todos los N valores fuesen distintos, para obtener su media se efectúa su suma y se divide por el número total N de individuos de la población. Pero, teniendo en cuenta que se pueden presentar valores repetidos, con las notaciones introducidas anteriormente, se definen la *media de la población*, \bar{x} , y la *varianza de la población*, σ^2 , como:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^m n_i x_i = \sum_{i=1}^m x_i F(x_i), \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^m (x_i - \bar{x})^2 n_i = \sum_{i=1}^m (x_i - \bar{x})^2 F(x_i),$$

donde los sumatorios se extienden al número de valores distintos que toma el carácter.

Si se desarrolla la expresión que da la varianza se obtiene otra, que simplifica los cálculos,

$$\sigma^2 = \overline{x^2} - (\bar{x})^2$$

con

$$\overline{x^2} = \frac{1}{N} \sum_{i=1}^m n_i x_i^2.$$

La varianza es una cantidad de segundo orden en unidades de la variable, por tanto, en general es más adecuado tomar como medida de dispersión su raíz cuadrada. Se denomina *desviación típica* y viene dada por la expresión

$$\sigma = \sqrt{\sum_{i=1}^m (x_i - \bar{x})^2 F(x_i)} = \sqrt{\overline{x^2} - (\bar{x})^2}.$$

Ejemplo

Para la distribución dada en la tabla A.2, el cálculo de la media y de la varianza se lleva a cabo de la forma

$$\bar{x} = \frac{1}{20} \sum_{i=1}^{11} n_i x_i = \sum_{i=1}^{11} x_i F(x_i) = 8,8 ; \quad \sigma^2 = \frac{1}{20} \sum_{i=1}^{11} (x_i - 8,8)^2 \cdot n_i = \sum_{i=1}^{11} (x_i - 8,8)^2 F(x_i) = 7,36 .$$

Muchas veces las desviaciones sólo incluyen una parte de los individuos de la población, es decir, una *muestra*.

En este caso, si se supone que el *tamaño de la muestra* -número de individuos que la constituyen- es n , con las mismas notaciones anteriores, se definen la *media muestral* o *media experimental* y la *varianza muestral* como

$$\bar{x}_e = \frac{\sum_{i=1}^m n_i x_i}{n} = \sum_{i=1}^m x_i F(x_i) , \quad s_e^2 = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^2 F(x_i) ,$$

donde \bar{x} es la media de la población.

Dado que, en general, se desconoce el valor de \bar{x} , puede estimarse la varianza de la muestra, utilizando la media experimental, a través de la expresión

$$s_e^2 \approx s_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^m n_i (x_i - \bar{x}_e)^2 .$$

El uso de la media experimental, en lugar de la media de la población, tiende a disminuir el valor medio de las diferencias $x_i - \bar{x}$ y por tanto resultaría una varianza inferior a la normal. Se dice entonces que el número de grados de libertad se ha reducido en una unidad y se introduce el término -1 en el denominador para minimizar este efecto.

La forma del gráfico de la distribución estadística indica cualitativamente la cantidad de dispersión en el conjunto de datos. Por ejemplo, la figura A.3 muestra la forma de dos diagramas de barras correspondientes a dos grupos de datos extremos: uno con una gran dispersión sobre la media y el otro con poca.

Las características anteriormente definidas, aunque no determinan completamente la distribución estadística, permiten hacerse una buena idea de ella, compararla fácilmente con otras distribuciones y, sobre todo, deducir conclusiones importantes sobre la población.

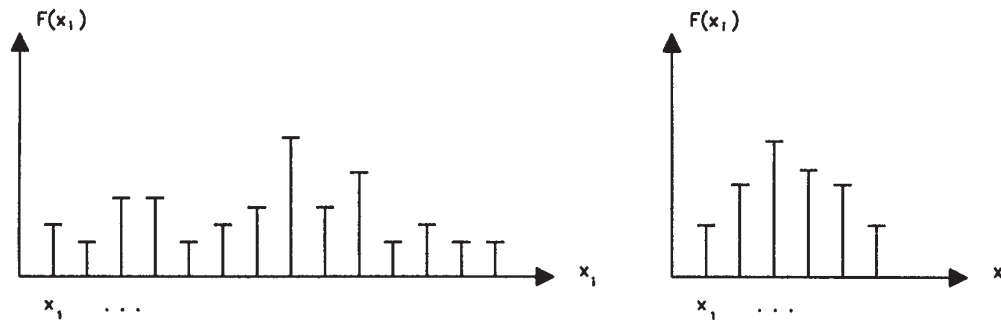


Fig. A.3 Diagrama de barras para dos grupos de datos con diferente cantidad de fluctuación interna

A.2 Modelos estadísticos

En este apartado se muestra que, con ciertas condiciones, se puede predecir la función de distribución que describe los resultados de varias repeticiones de una medida nuclear.

Considérese una *medida nuclear* donde, durante un tiempo t , se observa una muestra de N_0 núcleos radiactivos, y se pretende medir el número de núcleos que se han desintegrado durante dicho tiempo t . El resultado de esta medida se determinará a partir de la lectura del detector tomando en consideración la eficiencia de detección.

Para aplicar correctamente las herramientas estadísticas, es conveniente definir varios términos que van a ser utilizados de ahora en adelante.

A.2.1 Experiencia aleatoria. Suceso. Probabilidad de un suceso

Estadísticamente, la *medida nuclear* del ejemplo anterior debe entenderse como la realización de N_0 ensayos simultáneos de una *experiencia aleatoria*.

Una *experiencia aleatoria* es una experiencia que cumple:

- se puede repetir tantas veces como se quiera en las mismas condiciones;
- el resultado de cada repetición no se puede conocer a priori.

La tabla A.3 muestra tres ejemplos de experiencias aleatorias.

Los resultados posibles que se obtienen al efectuar una experiencia aleatoria se denominan: los *casos posibles* de la experiencia. Los casos posibles constituyen el nivel más detallado de los resultados que pueden considerarse. Cada vez que se repite la experiencia, se obtiene uno y sólo uno de los casos posibles.

Tabla A.3 Ejemplos de experiencias aleatorias

| EXPERIENCIA ALEATORIA | SUCESO | PROBABILIDAD DEL SUCESO |
|---|--|-------------------------|
| Lanzar una moneda | Obtener cara | 1/2 |
| Lanzar un dado | Obtener un número mayor que cuatro | 1/3 |
| Observar la desintegración de un núcleo radiactivo dado durante un tiempo "t" | El núcleo se desintegra durante la observación | $1 - e^{-\lambda t}$ |

Ejemplo

En la experiencia aleatoria *lanzar un dado*; hay seis casos posibles: obtener un uno, un dos, un tres, un cuatro, un cinco y un seis.

En la experiencia aleatoria *lanzar una moneda*, hay dos casos posibles: sacar cara, sacar cruz.

En la experiencia aleatoria *observar* la desintegración de un núcleo radiactivo dado durante un tiempo t, hay dos casos posibles: el núcleo se desintegra, el núcleo no se desintegra.

Un hecho que puede o no darse en una experiencia aleatoria se denomina un *suceso* de dicha experiencia. Los casos posibles también reciben el nombre de *sucesos elementales*.

Ejemplo

En la experiencia aleatoria *lanzar una moneda* un suceso sería *obtener cara*, otro *obtener cruz*, etc.

En la experiencia aleatoria *lanzar un dado* un suceso puede ser obtener un tres; otro, obtener un número par, etc.

En la experiencia aleatoria *observar* la desintegración de un núcleo radiactivo dado durante un tiempo t un suceso puede ser: el núcleo se desintegra durante la observación.

Un determinado suceso de una experiencia aleatoria se puede o no realizar al efectuar dicha experiencia. Los casos posibles que la realizan se denominan *casos favorables* a la realización del suceso.

Ejemplo

En la experiencia *lanzar una moneda* y para el suceso *sacar cara*, hay 1 caso favorable.

En la experiencia *lanzar un dado* y para el suceso *obtener un número par*, hay 3 casos favorables. En cambio, para el suceso *obtener un número mayor que 4* hay 2 casos favorables.

Supóngase que la experiencia aleatoria se repite N veces. Se llama *frecuencia absoluta de un suceso* A , y se puede indicar por $n(A)$, al número de veces que se ha realizado dicho suceso entre las N repeticiones de la experiencia. Se llama *frecuencia relativa del suceso* A , al cociente entre su frecuencia absoluta y el número total de veces que se ha repetido la experiencia. Se acostumbra a indicar por $f(A)$, y así de forma simbólica se escribe

$$f(A) = \frac{n(A)}{N} .$$

Algunas de las propiedades de las frecuencias relativas son:

- para cualquier suceso A , siempre se cumple que $0 \leq f(A) \leq 1$;
- la suma de la frecuencia relativa de todos los sucesos elementales vale 1.

Hay una ley que regula los fenómenos en que interviene el azar, fenómenos como los de los ejemplos anteriores. Es la llamada *ley empírica del azar*:

"En una experiencia aleatoria, a medida que el número de veces que se repite va siendo suficientemente grande, la frecuencia relativa de un suceso tiende a estabilizarse alrededor de un valor constante".

Admitida esta ley, a este valor constante se le llama *probabilidad del suceso*. Indicando dicho suceso por la letra A , su probabilidad se acostumbra a representar por $p(A)$.

Está claro que esta formulación conlleva algunas imprecisiones ya que no queda claro cuándo se podrá considerar que el número de veces que se ha repetido la experiencia es suficientemente grande, ni cuándo se podrá afirmar que la frecuencia relativa se ha estabilizado.

Sin embargo, hay algunas experiencias aleatorias en que razones de simetría y homogeneidad permiten formular un modelo teórico. Son aquellas en que a cada uno de los sucesos elementales se le puede asociar la misma probabilidad, o sea, que todos los sucesos elementales son equiprobables. Para este

tipo de experiencias se cumple la denominada *Ley de Laplace* y también la definición clásica de probabilidad: "La probabilidad de un suceso es el cociente entre el número de casos favorables a su realización y el número de casos posibles cuando se pueda asegurar que éstos son equiprobables".

$$p(A) = \frac{\text{número de casos favorables a la realización de } A}{\text{número de casos posibles}} .$$

Es importante resaltar el hecho de que esta definición se basa en la hipótesis de la equiprobabilidad de los sucesos elementales y sólo se puede aplicar en aquellas situaciones en que no se tengan dudas sobre la verosimilitud de dicha hipótesis.

Tal es el caso de las dos primeras experiencias aleatorias de la tabla A.3.

Ejemplo

En la experiencia *lanzar una moneda*, la probabilidad del suceso *obtener cara* es $1/2$, porque hay un caso favorable a su realización y dos casos posibles que son equiprobables.

En la experiencia *lanzar un dado*, en la que los sucesos elementales pueden suponerse equiprobables, las probabilidades de los sucesos

$$A = \text{obtener un número par} , \quad B = \text{obtener un número mayor que 4}$$

son

$$p(A) = \frac{3}{6} = \frac{1}{2} \quad ; \quad p(B) = \frac{2}{6} = \frac{1}{3} .$$

Pero no siempre es posible afirmar a priori la equiprobabilidad de los sucesos elementales. Así, por ejemplo en la experiencia "observar la desintegración de un núcleo radiactivo dado durante un tiempo t " los dos casos posibles son

$$\begin{aligned} A_1 &= \text{el núcleo se desintegra durante la observación,} \\ A_2 &= \text{el núcleo no se desintegra durante la observación,} \end{aligned}$$

y no puede suponerse que son sucesos equiprobables. De forma análoga que tampoco lo son los casos posibles de la experiencia aleatoria que considera los accidentes de circulación en una determinada vía a una determinada hora de un día laborable. Los dos casos posibles

$$\begin{aligned} A_1 &= \text{tener accidente,} \\ A_2 &= \text{no tener accidente,} \end{aligned}$$

está claro que no son equiprobables.

En estos casos para asignar una probabilidad es necesario recurrir a las sucesivas repeticiones de las experiencias y usar la ley empírica del azar.

Ejemplo

En la experiencia "observar la desintegración de un núcleo radiactivo durante un tiempo t ", considérense los dos sucesos

- A_1 = el núcleo se desintegra durante el tiempo t ,
- A_2 = el núcleo no se desintegra durante el tiempo t .

Tal como se ha indicado al comienzo de este apartado, en la práctica se observa durante un tiempo t una muestra de núcleos radiactivos que, en general, se podrá considerar que contiene un elevado número N_0 de núcleos. Esta observación es equivalente a la realización de N_0 ensayos de la experiencia "observar la desintegración de un núcleo radiactivo durante un tiempo t ".

La ley de la desintegración radiactiva permite determinar el número de veces que se ha realizado cada uno de los dos sucesos A_1 y A_2 en los N_0 ensayos.

Así, dicha ley da el número de núcleos presentes en la muestra, después de transcurrido el tiempo t de observación

$$N(t) = N_0 e^{-\lambda t},$$

donde N_0 es el número de núcleos iniciales, $N(t)$ el número de núcleos en el instante t y λ la constante de desintegración radiactiva de los núcleos observados.

En consecuencia el número de núcleos desintegrados es

$$N_0 - N_0 e^{-\lambda t} = N_0(1 - e^{-\lambda t})$$

Por consiguiente, según la ley del azar se pueden asignar a los sucesos A_1 y A_2 las probabilidades

$$p(A_1) = \frac{N_0(1 - e^{-\lambda t})}{N_0} = 1 - e^{-\lambda t}, \quad p(A_2) = \frac{N_0 e^{-\lambda t}}{N_0} = e^{-\lambda t}.$$

El tercer ejemplo de la tabla A.3 muestra las bases para aplicar los modelos estadísticos que siguen al recuento de procesos radiactivos. En este caso:

- una experiencia aleatoria consiste en observar la desintegración de un núcleo radiactivo dado durante un tiempo t ;
- el tamaño de la muestra es el número de ensayos efectuados de esta experiencia aleatoria, y es equivalente al número de núcleos en observación;

- una *medida nuclear* consiste en contabilizar aquellos núcleos que se desintegran en el tiempo t .

A.2.2 Distribuciones de probabilidad

Los sucesos correspondientes a las experiencias citadas en el apartado anterior eran de diferentes tipos: obtener cara y cruz en la moneda, sacar 1,2,...,6 puntos en el dado, etc. Algunos de esos sucesos se podían representar numéricamente y otros no. En general, el estudio es más cómodo en el primer caso, por ello se suele asignar un número a cada suceso de una experiencia aleatoria. Esto lleva a introducir el concepto de *variable aleatoria* (v.a.).

Una variable aleatoria es una función que asigna un número real a cada uno de los sucesos de una experiencia aleatoria.

Se llama aleatoria por estar su dominio formado por los sucesos de una experiencia aleatoria, y variable puesto que toma valores numéricos que pueden variar de un suceso a otro. Normalmente se utilizan letras mayúsculas, X, Y, \dots para designarlas.

Ejemplo

En la experiencia de tirar un dado, es natural hacer corresponder a los sucesos elementales: sacar 1, sacar 2... los valores 1,2,...6. Si indicamos por X esta variable aleatoria, se tiene que la v.a. X , que describe esta experiencia es:

sacar 1 \rightarrow 1,
sacar 2 \rightarrow 2,
sacar 3 \rightarrow 3,
sacar 4 \rightarrow 4,
sacar 5 \rightarrow 5,
sacar 6 \rightarrow 6.

Y así, se puede escribir $p(X=4)$ para indicar la probabilidad de que se realice alguno de los sucesos que tienen asignado el número 4; o bien $p(1 \leq X \leq 5)$ para indicar la probabilidad de que el resultado de la experiencia tenga asignado un número comprendido entre 1 y 5 ambos incluidos.

En experiencias en las que se estudian caracteres cualitativos, la asignación no es inmediata. Por ejemplo, si nos interesamos en los sucesos sacar cara o cruz al tirar una moneda, se pueden asignar el valor 0 al suceso sacar cara y el 1 al suceso sacar cruz.

Se tendrá pues la v.a.:

cara \rightarrow 0,
cruz \rightarrow 1.

Una variable aleatoria se llama *discreta* si tiene un número finito de valores posibles, o bien, de tener un número infinito, pueden ordenarse secuencialmente.

Una variable aleatoria se llama *continua* si sus posibles valores son todos los números reales de un intervalo.

En el apartado 1 se analizaban *distribuciones empíricas de frecuencias* que correspondían a la totalidad de la población o a una muestra. En el segundo caso, se trataba de una estimación de la distribución de la población correspondiente.

En muchos de los problemas estadísticos la muestra no es lo suficientemente grande para poder determinar la distribución de la población con mucha precisión. Pero, en general, existe información obtenida, experimentalmente, de otras fuentes. La combinación de esta información con la dada por la muestra, permite, normalmente, postular la naturaleza general de la distribución de la población, postulados que dan lugar a las distribuciones de probabilidad.

Una *distribución de probabilidad* es un modelo matemático para la distribución de frecuencia real.

De modo análogo al caso de las distribuciones de frecuencias se define el concepto de *distribución de probabilidad de una variable aleatoria discreta*.

Supóngase que la variable aleatoria X puede tomar los valores x_i , y que se simboliza por $p(X = x_i)$ la probabilidad de que la variable tome el valor x_i . Así, $p(X = x_i)$ es la probabilidad de obtener como resultado de la experiencia alguno de los sucesos que tienen asignado el valor x_i .

Una distribución de probabilidad f de una variable aleatoria discreta X es la función que hace corresponder a cada valor que toma X un número real $f(x_i)$ del intervalo $[0, 1]$ tal que

$$f(x_i) = p(X = x_i).$$

Se denomina también *función de probabilidad*.

Ejemplo

Considérese la experiencia aleatoria de tirar 2 dados, y la variable aleatoria X que asigna a cada suceso la suma de puntos de los dos dados, suma que varía de 2 a 12.

La función de probabilidad viene definida por:

$$\begin{array}{ll} x_1 = 2 \rightarrow f(2) = p(X=2) = 1/36, & x_7 = 8 \rightarrow f(8) = p(X=8) = 5/36, \\ x_2 = 3 \rightarrow f(3) = p(X=3) = 2/36, & x_8 = 9 \rightarrow f(9) = p(X=9) = 4/36, \\ x_3 = 4 \rightarrow f(4) = p(X=4) = 3/36, & x_9 = 10 \rightarrow f(10) = p(X=10) = 3/36, \\ x_4 = 5 \rightarrow f(5) = p(X=5) = 4/36, & x_{10} = 11 \rightarrow f(11) = p(X=11) = 2/36, \\ x_5 = 6 \rightarrow f(6) = p(X=6) = 5/36, & x_{11} = 12 \rightarrow f(12) = p(X=12) = 1/36, \\ x_6 = 7 \rightarrow f(7) = p(X=7) = 6/36, & \end{array}$$

Evidentemente se verifica

$$f(x_i) \geq 0 \text{ para cada } x_i,$$

$$\sum_{i=1}^{i=11} f(x_i) = 1 .$$

La forma como se representa gráficamente esta función es análoga al caso de las distribuciones de frecuencia, substituyendo la frecuencia relativa por la probabilidad (Fig. A.4).

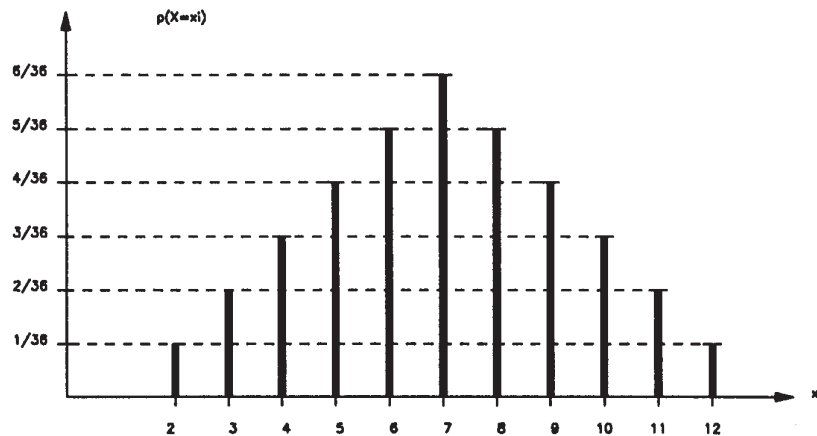


Fig. A.4 Función de probabilidad

A.2.3 Los parámetros de una distribución de probabilidad

De manera análoga al caso de las distribuciones de frecuencia se definen para las distribuciones de probabilidad unos parámetros que tienden a determinar numéricamente el centro y la dispersión de estas distribuciones. Nótese que las distribuciones de probabilidad son semejantes a las de frecuencia si se considera que el papel jugado por la probabilidad en las primeras se corresponde con el de las frecuencias relativas en estas últimas.

Así, se definen para una v.a, X , en términos de la función de probabilidad f correspondiente, los siguientes parámetros:

- la *media* como

$$\bar{x} = \sum x_i f(x_i) ,$$

donde el sumatorio se extiende a todos los posibles valores x_i de la v.a.;

- la *varianza* como

$$\sigma^2 = \sum (x_i - \bar{x})^2 f(x_i) .$$

La media de una v.a. X también se denomina *valor esperado* o *esperanza* y se representa por $E(X)$.

La raíz cuadrada positiva de la varianza se denomina *desviación típica* o *desviación estándar*.

Ejemplo

En el caso de la variable aleatoria X que asigna a cada suceso de la experiencia tirar 2 dados, la suma de puntuaciones obtenida, se tiene

$$\bar{x} = \sum_{i=1}^{i=11} x_i \cdot f(x_i) = 7 , \quad \sigma^2 = \sum_{i=1}^{i=11} (x_i - \bar{x})^2 \cdot f(x_i) = 5.83 .$$

A.2.4 La distribución binomial

Entre las distribuciones discretas de probabilidad la más importante por sus aplicaciones es la distribución binomial.

La distribución binomial es el modelo estadístico asociado a la variable aleatoria "contar el número de veces que se realiza un suceso A cuando se repite n veces, de forma independiente y en las mismas condiciones, un experimento aleatorio del que sólo se estudia la verificación o no de dicho suceso".

Si n es el número de ensayos de una experiencia aleatoria, y para cada ensayo de esa experiencia hay una probabilidad p de que se realice un determinado suceso, la probabilidad de que se realice exactamente x veces ese suceso es

$$P(x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} .$$

La función que hace corresponder a cada valor de x la probabilidad $P(x)$ es la función de probabilidad binomial, y está definida solamente para valores enteros de n y x .

Desafortunadamente, esta función es engorrosa de utilizar, sobre todo en los procesos de desintegración radiactiva donde el número de núcleos es muy grande (n es muy grande), por consiguiente es utilizada muy raramente en aplicaciones nucleares. Un ejemplo en el que la

distribución binomial debe ser usada, es en el análisis de datos adquiridos por recuento de un isótopo de vida muy corta, mediante un detector de eficiencia alta.

Dos propiedades importantes que se deducen de la distribución binomial son las expresiones que permiten estimar la media y la varianza,

$$\bar{x} = np, \quad \sigma^2 = \bar{x} \cdot (1-p).$$

En el caso de que n sea muy grande el uso de esta distribución, como ya se ha indicado, no es útil ya que requiere tediosas operaciones. Ello pone de manifiesto la necesidad de utilizar modelos matemáticos más manipulables. Estudios en esta dirección estuvieron en la génesis de la distribución de Poisson y la de Gauss o Normal.

A.2.5 La distribución de Poisson

Muchos procesos pueden ser caracterizados por la baja probabilidad de que un determinado suceso se realice en cada ensayo de la experiencia. En esta categoría están incluidos muchos experimentos nucleares de recuento, en los cuales la presencia de gran número de núcleos (del orden del número de Avogadro) hace aumentar el tamaño de la muestra o el número de ensayos de la experiencia aleatoria, mientras que una fracción relativamente pequeña de éstos origina un recuento. En estas condiciones la aproximación $p \ll 1$ es válida.

En el límite, cuando en una distribución binomial se hace que p tienda a cero, mientras n tiende a infinito, de tal manera que np tiende a una constante ν , se puede probar que la distribución binomial se reduce a la forma

$$P(x) = \frac{\nu^x e^{-\nu}}{x!}.$$

En el paso al límite aquí considerado, la distribución binomial tiende a una distribución límite de tipo discreto donde los valores posibles de la v.a. X forman una sucesión indefinida $x = 0, 1, 2, \dots$ a los que se les asigna la probabilidad $p(X=x)$ dada por la expresión anterior.

De este resultado se puede deducir que:

- la distribución está normalizada

$$\sum_{x=0}^{\infty} P(x) = 1 ;$$

- la media es

$$\bar{x} = \sum_{x=0}^{\infty} x \cdot P(x) = \nu ;$$

- la varianza es

$$\sigma^2 = \sum_{x=0}^{\infty} (x-\bar{x})^2 \cdot P(x) = \bar{x} .$$

Así, la distribución de Poisson se escribe

$$P(x) = \frac{(\bar{x})^x e^{-\bar{x}}}{x!} .$$

Es interesante señalar que para esta distribución, la desviación tipo σ es la raíz cuadrada de la media \bar{x}

$$\sigma = \sqrt{\bar{x}} .$$

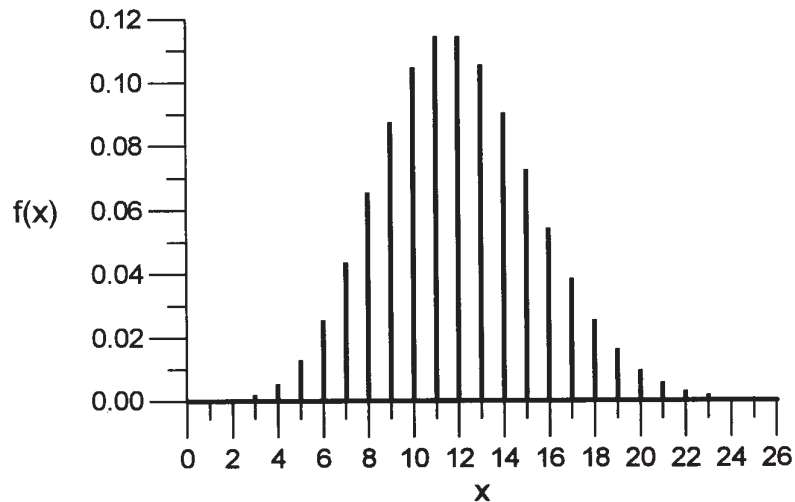


Fig. A.5 Distribución de Poisson con $\bar{x} = np = 12$

Así, mientras que una distribución binomial requiere *conocer* los valores de dos parámetros: el número de ensayos de la experiencia aleatoria n y la probabilidad p de que se realice un determinado suceso, una distribución de Poisson sólo requiere un parámetro, la media \bar{x} . Ello representa una gran ayuda en el análisis de los procesos en los cuales se desconoce la probabilidad p o el tamaño de la muestra n , y en los que sólo se sabe que n es muy elevada y p muy pequeña, pero que por algún método se mide o estima el valor medio. Este es el caso usual en medidas nucleares.

En la figura A.5 viene representada una distribución de Poisson con $\bar{x} = n \cdot p = 12$. En ella los valores posibles de la variable forman la sucesión indefinida $x = 0, 1, 2, \dots$

A.2.6 La distribución Normal

En el apartado anterior se ha analizado la conducta de una distribución binomial, cuando se hace tender p a cero, mientras n tiende a infinito de manera que $n \cdot p$ tiende a una constante, y se ha visto que entonces la distribución binomial se transforma en otra distribución discreta: la distribución de Poisson.

Realizando un paso al límite de forma algo distinta, resulta otra distribución límite interesante pero de tipo continuo: la distribución normal o de Gauss.

Considérese una distribución binomial

$$P(x) = \frac{n!}{(n-x)! x!} p^x (1-p)^{n-x},$$

cuya media y varianza serán, según se ha visto en el apartado 4,

$$\bar{x} = np, \quad \sigma^2 = np(1-p).$$

Introdúzcase una *variable normalizada o tipificada* correspondiente a la variable X ,

$$\frac{X - \bar{x}}{\sigma},$$

que expresa la desviación de X respecto a su valor medio, medida en unidades de la desviación típica. Esta nueva variable tiene un valor medio nulo y una varianza igual a la unidad.

De esta forma se obtiene una distribución binomial tipificada. Pues bien, si se hace que n tienda a infinito mientras que p se mantiene constante, se puede probar que la distribución binomial tipificada, que es una distribución discreta, se transforma entonces en una distribución límite continua: la distribución normal. Esta distribución viene caracterizada por la ley

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} .$$

Esta función tiene la forma denominada *campana de Gauss*, cuyos rasgos más característicos son (ver figuras A.6 y A.7):

- presenta un eje de simetría, que es la recta $x = \bar{x}$;
- presenta un único máximo que corresponde a una abscisa $x = \bar{x}$;
- el valor de σ determina la forma de la campana
 - . valores pequeños de σ hacen que la curva sea más *estrecha*, con valores más elevados del máximo;
 - . valores grandes de σ hacen que la curva sea más *ancha* con valores más pequeños del máximo.

La función f es un ejemplo de la denominada *función densidad de probabilidad* asociada a una variable aleatoria continua, que substituye a la función de probabilidad del caso discreto, función que daba la probabilidad de que la variable tomase cada uno de los valores que eran posibles.

En general, para una variable aleatoria continua, la función densidad de probabilidad es una función f tal que la probabilidad de que la variable X tome un valor que pertenezca al intervalo cerrado de extremos a y b es, (ver Fig. A.8),

$p [a \leq X \leq b] =$ área de la zona limitada por el gráfico de la función, f el eje de las x y las ordenadas correspondientes a $x = a$ y $x = b$.

De manera formal la anterior igualdad se indica

$$p [a \leq x \leq b] = \int_a^b f(x) dx .$$

Es importante resaltar que si x es uno de los valores que puede tener X , se tiene que

$$p[X = x] = 0 .$$

En el caso de la distribución normal o en una situación práctica que se pueda modelizar con esta

distribución, la interpretación de la desviación típica como una medida de la dispersión de los datos respecto a la media tiene un gran interés.

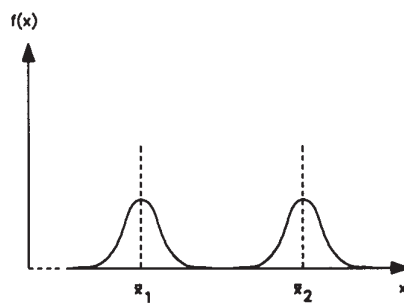


Fig. A.6 Dos curvas normales con distinta media y la misma σ

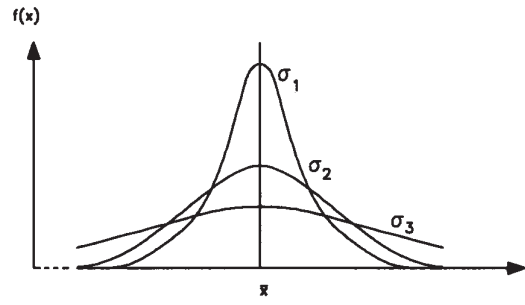


Fig. A.7 Curvas normales con la misma media y distinta σ

En efecto, para una variable aleatoria con distribución normal se verifica

$$p [\bar{x} - \sigma \leq X \leq \bar{x} + \sigma] = 0,683,$$

$$p [\bar{x} - 2\sigma \leq x \leq \bar{x} + 2\sigma] = 0,954,$$

$$p [\bar{x} - 3\sigma \leq x \leq \bar{x} + 3\sigma] = 0,997,$$

es decir, la probabilidad de que la variable aleatoria tome un valor comprendido en una zona que abarca una σ a izquierda y derecha de la media es de 0,683; que abarque 2σ es de 0,954 y que abarque 3σ es de 0,997.

Se observa que:

- hay una fuerte agrupación de los datos en torno a la media;
- el intervalo de valores significativos se reduce, prácticamente, para la distribución teórica, al intervalo $[\bar{x} - 3\sigma, \bar{x} + 3\sigma]$.

Estas probabilidades para distintos intervalos están tabuladas. Algunos de estos valores, para distintos intervalos centrados en la media se han recogido en la tabla A.4.

Como ejemplo ilustrativo se presenta el caso de una distribución normal de media $\bar{x} = 27,4$ y desviación típica $\sigma = 5,23$. Su función densidad de probabilidad es la de la figura A.9.

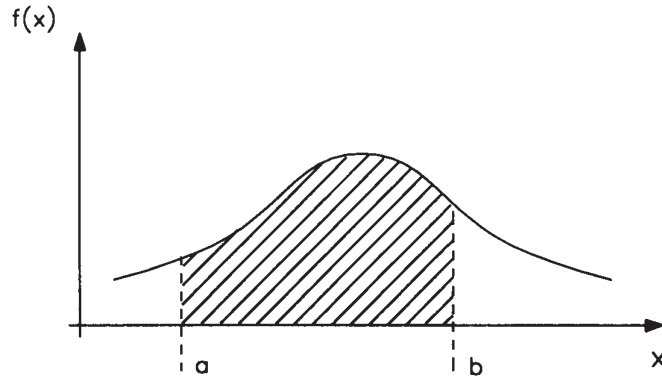


Fig. A.8 Ilustración gráfica de la probabilidad $p[a \leq X \leq b]$, que la variable aleatoria X tome un valor que pertenezca al intervalo cerrado de extremos a y b

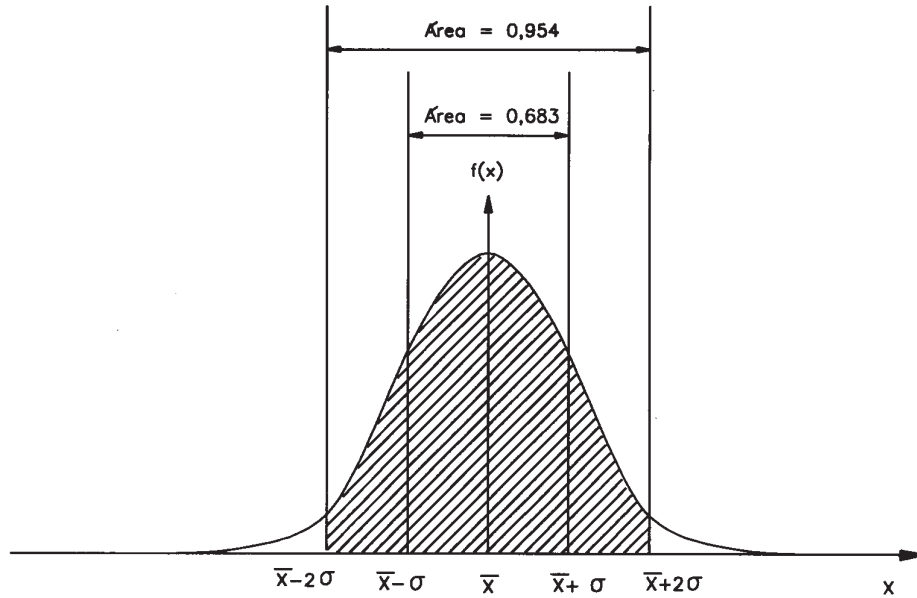
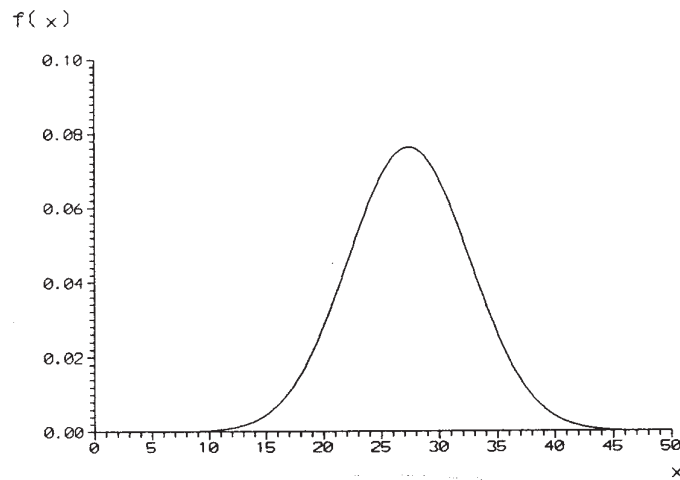


Fig. A.9 Interpretación de la desviación típica como una medida de la dispersión de los datos respecto a la media

Tabla A.4 Probabilidad según la distribución de Gauss

| $p(\bar{x}-k \cdot \sigma < x < \bar{x}+k \cdot \sigma)$ [%] | $k \in \mathbb{R}^+$ |
|---|----------------------|
| 50,0 | 0,674 |
| 68,3 | 1,000 |
| 90,0 | 1,640 |
| 95,0 | 1,960 |
| 95,4 | 2,000 |
| 99,0 | 2,580 |
| 99,7 | 3,000 |

Fig. A.10 Función distribución de Gauss o Normal, para un valor medio $\bar{x} = 27,4$ y $\sigma = 5,23$

La tabla A.5 muestra para esta distribución algunos intervalos centrados en torno a la media y las probabilidades asociadas a cada uno.

Tabla A.5 Probabilidad según la distribución de Gauss. Intervalos de incertidumbre. Aplicación a una distribución de media $\bar{x} = 27,4$ $\sigma = 5,23$

| INTERVALO | INTERVALO | $p(\bar{x} - k \cdot \sigma < x < \bar{x} + k \cdot \sigma)$ |
|----------------------------|---------------|--|
| $\bar{x} \pm 0,674 \sigma$ | 23,87 - 30,93 | 50,0 % |
| $\bar{x} \pm \sigma$ | 22,16 - 32,63 | 68,3 % |
| $\bar{x} \pm 1,64 \sigma$ | 18,82 - 35,98 | 90,0 % |
| $\bar{x} \pm 1,96 \sigma$ | 17,14 - 37,66 | 95,0 % |
| $\bar{x} \pm 2 \sigma$ | 16,94 - 37,86 | 95,4 % |
| $\bar{x} \pm 2,58 \sigma$ | 13,89 - 40,90 | 99,0 % |
| $\bar{x} \pm 3,00 \sigma$ | 11,70 - 43,10 | 99,7 % |

Es importante disponer de algún criterio para poder decidir cuándo se puede aproximar una distribución binomial por una distribución normal.

En general, una distribución binomial de parámetros n y p tales, que np y $n(1-p)$ sean valores *grandes* puede ser aproximada correctamente por una distribución normal de media $\bar{x} = np$ y de desviación típica $\sigma = \sqrt{np(1-p)}$. En general, se puede pensar que si $np \geq 5$ y $n(1-p) \geq 5$, la aproximación ya es aceptable. Ahora bien, muchos tratados de estadística, para obtener un mayor grado de fiabilidad, imponen la condición $np \geq 15$ y $n(1-p) \geq 15$.

Para que la aproximación sea correcta es necesario considerar la llamada *corrección de continuidad*: si X_b es la variable aleatoria con distribución binomial y X_n la variable aleatoria con distribución normal con la que se modela, los valores análogos serán

$$p[X_b = x] \approx p[x - 1/2 \leq X_n \leq x + 1/2]$$

A.3 La propagación de errores. Aplicación al recuento de muestras radiactivas

Es bastante *convencional* en las medidas nucleares expresar la contribución estadística de la *incertidumbre o error* de una simple medida como el valor de la desviación estándar σ . Así el resultado de una medida se suele dar en la forma $x \pm \sigma$.

En el apartado anterior se han presentado los modelos estadísticos que permiten estimar la contribución estadística de la incertidumbre asociada a una magnitud cuando ésta es medida directamente.

En el caso de medidas nucleares, cuando se posee una única medida "x" de un proceso de recuento, si ésta es suficientemente grande, se considera que es la mejor estimación de la media y se acepta $\bar{x} = x$.

Adoptando una distribución de Poisson se determina la varianza, $\sigma^2 = \bar{x} = x$. Por tanto, la incertidumbre o error asociado a una simple medida se podrá expresar como

$$x \pm \sigma \text{ es decir } x \pm \sqrt{x} .$$

Es importante observar que este resultado *sólo es válido* si la magnitud x representa el número de veces que se ha realizado un suceso, cuando se han efectuado n ensayos de una experiencia aleatoria, es decir, x representa el número de contajes registrados por un detector durante un tiempo de observación. La mayoría de las equivocaciones que se producen en los recuentos estadísticos provienen de una mala aplicación del resultado anterior.

Esta estimación de la incertidumbre no puede aplicarse directamente a:

- tasas de recuento;
- sumas o diferencias de recuentos;
- medias de recuentos independientes;
- cualquier magnitud derivada.

En todos estos casos, el error asociado con estas magnitudes debe ser calculado mediante la teoría de propagación de errores.

A.3.1 Propagación de errores

Frecuentemente no puede medirse la magnitud de interés directamente, sino que debe deducirse de otras magnitudes medibles. Si se denomina U a la magnitud de interés y x, y, z, \dots son las magnitudes que se miden, el problema es determinar el error en el valor calculado de $U = F(x, y, z, \dots)$, debido a las incertidumbres de medida de las magnitudes x, y, z, \dots , cuyo error sí que es conocido ($\sigma_x, \sigma_y, \sigma_z, \dots$).

Si las magnitudes x, y, z, \dots son estadísticamente independientes, es decir no están correlacionadas, la incertidumbre en U será

En la tabla A.6 se presentan algunos casos particulares correspondientes a funciones sencillas, suponiendo que las variables x , e y y son independientes.

Tabla A.6 Fórmulas de propagación de errores para algunas funciones sencillas (x e y estadísticamente independientes)

| FUNCIÓN $U = F(x,y)$ | VARIANZA σ_u^2 |
|----------------------|---|
| $Ax + By$ | $A^2\sigma_x^2 + B^2\sigma_y^2$ |
| $x \cdot y$ | $(x \cdot y)^2 \left(\frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2} \right)$ |
| $\frac{x}{y}$ | $\left(\frac{x}{y} \right)^2 \left(\frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2} \right)$ |
| \sqrt{x} | $\frac{1}{4} \cdot \frac{\sigma_x^2}{x}$ |
| $\ln x$ | $\frac{\sigma_x^2}{x^2}$ |

A.3.2 Aplicaciones al recuento de muestras radiactivas

A continuación se presentan varios ejemplos para ilustrar la aplicación de algunos de los conceptos anteriormente introducidos, en el análisis de los datos obtenidos en el recuento de muestras radiactivas.

1. Recuento en presencia de fondo

Una aplicación común es el recuento de la actividad de una muestra en presencia de fondo. Sea N el número de cuentas *netas* de la muestra, G las cuentas totales (muestra + fondo) y F el número de cuentas de fondo, obtenidas durante el mismo tiempo de medida. Así

$$N = G - F .$$

Por la fórmula de propagación de errores se tendrá que

$$\sigma_N^2 = \sigma_G^2 + \sigma_F^2 .$$

Dado que G y F son números de cuentas medidas directamente, que siguen la distribución de Poisson, la desviación estándar estimada de cada uno de ellos será su propia raíz cuadrada

$$\sigma_G = \sqrt{G} = \sqrt{G} , \quad \sigma_F = \sqrt{F} = \sqrt{F} .$$

Por tanto

$$\sigma_N = \sqrt{G+F} .$$

Ejemplo

Se supone que se han obtenido de las medidas del número de cuentas totales y del fondo durante tiempos iguales de recuento los resultados siguientes: cuentas totales $G = 1071$, fondo $F = 521$. ¿Cuáles son las cuentas netas y su incertidumbre?.

De los datos se obtiene que las cuentas netas son $N = 550$, por tanto

$$\sigma_N = \sqrt{\sigma_G^2 + \sigma_F^2} = \sqrt{G+F} = \sqrt{1592} = 39,9 .$$

El resultado se expresa

$$\text{cuentas netas } N = 550 \pm 39,9 .$$

2. Determinación de la tasa de recuento

Si x son las cuentas obtenidas durante un tiempo de medida t , entonces la tasa de recuento r es

$$r = \frac{x}{t} .$$

Se acostumbra a aceptar que el tiempo es medido con muy poca incertidumbre y por consiguiente, t se considera una constante. Se trata entonces de una función que depende de una sola variable aleatoria

$$r = F(x) = \frac{1}{t} \cdot x .$$

Por la fórmula de propagación de errores se tendrá

$$\sigma_r^2 = \left(\frac{1}{t}\right)^2 \cdot \sigma_x^2 .$$

Ejemplo

Se supone que se ha obtenido en una medida efectuada durante $t = 10$ s un número de cuentas $x = 2250$. Determinar la tasa de contaje y su incertidumbre.

La tasa de contaje es

$$r = \frac{2250}{10} = 225 \text{ s}^{-1} .$$

La desviación estándar asociada será

$$\sigma_r = \frac{\sigma_x}{t} = \frac{\sqrt{x}}{t} = \frac{\sqrt{x}}{t} = \frac{\sqrt{2250}}{10} = 4,7 \text{ s}^{-1} \approx 5 \text{ s}^{-1} .$$

Por tanto, la tasa de contaje con indicación de la incertidumbre de la medida es $r = 225 \pm 5$ cuentas por segundo.

3. Tasa de recuento en presencia de fondo

Este caso corresponde a la superposición de los dos anteriores. Las cuentas brutas G se miden durante un tiempo t_G , el fondo F se mide durante un tiempo t_F . La tasa de recuento neta será

$$R_N = R_G - R_F = \frac{G}{t_G} - \frac{F}{t_F} .$$

Si se acepta que el tiempo se ha medido con poca incertidumbre, puede considerarse como una constante, resultando la tasa de recuento neta una función de dos variables $R_N = f(G, F)$. Aplicando la fórmula de propagación de errores

$$\sigma_{R_N}^2 = \left(\frac{1}{t_G}\right)^2 \sigma_G^2 + \left(\frac{1}{t_F}\right)^2 \sigma_F^2 .$$

Ejemplo

Se supone que durante un tiempo $t_G = 20$ min se ha observado una muestra radiactiva y se ha obtenido $G = 40$ cuentas. Después durante un tiempo $t_F = 1000$ min se ha realizado una medida del fondo obteniéndose $F = 1000$ cuentas. ¿Cuál es la tasa de recuento neta y su incertidumbre?

La tasa de recuento neta será

$$R_N = \frac{40}{20} - \frac{1000}{1000} = 1 \text{ cpm} .$$

La desviación estándar asociada será

$$\sigma_{R_N} = \sqrt{\left(\frac{1}{t_G}\right)^2 \sigma_G^2 + \left(\frac{1}{t_F}\right)^2 \sigma_F^2} = \sqrt{\frac{G}{t_G^2} + \frac{F}{t_F^2}} , \quad \sigma_{R_N} = \sqrt{\frac{40}{20^2} + \frac{1000}{1000^2}} = 0,32 \text{ cpm} .$$

El resultado se expresa

$$R_N = (1 \pm 0,32) \text{ cpm} .$$

A.3.3 Valor medio de recuentos múltiples

Supóngase que se han realizado n medidas repetidas de la misma muestra radiactiva durante intervalos de tiempo iguales, obteniéndose valores x_1, x_2, \dots, x_n . Si se indica la suma de estos valores por Σ se tiene

$$\Sigma = x_1 + x_2 + \dots + x_n .$$

Si se aplica la fórmula de propagación de errores se obtiene

$$\sigma_{\Sigma}^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \dots + \sigma_{x_n}^2 .$$

Pero dado que, para cada recuento independiente $\sigma_{x_i} = \sqrt{x_i}$ se tendrá

$$\sigma_{\Sigma}^2 = x_1 + x_2 + \dots + x_n = \Sigma ,$$

por tanto

$$\sigma_{\Sigma} = \sqrt{\Sigma} .$$

Este resultado prueba que la desviación estándar de la suma es la misma que se habría obtenido en un sólo recuento extendido a todo el período de tiempo de todos los recuentos.

Si ahora se calcula el valor medio de esas n medidas independientes

$$\bar{x} = \frac{\Sigma}{n},$$

teniendo en cuenta que n es una constante puede aplicarse de nuevo la ley de propagación de errores y se obtiene

$$\sigma_{\bar{x}} = \frac{\sigma_{\Sigma}}{n} = \frac{\sqrt{\Sigma}}{n} = \frac{\sqrt{n \cdot \bar{x}}}{n},$$

$$\sigma_{\bar{x}} = \sqrt{\frac{\bar{x}}{n}}.$$

Dado que una observación x_i no diferirá mucho de \bar{x} y que $\sigma_{x_i} = \sqrt{x_i}$ se puede concluir que

$$\sigma_{\bar{x}} = \frac{\sigma_{x_i}}{\sqrt{n}},$$

por tanto el error disminuye con el número de medidas en un factor \sqrt{n} .

Una conclusión general es que, si se desea disminuir el error de una medida dada en un factor de dos, se debe aumentar cuatro veces el tiempo de recuento.

A.3.4 Límites de detección

Cuando se trata de evaluar niveles bajos de una magnitud (actividad, dosis,...), frecuentemente se plantea la siguiente pregunta:

¿Cuál es el mínimo nivel de actividad que el sistema de medida es capaz de detectar?

En función de la respuesta a esta pregunta puede ser necesario buscar otro sistema de medida que sea capaz de detectar los niveles requeridos para una aplicación concreta.

Se define el *límite inferior de detección LID* como la menor actividad de una muestra que dará lugar a un recuento neto para el que se detectará la presencia de actividad con un nivel de confianza predeterminado.

A continuación, se va a obtener una expresión que permite determinar el *LID*, y se van a presentar algunos ejemplos para tratar de clarificar este concepto. Asimismo, se van a introducir los conceptos de error de primera y segunda especie, dado que existen dos posibilidades de adoptar una decisión errónea:

Error de primera especie α : Consiste en suponer que hay señal cuando de hecho no la hay.

Error de segunda especie β : Se trata de suponer que no hay señal cuando ésta existe realmente.

Para analizar una muestra se efectúan dos recuentos, en el primero se determina el fondo F , y en el segundo las cuentas brutas G ; por tanto las cuentas netas serán

$$N = G - F.$$

Se supone que la muestra tiene actividad nula, por tanto, que N tiene de valor medio cero y se distribuye normalmente alrededor de este valor.

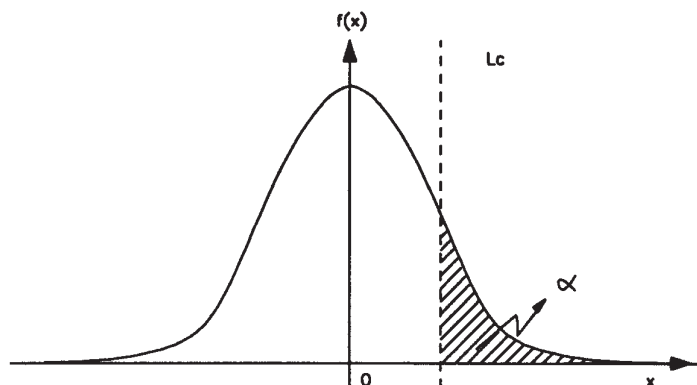


Fig. A.11 Distribución normal de la tasa de recuento, límite crítico y error de primera especie

Si ahora se mide otra muestra, y se obtiene como cuentas netas $N' > 0$, debe preguntarse si estas cuentas son consecuencia de que la muestra tiene actividad, o bien que la muestra no tiene actividad pero el N' observado es consecuencia de la fluctuación estadística de la medida. Por ello se define el *límite crítico L_c* asociado a un error de primera especie α (Fig. A.11), como

$$L_c = K_\alpha \cdot \sigma_0$$

para la curva normal tipificada,

donde K_α es un factor que depende del nivel de confianza (según tabla A.7); σ_0 es la desviación estándar de la muestra sin actividad

Si $N' > L_c$ se dice que la muestra es activa con un nivel de confianza $(1-\alpha)\%$.

Si $N' < L_c$ se dice que la muestra no es activa con un nivel de confianza $(1-\alpha)\%$.

L_c se denomina *límite crítico a posteriori*.

Ejemplo

Para un sistema de medida se ha obtenido una desviación estándar de la muestra sin actividad de $\sigma_0 = 9$. Se efectúa la medida de una muestra y se obtienen $N = 16$ cuentas netas. Si se adopta como criterio un nivel de confianza del 95%, ¿qué puede decirse sobre la actividad de dicha muestra?

Un nivel de confianza del 95% corresponde a

$$1 - \alpha = 95\% \Rightarrow K_\alpha = 1,645.$$

El límite crítico será

$$L_c = 1,645 \cdot 9 = 14,8.$$

Como $N > L_c$, se acepta que la muestra tiene actividad con un nivel de confianza del 95%.

Hay, pues, una probabilidad del 5% de que no exista actividad (error de primera especie).

Hasta aquí no se ha hecho referencia al error de segunda especie. Supóngase ahora que se pregunta: ¿cuál tiene que ser la actividad mínima de una muestra (de su valor medio), para que al efectuarse una medida se tenga una probabilidad β de cometer un error de segunda especie, es decir, de que la medida sea menor que L_c y por tanto considerar que no hay actividad cuando en realidad sí la hay? Esta actividad mínima se llama *límite inferior de detección (LID)*.

En la figura A.12 quedan reflejadas estas condiciones.

Analíticamente

$$LID = L_c + K_\beta \cdot \sigma_D = K_\alpha \cdot \sigma_0 + K_\beta \cdot \sigma_D,$$

donde

K_β depende del nivel de confianza (según tabla A.7);

σ_D desviación estándar de las cuentas netas de la muestra con actividad.

En algunas ocasiones puede simplificarse la expresión anterior:

- Se supone $\sigma_0 = \sigma_D \Rightarrow LID = (K_\alpha + K_\beta) \sigma_D$.

- Si además se adoptan valores iguales para los errores de primera y segunda especie

$$K_\alpha = K_\beta = K \implies LID = 2 \cdot K \cdot \sigma_D .$$

- Si la actividad bruta y la de fondo son muy próximas

$$\sigma_D = \sqrt{\sigma_G^2 + \sigma_F^2} = \sigma_F \cdot \sqrt{2}$$

y ya que $\sigma_F = \sqrt{F}$, se obtiene

$$LID = 2 \cdot K \cdot \sqrt{2 \cdot F} .$$

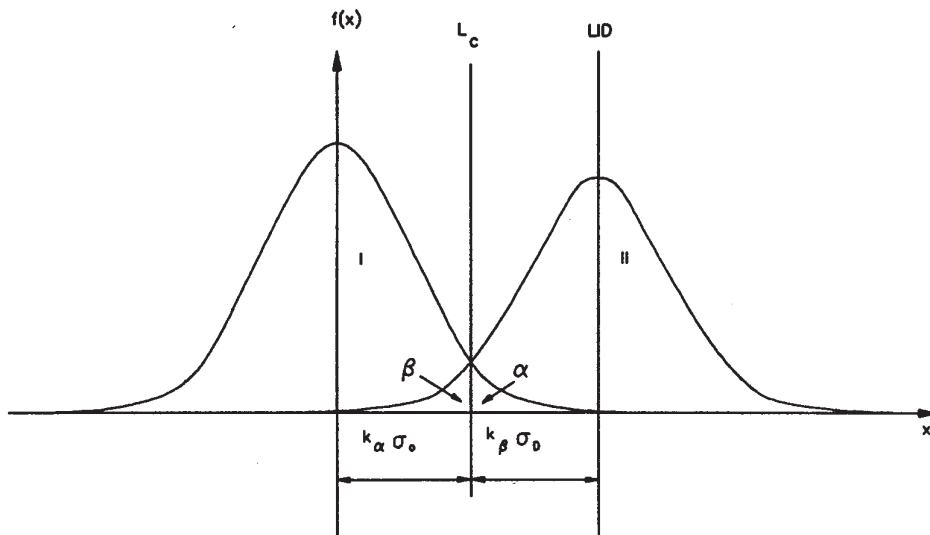


Fig. A.12 Distribuciones de Gauss de las tasas de recuento de muestra y fondo, representando los valores de K_α , K_β , L_c y LID

Tabla A.7 Valores de K para algunos niveles de confianza

| α | $1 - \alpha$ | K |
|----------|--------------|-------|
| 0,01 | 0,99 | 2,327 |
| 0,05 | 0,95 | 1,645 |
| 0,10 | 0,90 | 1,282 |
| 0,50 | 0,50 | 0,0 |

Es usual adoptar un nivel de confianza del 95 % para los dos tipos de errores, con lo que resulta el límite inferior de detección

$$LID = 4,65\sqrt{F} .$$

Ejemplo

Se dispone de un detector con un fondo de 0,2 cpm. El tiempo de medida para el fondo y para las muestras es siempre de 400 min. Se tendrá

$$F = 0,2 \cdot 400 = 80 \text{ cuentas} , \quad \sigma_F = \sqrt{80} = 8,94 .$$

Si se toman los errores de primera y segunda especie iguales a 5 %, se tendrá

$$LID = 4,65 \cdot 8,94 = 42 \text{ cuentas} \quad (\text{ó } \frac{42}{400} = 0,1 \text{ cpm}) .$$

Supóngase que se realiza ahora la medida de dos muestras, y se obtiene los valores

$$G_1 = 90 \text{ cuentas} , \quad G_2 = 140 \text{ cuentas} .$$

En el primer caso

$$N_1 = G_1 - F = 90 - 80 = 10 < LID ,$$

el resultado se expresará

$$R_1 < LID \quad (\text{con } LID = 0,1 \text{ cpm}) .$$

En el segundo caso

$$N_2 = G_2 - F = 140 - 80 = 60 \text{ cuentas} > LID, \quad \sigma_{N_2} = \sqrt{140 + 80} = 14,8 \text{ cuentas}$$

el resultado se expresará

$$R_2 = \frac{140}{400} \pm \frac{14,8}{400} = (0,35 \pm 0,04) \text{ cpm} .$$

B Unidades

Tabla B.1 Unidades fundamentales del sistema internacional

| Magnitud | Nombre | Símbolo |
|-------------------------|-----------|---------|
| Longitud | metro | m |
| Masa | kilogramo | kg |
| Tiempo | segundo | s |
| Intensidad de corriente | amperio | A |
| Temperatura | kelvin | K |
| Intensidad luminosa | candela | cd |
| Cantidad de materia | mol | mol |

Tabla B.2 Unidades derivadas del sistema internacional

| Magnitud | Nombre | | Símbolo | Dimensiones | |
|------------------------------|------------|---------------|----------|--|-------------------|
| | Español | Internacional | | Básica | Usual |
| Fuerza | neutonio | newton | N | $m \cdot kg \cdot s^{-1}$ | -- |
| Energía | julio | joule | J | $m^2 \cdot kg \cdot s^{-2}$ | N · m |
| Potencia | vatio | watt | W | $m^2 \cdot kg \cdot s^{-3}$ | J/s |
| Presión | pascal | pascal | Pa | $m^{-1} \cdot kg \cdot s^{-2}$ | N/m ² |
| Frecuencia | hercio | hertz | Hz | s^{-1} | -- |
| Cantidad de electricidad | culombio | coulomb | C | $s \cdot A$ | A · s |
| Tensión eléctrica | voltio | volt | V | $m^2 \cdot kg \cdot s^{-3} \cdot A^{-1}$ | W/A |
| Resistencia eléctrica | ohmio | ohm | Ω | $m^2 \cdot kg \cdot s^{-3} \cdot A^{-2}$ | V/A |
| Capacidad eléctrica | faradio | farad | F | $m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$ | C/V |
| Inducción magnética | tesla | tesla | T | $kg \cdot s^{-2} \cdot A^{-1}$ | Wb/m ² |
| Inductancia | henrio | henry | H | $m^2 \cdot kg \cdot s^{-2} \cdot A^{-2}$ | Wb/A |
| Flujo de inducción magnética | weberio | weber | Wb | $m^2 \cdot kg \cdot s^{-2} \cdot A^{-1}$ | V · s |
| Actividad | bequerelio | becquerel | Bq | s^{-1} | -- |
| Dosis absorbida | gray | gray | Gy | $m^2 \cdot s^{-2}$ | J/kg |
| Dosis equivalente | sievert | sievert | Sv | $m^2 \cdot s^{-2}$ | J/kg |