Capítulo 4 (Continuación)

MÉTODOS ESTADÍSTICOS

Autor: José María García Palanco

4.8 Métodos Estadísticos

Ya hemos visto en los apartados anteriores, que un procedimiento aproximado para la estimación del error se basa en el establecimiento de una cota superior para el error. Es un procedimiento simple de aplicar, pero corremos el riesgo de estimar el error cometido por exceso.

Por fortuna, existe otro procedimiento alternativo para estimar los errores de un experimento de medida. Esta segunda posibilidad está basada en el uso de métodos estadísticos. Resulta por tanto más complicado y laborioso que el basado en las cotas de error y será de aplicación más restringida. Más aún si tenemos en cuenta que la condición previa para poder realizar este tipo de estimación del error es disponer un conjunto de medidas repetidas sobre el mismo sistema.

La Estadística es una ciencia amplia y compleja, cuyo estudio, incluso introductorio, va más allá del alcance de este curso. A continuación simplemente presentaremos, sin ningún tipo de justificación teórica, algunos de los procedimientos necesarios para calcular un error de carácter estadístico.

4.9 Probabilidad de un Suceso al Azar

El punto de partida es asumir que los inevitables errores de medida son aleatorios (se producen al azar), es decir que unas veces serán por exceso y otras por defecto con respecto al valor exacto, que lógicamente resulta ser desconocido.

Los sucesos aleatorios ya fueron estudiados por Gauss, resultando posible prever la probabilidad con la que puede darse uno de estos sucesos. Para ver más claramente este punto conviene referirnos a la llamada *función de distribución normal* o *gaussiana*, cuya expresión matemática y su representación gráfica son las que aparecen a continuación:

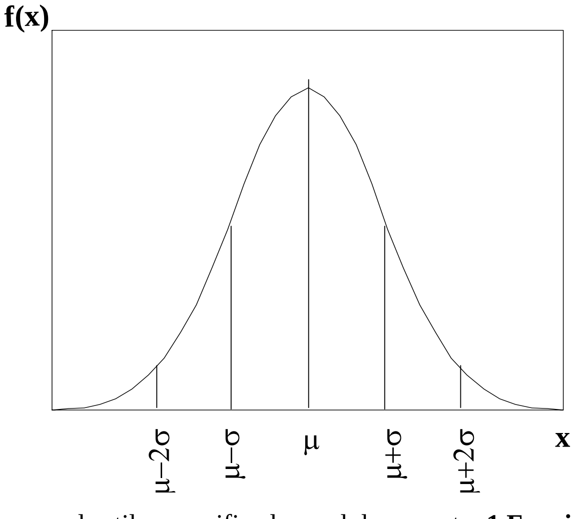

Fig. ¡Error! No hay texto con el estilo especificado en el documento.**.1 Función de distribución gaussiana**

[¡Error! No hay texto con el estilo especificado en el documento.**.1]** $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma}}$$

Observamos cómo esta función se define a partir de dos parámetros, que posteriormente cobrarán especial importancia: la *media* o *promedio* (μ) y la *desviación estándar* (σ).

A partir de esta función de distribución, puede demostrarse que resulta posible calcular la probabilidad $P(x)$, de que un suceso al azar x, tome un valor comprendido dentro de un cierto intervalo. En concreto, podríamos destacar los siguientes valores significativos:

[¡Error! No hay texto con el estilo especificado en el documento.**.2]**

$$P(\mu-\sigma \le x \le \mu+\sigma) = 0,683 \quad \rightarrow \quad 68,3\%$$
$$P(\mu-2\sigma \le x \le \mu+2\sigma) = 0,954 \quad \rightarrow \quad 95,4\%$$
$$P(\mu-3\sigma \le x \le \mu+3\sigma) = 0,997 \quad \rightarrow \quad 99,7\%$$

4.10 Cálculo de Errores Estadísticos

Estos valores de probabilidad son los que nos van a permitir estimar el error cometido en una serie de medidas. En concreto, bastará con calcular la *media* y la *desviación estándar* de nuestro conjunto de datos, según:

[¡Error! No hay texto con el estilo especificado en el documento.**.3]**

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$
$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}$$

siendo n el número total de datos y x_i cada uno de los valores experimentales disponibles.

Supongamos, por ejemplo, que hemos realizado un estudio para determinar el volumen envasado en botellines de cerveza, midiéndose un total de 52 muestras. Si la media obtenida es de 19,7 cm^3 y la desviación estándar de 0,3 cm^3, podemos concluir que si tomamos uno de los botellines al azar, su volumen debe ser algún valor comprendido dentro del intervalo $\mu\pm2\sigma$ con una probabilidad del 95,4%.

Volumen $\in$ [19,7 – 0,6 , 19,7 + 0,6] cm^3

Dicho de otra forma, podemos asegurar casi con certeza que los botellines contienen 19,7 cm^3, con un error absoluto de ±0,6 cm^3.

Nos queda aún por señalar una última cuestión de carácter práctico. En realidad, el cálculo del error basado en la función de distribución gaussiana, solo es válido para muestras suficientemente grandes, del orden de al menos 30 elementos. La mayoría de los ensayos realizados en un laboratorio no suelen ser tan amplios, por lo que el uso de este método está imposibilitado.

Afortunadamente, también puede demostrarse que para muestras más pequeñas se cumple otro tipo de distribución estadística, conocida como *t de Student*. El procedimiento de trabajo en este segundo caso, resulta ser muy semejante al anterior. En concreto, comenzaremos por definir la *media* μ y la *desviación estándar de la muestra s*, según:

[¡Error! No hay texto con el estilo especificado en el documento.**.4]**
$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$
$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}$$

Definimos también los *grados de libertad* (ν) del sistema, restando la unidad al número de medidas disponibles:

[¡Error! No hay texto con el estilo especificado en el documento.**.5]** $\nu = n - 1$

En función de los grados de libertad y del grado de confianza (probabilidad de acertar) que queramos en nuestros resultados (usualmente el 95% o el 99%), podemos determinar el llamado *límite de confianza* (*LC*):

[¡Error! No hay texto con el estilo especificado en el documento.**.6]** $LC = \overline{x} \pm t\,\frac{s}{\sqrt{n}}$

El parámetro *t* tiene que ser determinado a partir de los datos disponibles en la Tabla 4.1 que aparece en la página siguiente.

Tabla ¡Error! No hay texto con el estilo especificado en el documento.**.1 Niveles de confianza para una distribución *t* de Student**

Grados de Libertad	**Grado de Confianza**				**Grados de Libertad**	**Grado de Confianza**			
	90 %	**95 %**	**99 %**	**99,9 %**		**90 %**	**95 %**	**99 %**	**99,9 %**
1	6,314	12,71	63,66	636,6	16	1,746	2,120	2,921	4,015
2	2,920	4,303	9,925	31,60	17	1,740	2,110	2,898	3,965
3	2,353	3,182	5,841	12,94	18	1,734	2,101	2,878	3,922
4	2,132	2,776	4,604	8,610	19	1,729	2,093	2,861	3,883
5	2,015	2,571	4,032	6,859	20	1,725	2,086	2,845	3,850
6	1,943	2,447	3,707	5,959	21	1,721	2,080	2,831	3,819
7	1,895	2,365	3,499	5,405	22	1,717	2,074	2,819	3,792
8	1,860	2,306	3,355	5,041	23	1,714	2,069	2,807	3,767
9	1,833	2,262	3,250	4,781	24	1,711	2,064	2,797	3,745
10	1,812	2,228	3,169	4,537	25	1,708	2,060	2,787	3,725
11	1,796	2,201	3,106	4,437	26	1,706	2,056	2,779	3,707
12	1,782	2,179	3,055	4,318	27	1,703	2,052	2,771	3,690
13	1,771	2,160	3,012	4,221	28	1,701	2,048	2,763	3,674
14	1,761	2,145	2,977	4,140	29	1,699	2,045	2,756	3,659
15	1,753	2,131	2,947	4,073	30	1,697	2,042	2,750	3,646
					40	1,645	1,960	2,576	3,291