
Introducción al Análisis de Errores

El estudio de las incertidumbres
en las mediciones físicas

Traducción de la segunda edición original

JOHN R. TAYLOR

Profesor de Física
UNIVERSIDAD DE COLORADO



EDITORIAL
REVERTÉ

Barcelona · Bogotá · Buenos Aires · Caracas · México

Registro bibliográfico (ISBD)

Taylor, John R.

[An Introduction to Error Analysis : The Study of Uncertainties in Physical Measurements. Español]
Introducción al análisis de errores : el estudio de las incertidumbres en las mediciones físicas / John R Taylor ;
versión española traducida por: José María Oller Sala. – Barcelona : Reverté, 2014.

XVII , 342 p. : il. ; 25 cm.

Traducción de : Introduction to Error Analysis. Second Edition.

Bibliog.: p. 309. – Índice.

DL B 12029-2014. – ISBN 978-84-291-5184-8

I. Análisis de errores. I. Oller Sala, José María, trad. II. Título.
519.2

Título de la obra original:

An Introduction to Error Analysis. The Study of Uncertainties in Physical Measurements. Second Edition.

Edición original en lengua inglesa publicada por

University Science Books. Sausalito, California.

Copyright © 1997 by University Science Books. All Rights Reserved

Edición en español:

© Editorial Reverté, S. A., 2014

ISBN: 978-84-291-5184-8

Versión española traducida por:

José María Oller Sala

Catedrático de Estadística de la Universidad de Barcelona

DISEÑO DE LA CUBIERTA: David Kimura + Gabriela Varela

MAQUETACIÓN: Reverté-Aguilar, S. L.

Propiedad de:

EDITORIAL REVERTÉ, S. A.

Loreto, 13-15. Local B

Tel: (34) 93 419 33 36

Fax: (34) 93 419 51 89

08029 Barcelona. España

reverte@reverte.com

www.reverte.com

Reservados todos los derechos. La reproducción total o parcial de esta obra, por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos, queda rigurosamente prohibida sin la autorización escrita de los titulares del copyright, bajo las sanciones establecidas por las leyes.

Impreso en España - *Printed in Spain*

ISBN: 978-84-291-5184-8

DL B 12029-2014

Impresión y encuadernación: Liberdúplex, S. L. U..

1415

Contenidos

Parte I

Capítulo 1 Descripción preliminar del análisis de errores 3

- 1.1 Errores como incertidumbres 3
- 1.2 La inevitabilidad de la incertidumbre 3
- 1.3 La importancia de conocer las incertidumbres 5
- 1.4 Más ejemplos 7
- 1.5 Estimación de incertidumbres al leer escalas 8
- 1.6 Estimar las incertidumbres en mediciones repetidas 10

Capítulo 2 Cómo expresar y utilizar las incertidumbres 13

- 2.1 Mejor estimación \pm incertidumbre 13
- 2.2 Cifras significativas 15
- 2.3 Discrepancia 17
- 2.4 Comparación entre valores medidos y aceptados 19
- 2.5 Comparación entre dos valores medidos 21
- 2.6 Comprobación gráfica de relaciones entre variables 25
- 2.7 Incertidumbre relativa 30
- 2.8 Cifras significativas e incertidumbres relativas 31
- 2.9 Producto de dos valores medidos 33
- Principales definiciones y ecuaciones del capítulo 2 36
- Problemas del capítulo 2 37

Capítulo 3 Propagación de incertidumbres 47

- 3.1 Incertidumbre de las mediciones directas 48
- 3.2 La regla de la raíz cuadrada para experimentos de recuento 50

3.3	Sumas y restas; productos y cocientes	52
3.4	Dos importantes casos particulares	56
3.5	Incertidumbres independientes en una suma	60
3.6	Más sobre incertidumbres independientes	62
3.7	Funciones arbitrarias de una variable	65
3.8	Propagación de errores paso a paso	69
3.9	Ejemplos	71
3.10	Un ejemplo más complicado	74
3.11	Fórmula general para la propagación de errores	76
	Principales definiciones y ecuaciones del capítulo 3	80
	Problemas del capítulo 3	82
Capítulo 4	Análisis estadístico de incertidumbres aleatorias	97
4.1	Errores aleatorios y sistemáticos	98
4.2	La media y la desviación estándar	102
4.3	Desviación estándar como incertidumbre de una medición individual	106
4.4	La desviación estándar de la media	107
4.5	Ejemplos	109
4.6	Errores sistemáticos	111
	Principales definiciones y ecuaciones del capítulo 4	115
	Problemas del capítulo 4	116
Capítulo 5	La distribución normal	127
5.1	Histogramas y distribuciones	128
5.2	Distribuciones límite	132
5.3	La distribución normal o de Gauss	136
5.4	La desviación estándar como límite de confianza del 68%	141
5.5	Justificación de la media como mejor estimación	144
5.6	Justificación de la suma cuadrática	148
5.7	La desviación estándar de la media	154
5.8	Aceptabilidad de una cantidad estimada	156
	Principales definiciones y ecuaciones del capítulo 5	159
	Problemas del capítulo 5	161

Parte II

Capítulo 6 Exclusión de datos 173

- 6.1 El problema de excluir datos 173
- 6.2 El criterio de Chauvenet 174
- 6.3 Discusión 177
- Principales definiciones y ecuaciones del capítulo 6 178
- Problemas del capítulo 6 178

Capítulo 7 Medias ponderadas 181

- 7.1 El problema de combinar mediciones diferentes 181
- 7.2 La media ponderada 182
- 7.3 Un ejemplo 184
- Principales definiciones y ecuaciones del capítulo 7 185
- Problemas del capítulo 7 186

Capítulo 8 Ajuste mínimo cuadrático 189

- 8.1 Datos que deberían ajustarse a una línea recta 189
- 8.2 Cálculo de las constantes A y B 190
- 8.3 Incertidumbre en la medición de y 194
- 8.4 La incertidumbre de las constantes A y B 196
- 8.5 Un ejemplo 198
- 8.6 Ajuste mínimo cuadrático a otras curvas 201
- Principales definiciones y ecuaciones del capítulo 8 206
- Problemas del capítulo 8 208

Capítulo 9 Covarianza y correlación 217

- 9.1 Revisión de la propagación de errores 217
- 9.2 La covarianza en la propagación de errores 219
- 9.3 Coeficiente de correlación lineal 223
- 9.4 Análisis cuantitativo del nivel de significación de r 227
- 9.5 Ejemplos 229
- Principales definiciones y ecuaciones del capítulo 9 230
- Problemas del capítulo 9 231

Capítulo 10 La distribución binomial 237

- 10.1 Distribuciones 237
- 10.2 Probabilidades en un lanzamiento de dados 238
- 10.3 Definición de la distribución binomial 239
- 10.4 Propiedades de la distribución binomial 242
- 10.5 La distribución de Gauss de errores aleatorios 245
- 10.6 Aplicaciones; contraste de hipótesis 246
- Principales definiciones y ecuaciones del capítulo 10 251

Capítulo 11 La distribución de Poisson 255

- 11.1 Definición de la distribución de Poisson 255
- 11.2 Propiedades de la distribución de Poisson 259
- 11.3 Aplicaciones 262
- 11.4 Eliminación de ruidos 264
- Principales definiciones y ecuaciones del capítulo 11 266
- Problemas del capítulo 11 267
- Problemas del capítulo 10 252

Capítulo 12 El test ji-cuadrado para una distribución 271

- 12.1 Introducción a la ji cuadrado 271
- 12.2 Definición general de la ji cuadrado 275
- 12.3 Grados de libertad y ji cuadrado reducida 279
- 12.4 Probabilidades para la distribución ji cuadrado 281
- 12.5 Ejemplos 284
- Principales definiciones y ecuaciones del capítulo 12 288
- Problemas del capítulo 12 289

Apéndices 295

- Apéndice A Integral de error normal, I 296
- Apéndice B Integral de error normal, II 298
- Apéndice C Probabilidades para los coeficientes de correlación 300
- Apéndice D Probabilidades para la distribución ji cuadrado 302
- Apéndice E Dos demostraciones relativas a las desviaciones estándar muestrales 304
- Bibliografía 309

Respuestas a los tests rápidos y a los problemas de numeración
impar 311

Índice de materias 333

Prefacio a la segunda edición

Escribí por primera vez *Introducción al análisis de errores* porque mi experiencia docente en clases introductorias de laboratorio durante varios años me convenció de la gran necesidad de un libro que realmente *introdujera* el tema a estudiantes de ciencias universitarios. Varios buenos libros sobre el tema estaban ya disponibles, pero ninguno era realmente adecuado para un estudiante que se iniciase en la materia. La buena acogida de la primera edición confirmó la existencia de dicha necesidad y sugiere que el libro la satisfizo.

El continuo éxito de la primera edición sugiere que todavía responde a esta necesidad. Sin embargo, después de más de una década, todo autor de un libro de texto universitario debe, sin duda, sentirse obligado a mejorar y actualizar la versión original. Las ideas para la modificación provienen de distintas fuentes: sugerencias de los lectores, la necesidad de adaptar el libro a la amplia disponibilidad de calculadoras y ordenadores, así como a mis propias experiencias en la enseñanza del libro y el hallazgo de aspectos mejorables.

Debido a la reacción abrumadoramente favorable a la primera edición, he mantenido su nivel básico y el enfoque general. Por lo tanto, muchas revisiones son simplemente cambios en la redacción para mejorar la claridad. Unos pocos cambios son importantes, los más importantes de los cuales son los siguientes:

1. El número de problemas al final de cada capítulo casi se duplica para ampliar la oferta a los usuarios y para que los profesores tengan la posibilidad de variar sus problemas asignados de año en año. Huelga decir que no es necesario que el lector resuelva los cerca de 264 problemas que se ofrecen; al contrario, una media docena de problemas por capítulo es probablemente suficiente.
2. Varios lectores me recomendaron la colocación de unos cuantos ejercicios sencillos repartidos regularmente a lo largo del texto para que comprueben que realmente entienden las ideas que se le acaban de presentar. Estos ejercicios aparecen ahora como “Test rápido”, e insto encarecidamente a los estudiantes que se introducen en la materia que traten de resolverlos todos. Si algún “Test rápido” le requiere al lector mucho más de uno o dos minutos, probablemente éste tenga que releer los párrafos anteriores. Las respuestas de todos los “Test rápidos” se proporcionan en la sección de respuestas al final del libro. Quienes encuentren este tipo de ejercicio como una distracción, pueden simplemente saltárselos.

3. Otra de las novedades de esta edición son los resúmenes completos de todas las ecuaciones importantes al final de cada capítulo que complementan las breves sinopsis de la primera edición situadas detrás de la portada y la contraportada. En estos nuevos resúmenes se enumeran todas las ecuaciones fundamentales del capítulo y también de la colección de problemas.

4. En esta edición hay muchos gráficos nuevos, sobre todo en los primeros capítulos. Los gráficos ayudan a hacer que el texto parezca menos intimidatorio y reflejan mi esfuerzo consciente para estimular a los estudiantes que a piensen más visualmente sobre las incertidumbres. He observado, por ejemplo, que muchos estudiantes les ayuda a entender cuestiones tales como la coherencia de las mediciones si piensan visualmente en términos de barras de error.

5. He reorganizado las colecciones de problemas al final de cada capítulo de tres maneras. En primer lugar, la sección de Soluciones situada al final del libro contiene las respuestas de todos los problemas impares. (La primera edición sólo proporcionaba la solución a problemas seleccionados). El nuevo sistema es más simple y más tradicional. En segundo lugar, como una guía aproximada para indicar su nivel de dificultad, he etiquetado los problemas con un sistema de estrellas: una estrella (★) indica un ejercicio simple cuya resolución no debería requerir más de un par de minutos si se entiende el tema; dos estrellas (★★) indican un problema algo más difícil; y tres (★★★) indican un problema realmente más complejo cuya resolución requiere la utilización de varios conceptos diferentes, así como más tiempo. Admito que la clasificación es muy aproximada, pero los alumnos que estudian por su cuenta encontrarán que le es útil, así como los maestros a la hora de asignar problemas a sus alumnos.

En tercer lugar, he ordenado los problemas por el número de sección. Tan pronto como el lector haya leído la sección N , deberá estar preparado para intentar resolver cualquier problema listado en esa sección. Aunque este sistema es conveniente tanto para el estudiante como el profesor, parece ser que en la actualidad está en desuso. Supongo que esta desaprobación está basada en el argumento de que dicho sistema podría excluir a los problemas más complejos que implican muchas ideas de diferentes secciones. Considero que éste es un argumento engañoso: un problema listado en la sección N puede, por supuesto, involucrar ideas de muchas secciones previas y, por consiguiente, ser tan general y profundo como cualquier problema enumerado bajo un título más general.

6. He añadido los problemas que requieren el uso de programas informáticos tales como hojas de cálculo de Lotus 123 o Excel. Ninguno de estos problemas es específico de un determinado sistema; por el contrario, instan al estudiante a que aprenda a hacer distintas tareas utilizando cualquier sistema disponible. Del mismo modo, varios problemas los animan a aprender a usar las funciones incorporadas en sus calculadoras para calcular las desviaciones estándar y similares.

7. He añadido un apéndice (Apéndice E) para mostrar dos demostraciones que se refieren a las desviaciones estándar de la muestra: la primera, que, basándose en N mediciones de una magnitud, la mejor estimación de la anchura real de su distribución es la desviación estándar de la muestra con $(N - 1)$ en el denomi-

nador, y la segunda, que la incertidumbre en esta estimación viene dada por la ecuación (5.46). Estas demostraciones son sorprendentemente difíciles y no son fáciles de encontrar en la bibliografía.

Es un placer dar las gracias a las numerosas personas que me han hecho sugerencias para esta segunda edición. Entre mis amigos y colegas de la Universidad de Colorado, los que más generosamente me dieron su tiempo y sus conocimientos fueron David Alexander, Dana Anderson, David Bartlett, Barry Bruce, John Cumalat, Mike Dubson, Bill Ford, Mark Johnson, Jerry Leigh, Uriel Nauenberg, Bill O'Sullivan, Bob Ristinen, Rod Smythe y Chris Zafiratos. En otras instituciones, quiero dar especialmente las gracias a R. G. Chambers de Leeds, Inglaterra, a Sharif Heger de la Universidad de Nuevo México, a Steven Hoffmaster de Gonzaga University, a Hilliard Macomber de Universidad de Northern Iowa, a Mark Semon de Bates College, a Peter Timbie de Brown University y a David Van Dyke de la Universidad de Pennsylvania. Me siento profundamente en deuda con todas estas personas por su generosa ayuda. Estoy también muy agradecido a Bruce Armbruster de la Universidad de Science Books por su generoso ánimo y apoyo. Por encima de todo, quiero dar las gracias a mi esposa Debby; yo no sé cómo aguanta las tensiones y el estrés que suponen la escritura de un libro, pero estoy muy agradecido de como lo hace.

J. R. Taylor
Septiembre de 1996
Boulder, Colorado

Prefacio a la primera edición

Todas las medidas, por muy cuidadas y científicas que sean, están sujetas a ciertas incertidumbres. El análisis de errores es el estudio y la evaluación de estas incertidumbres, y sus dos principales funciones son permitir a los científicos estimar el tamaño de sus incertidumbres, y ayudarlos a reducirlas cuando sea necesario. El análisis de las incertidumbres, o “errores”, es una parte vital de cualquier experimento científico, y por lo tanto el análisis de errores es una parte importante de cualquier curso universitario de ciencias experimentales. También puede ser una de las partes más interesantes del curso. El reto de estimar incertidumbres y reducirlas a un nivel que permita sacar una conclusión correcta puede hacer que un conjunto aburrido y rutinario de mediciones se convierta en un ejercicio realmente interesante.

Este libro es una introducción al análisis de errores para su uso en aquellos cursos universitarios de introducción a la física experimental a los que por lo general asisten estudiantes de primero o de segundo año de ciencias o ingeniería. Ciertamente no afirmo que el análisis de errores es la parte más importante (ni mucho menos la única) de este curso, pero a menudo he encontrado que es la parte más maltratada y descuidada. En muchos de esos cursos, el análisis de errores se “enseña” entregando un par de páginas con notas y unas pocas fórmulas, y se espera que el estudiante siga adelante con el trabajo solo. El resultado es que el análisis de errores se convierte en un ritual sin sentido, en el que el estudiante agrega unas pocas líneas de cálculo al final de cada informe de laboratorio, no porque entienda el porqué, sino, simplemente, porque el instructor le ha dicho que lo haga.

Escribí este libro con la convicción de que cualquier estudiante, incluso uno que nunca hubiera oído hablar de este tema, debería ser capaz de aprender qué es el análisis de errores, por qué es interesante e importante, y cómo se utilizan las herramientas básicas de la materia en los informes de laboratorio. La parte I del libro (capítulos 1 a 5) trata de hacer todo esto, con muchos ejemplos de los experimentos que se realizan en los laboratorios de enseñanza. El estudiante que domine este material debería entonces conocer y entender casi todo lo que espera aprender sobre análisis de errores en un curso de laboratorio de primer año: la propagación de errores, el uso de la estadística elemental y su justificación en términos de la distribución normal.

La segunda parte contiene una selección de temas más avanzados: ajuste de mínimos cuadrados, el coeficiente de correlación, la prueba de χ^2 y otros. Éstos

con toda seguridad no estarán incluidos oficialmente en un curso de laboratorio de primer año, aunque algunos estudiantes podrían interesarse en algunos de ellos. Sin embargo, varios de estos temas serán necesarios en un segundo curso de laboratorio, y es principalmente por esta razón que los he incluido.

Soy muy consciente de que hay muy poco tiempo para dedicar a un tema como el análisis de errores en la mayoría de los cursos de laboratorio. En la Universidad de Colorado damos una conferencia de una hora en cada una de las primeras seis semanas de nuestro curso de laboratorio de primer año. Estas conferencias, junto con unas pocas tareas con los problemas al final de los capítulos, nos ha llevado a cubrir los capítulos 1 al 4 en detalle y brevemente el capítulo 5. Esto da a los estudiantes un conocimiento práctico de la propagación de errores y de los elementos de la estadística, además de un conocimiento superficial de la teoría subyacente de la distribución normal.

De los comentarios de varios estudiantes de Colorado, se hizo evidente que las conferencias fueron un lujo innecesario para al menos algunos de los estudiantes, quienes probablemente podrían haber aprendido el material necesario de las lecturas asignadas y de los problemas propuestos. Ciertamente creo que el libro podría ser estudiado sin la ayuda de conferencias.

La parte II se podría enseñar en unas pocas charlas al inicio de un curso de laboratorio de segundo año (una vez más, complementado con algunos de los problemas asignados). Pero fue pensada, incluso más que la parte I, para ser leída por el estudiante en cualquier momento según sus propias necesidades e intereses. Sus siete capítulos son casi completamente independientes entre sí, con el fin de fomentar este tipo de uso.

Al final de cada capítulo he incluido una selección de problemas; el lector tiene que trabajar varios de ellos para dominar las técnicas. La mayoría de los cálculos de errores son bastante sencillos. Un estudiante que hace muchos cálculos complicados (ya sea en los problemas de este libro o en los informes de laboratorio) es casi seguro que está haciendo algo de una manera innecesariamente difícil. Con el fin de dar a los profesores y lectores una buena opción, he incluido muchos más problemas de los que el lector medio necesita tratar. Un lector que haga una tercera parte de los problemas lo estará haciendo bien.

En el interior de la cubierta y contracubierta figuran resúmenes de todas las fórmulas principales. Espero que estas referencias sean útiles para el lector, tanto en el estudio del libro como después. Los resúmenes están organizados por capítulos, y también espero que sirvan como breves reseñas a las que el lector pueda acudir después de estudiar cada capítulo.

En el texto, algunas declaraciones —ecuaciones y reglas de procedimiento— se han destacado con un fondo sombreado. Esta forma de destacar está reservada para las declaraciones que son importantes y están en su forma final (es decir, no se modificarán por el trabajo posterior). Definitivamente, usted tendrá que recordar estas afirmaciones, por lo que se han destacado para llamar su atención.

El nivel de matemáticas que se espera del lector sube lentamente a lo largo del libro. Los dos primeros capítulos sólo requieren álgebra; el capítulo 3 requiere diferenciación (y la diferenciación parcial en la sección 3.11, que es opcional), el

capítulo 5 necesita un conocimiento de la integración y la función exponencial. En la parte II supongo que el lector está del todo cómodo con todas estas ideas.

El libro contiene numerosos ejemplos de experimentos de física, pero no es esencial comprender la teoría subyacente. Además, los ejemplos se toman en su mayoría de la mecánica elemental y la óptica para que sea más probable que el estudiante ya haya estudiado la teoría. El lector que lo necesite puede encontrar una explicación de la teoría buscando en el índice de cualquier texto introductorio de física.

El análisis de errores es un tema que apasiona a mucha gente, y ningún tratamiento sencillo puede aspirar a complacer a todos. Mi propio prejuicio es que, cuando hay que elegir entre la facilidad de comprensión y el estricto rigor, un texto de física debe escoger la primera. Por ejemplo, en la controvertida cuestión de la combinación de errores en cuadratura en comparación con la adición directa, he elegido tratar en primer lugar la adición directa, ya que el alumno puede comprender fácilmente los argumentos que conducen a ella.

En los últimos años se ha producido un cambio drástico en los laboratorios de estudiantes con la aparición de la calculadora de bolsillo. Esto tiene algunas consecuencias desafortunadas —sobre todo, el hábito atroz de citar ridículamente cifras *no significativas* simplemente porque la calculadora las produce— pero es desde casi todos los puntos de vista una gran ventaja, especialmente en el análisis de errores. La calculadora de bolsillo le permite a uno calcular, en pocos segundos, medias y desviaciones estándar que anteriormente habrían tomado horas. Se hacen innecesarias muchas tablas, ya que ahora se pueden calcular funciones como la de Gauss con mayor rapidez que la que uno necesitaría para encontrarla en un libro de tablas. He tratado de explotar esta herramienta maravillosa en la medida de lo posible.

Es un placer agradecer a varias personas por sus valiosos comentarios y sugerencias. Una edición preliminar del libro se utilizó en varias universidades y estoy agradecido a muchos estudiantes y colegas por sus críticas. Especialmente útiles fueron los comentarios de John Morrison y David Nesbitt de la Universidad de Colorado, los profesores Pratt y Schroeder de la Michigan State, el profesor Shugart de la U. C. Berkeley y el profesor Semon del Bates College. Diane Casparian, Linda Frueh y Connie Gurule escribieron maravillosamente y a gran velocidad los sucesivos proyectos. Sin mi suegra, Frances Kretschmann, la corrección de pruebas no se hubiera hecho a tiempo. Doy las gracias a todas estas personas por su ayuda; pero sobre todo doy gracias a mi esposa, cuya laboriosa y rigurosa edición mejoró el libro entero más allá de toda medida.

J. R. Taylor
1 de Noviembre de 1981
Boulder, Colorado

Parte I

-
1. Descripción preliminar del análisis de errores
 2. Cómo expresar y utilizar las incertidumbres
 3. Propagación de incertidumbres
 4. Análisis estadístico de incertidumbres aleatorias
 5. La distribución Normal
-

La parte I presenta las ideas básicas del análisis de errores, tal como se necesitan en un laboratorio típico de primer año de Física en la universidad. Los dos primeros capítulos describen lo que es el análisis de errores, por qué es importante y cómo se puede utilizar en un informe estándar de laboratorio. El capítulo 3 describe la propagación de errores, la forma en que las incertidumbres en las mediciones originales “se propagan” a través de los cálculos realizados provocando incertidumbre en los resultados finales obtenidos. Los capítulos 4 y 5 introducen los métodos estadísticos para evaluar las llamadas incertidumbres aleatorias.

Capítulo 1

Descripción preliminar del análisis de errores

El análisis de errores es el estudio y evaluación de la incertidumbre correspondiente a cualquier medición. La experiencia ha demostrado que ninguna medición, por muy cuidadosa que sea, puede estar completamente libre de incertidumbres. Puesto que toda la estructura y aplicación de la Ciencia se basa en mediciones experimentales, la capacidad de evaluarlas y de minimizar su influencia es de crucial importancia.

En este primer capítulo se describen algunas mediciones simples que ilustran la inevitable aparición de incertidumbres experimentales y se muestra la importancia de cuantificar su magnitud. A continuación el capítulo describe cómo (por lo menos en algunos casos sencillos) la magnitud de dichas incertidumbres experimentales se puede estimar con realismo, a menudo mediante poco más que simple sentido común.

1.1 Errores como incertidumbres

En la ciencia, la palabra *error* no tiene las connotaciones habituales de las expresiones *equivocación* o *metedura de pata*. El error en la medición científica significa la inevitable incertidumbre que asiste a todas las mediciones. Como tal, los errores no son equivocaciones, no se pueden eliminar por más cuidadosos que seamos. Lo máximo que podemos aspirar es asegurarnos de que los errores sean tan pequeños como sea razonablemente posible y tener una estimación fiable de su magnitud. La mayoría de los libros de texto introducen otras definiciones de error, que se discutirán más adelante. Por ahora, el error se utiliza exclusivamente en el sentido de la incertidumbre, y las dos palabras se usan indistintamente.

1.2 La inevitabilidad de la incertidumbre

Para ilustrar la aparición inevitable de incertidumbre, sólo tenemos que examinar con detenimiento cualquier medición cotidiana. Consideremos, por ejemplo, un carpintero que debe medir la altura de un portal antes de instalar una puerta. Como una primera medida aproximada, puede simplemente mirar el portal y estimar su altura en 210 cm. Esta tosca “medida” está, sin duda, sujeta a incertidumbre. Si se le apura, el carpintero podría expresar dicha incertidumbre admitiendo que la altura puede ser cualquier valor entre 205 cm y 215 cm.

Si quisiera una medida más precisa, usaría una cinta de medir y tal vez encontraría que la altura es 211,3 cm. Esta medida es, sin duda, más precisa que la estimación original, pero obviamente aún está sujeta a cierta incertidumbre, ya que le es imposible saber si la altura es exactamente 211,3000 cm en lugar de, por ejemplo, 211,3001 cm.

Esta incertidumbre remanente tiene muchas causas, varias de las cuales se tratarán en el presente libro. Algunas de las causas podrían ser eliminadas si el carpintero realmente lo necesitara. Por ejemplo, una fuente de incertidumbre podría deberse a una insuficiente iluminación que dificultara la lectura de la cinta métrica; este problema se podría corregir mediante la mejora de la iluminación.

Por otro lado, algunas fuentes de incertidumbre son intrínsecas al proceso de medición y nunca pueden ser completamente eliminadas. Por ejemplo, supongamos que la cinta métrica del carpintero está graduada en medios centímetros. La parte superior de la puerta probablemente no coincidirá exactamente con una de las marcas de medio centímetro, y si no lo hace, el carpintero deberá *estimar* justo donde se encuentra dicha parte superior entre dos marcas adyacentes. Incluso si la parte superior coincide con una de las marcas, la propia marca es quizás de un milímetro de grosor; de modo que deberá estimar justo donde se encuentra dicha parte superior dentro de la marca. En cualquier caso, el carpintero finalmente deberá estimar donde se encuentra la parte superior de la puerta en relación con las marcas en la cinta, y esta necesidad provoca cierta incertidumbre en la medición.

Comprando una cinta métrica mejor con las marcas más finas y más próximas entre sí, el carpintero puede reducir su incertidumbre pero no puede eliminarla por completo. Si se obsesiona con la idea de encontrar la altura de la puerta con la mayor precisión técnicamente posible, podría comprar un interferómetro láser caro. Pero incluso la precisión de un interferómetro está limitada a distancias del orden de la longitud de onda de la luz (aproximadamente $0,5 \times 10^{-6}$ metros). Aunque el carpintero ahora sería capaz de medir la altura con una precisión fantástica, todavía no conoce la altura del portal *exactamente*.

Además, como nuestro carpintero trate de lograr una mayor precisión, se topará con un importante problema de principio. Sin duda encontrará que la altura es diferente en diferentes lugares. Incluso en un mismo lugar, encontrará que la altura varía si la temperatura y la humedad varían, o incluso si accidentalmente se pega a la puerta una fina capa de suciedad. En otras palabras, encontrará que no hay nada que sea *la* altura de la puerta. Este tipo de problema se llama *problema de definición* (la altura de la puerta no es una magnitud bien definida) y juega un papel importante en muchas mediciones científicas.

Nuestra experiencia con el carpintero ilustra un hecho que es generalmente cierto: ninguna magnitud física (longitud, tiempo o temperatura, por ejemplo) se puede medir con total certeza. Con cuidado, podemos ser capaces de reducir las incertidumbres hasta que sean sumamente pequeñas, pero es imposible eliminarlas por completo.

Por lo general, en las mediciones cotidianas no nos molestamos en hablar de las incertidumbres. A veces las incertidumbres simplemente no son interesantes. Si decimos que la distancia entre nuestra casa y la escuela es de 3 kilómetros, el

que esto signifique “alguna cifra entre 2,5 y 3,5 kilómetros” o “algún valor entre 2,99 y 3,01 kilómetros”, generalmente carece de importancia. A menudo las incertidumbres son importantes, pero se pueden admitir de forma instintiva y sin consideración explícita. Cuando mide la puerta, nuestro carpintero debe conocer su altura con una incertidumbre inferior a 1 mm. Siempre que la incertidumbre sea inferior a ese valor, la puerta (para todos los fines prácticos) encajará de forma perfecta, y su preocupación por el análisis de errores concluirá.

1.3 La importancia de conocer las incertidumbres

Nuestro ejemplo del carpintero midiendo una puerta ilustra cómo las incertidumbres están siempre presentes en las mediciones. Veamos ahora un ejemplo que ilustra más claramente la importancia crucial de conocer la magnitud de dichas incertidumbres.

Supongamos que nos encontramos ante un problema como el que se dice que resolvió Arquímedes. Se nos pide que averigüemos si una corona está hecha de oro de 18 quilates, como se afirma, o de una aleación más barata. Siguiendo a Arquímedes, nos decidimos a determinar la densidad de la corona ρ sabiendo que las densidades del oro de 18 quilates y de la presunta aleación son

$$\rho_{\text{oro}} = 19,3 \text{ gramos/cm}^3$$

y

$$\rho_{\text{aleación}} = 13,8 \text{ gramos/cm}^3.$$

Si somos capaces de medir la densidad de la corona, deberíamos ser capaces de decidir (tal como Arquímedes sugirió) si la corona es realmente de oro mediante la comparación de ρ con las densidades conocidas ρ_{oro} y $\rho_{\text{aleación}}$.

Supongamos que convocamos a dos expertos en la medición de densidades. El primer experto, Jorge, podría efectuar una medición rápida de ρ e informar que su mejor estimación de ρ es 15 y que es casi seguro que el verdadero valor se encuentra entre 13,5 y 16,5 gramos/cm³. Nuestro segundo experto, Marta, podría tardar un poco más y luego comunicar su mejor estimación de 13,9 y un intervalo o rango probable de 13,7 a 14,1 gramos/cm³. Las conclusiones de nuestros dos expertos se resumen en la figura 1.1.

El primer aspecto a destacar acerca de estos resultados es que, aunque la medición de Marta es mucho más precisa, la medición de Jorge es, probablemente, también correcta. Cada experto indica un intervalo o rango dentro del cual confía en que esté el verdadero valor de ρ , y estos intervalos se solapan; por tanto es perfectamente posible (e incluso probable) que ambas afirmaciones sean correctas.

Señalamos a continuación que la incertidumbre en la medición de Jorge es tan grande que sus resultados no sirven para nada. Las densidades del oro de 18 quilates y de la aleación se encuentran ambas dentro del intervalo estimado, desde 13,5 hasta 16,5 gramos/cm³; por consiguiente no se pueden sacar conclusiones a partir de las mediciones de Jorge. Por otro lado, las mediciones de Marta indican claramente que la corona no es de auténtico oro, la densidad de la presunta aleación, 13,8, se encuentra holgadamente dentro del intervalo estimado por Marta de 13,7 a 14,1, pero la de oro de 18 quilates, 19,3, está claramente fue-

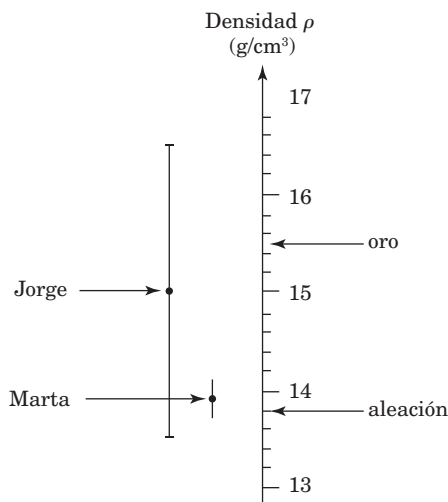


Figura 1.1. Dos determinaciones de la densidad de una corona supuestamente de oro. Los dos puntos negros muestran las mejores estimaciones de la densidad de Jorge y Marta; las dos barras verticales de error muestran sus respectivos márgenes de error, los intervalos o rangos dentro de los cuales creen que se halla el verdadero valor de la densidad. La incertidumbre de Jorge es tan grande que el oro y la presunta aleación caen dentro de sus márgenes de error; por lo tanto, su medida no determina qué metal se utilizó. La incertidumbre de Marta es sensiblemente menor, y su medición muestra claramente que la corona no es de oro.

ra del mismo. Evidentemente, si las mediciones se efectúan para sacar una conclusión, las incertidumbres experimentales no deberían ser demasiado grandes. No obstante, tampoco necesitan ser extremadamente pequeñas. En este sentido, nuestro ejemplo es típico de muchas mediciones científicas, que requieren incertidumbres razonablemente bajas (tal vez un pequeño porcentaje del valor medido), pero que a menudo las precisiones extremas son innecesarias.

Dado que nuestra decisión depende de la afirmación de Marta de que ρ se encuentra entre 13,7 y 14,1 gramos/ cm^3 , ella debe darnos razones para que podamos confiar en su afirmación. En otras palabras, debe *justificar* la amplitud del intervalo de valores de la densidad. Los estudiantes principiantes a menudo pasan por alto este punto y se limitan a indicar sus incertidumbres omitiendo cualquier justificación. Sin una breve explicación de cómo se estima la incertidumbre, la afirmación es casi inútil.

La conclusión más importante acerca de las mediciones de nuestros dos expertos es la siguiente: al igual que la mayoría de las mediciones científicas, ambas no hubieran servido para nada si no se hubieran proporcionado cuantificaciones fiables de sus incertidumbres. De hecho, si sólo hubiésemos conocido las dos mejores estimaciones (15 proporcionada por Jorge y 13,9 por Marta), no sólo no hubiéramos podido llegar a una conclusión válida, sino que en realidad podríamos haber sido inducidos a error, ya que el resultado de Jorge (15) parece sugerir que la corona es de oro auténtico.

1.4 Más ejemplos

Los ejemplos de las dos secciones anteriores no se eligieron por su gran importancia, sino para introducir algunas de las características principales del análisis de errores. Por lo tanto, el lector queda excusado si considera que son ejemplos un poco artificiales. Es fácil, sin embargo, pensar en ejemplos de gran importancia en casi cualquier rama de la ciencia básica o aplicada.

En ciencias aplicadas, por ejemplo, los ingenieros de diseño de una planta de energía deben conocer las características de los materiales y combustibles que se van a utilizar. El fabricante de una calculadora de bolsillo debe conocer las propiedades de los diversos componentes electrónicos. En cada caso, alguien tiene que medir los parámetros requeridos, y de haberlos medido, debe establecer su fiabilidad, lo que requiere el análisis de errores. Los ingenieros que se ocupan de la seguridad de los aviones, trenes o coches deben comprender las incertidumbres asociadas a los tiempos de reacción de los conductores, a las distancias de frenado y a un sinfín de otras variables; si no se hace un análisis de errores, pueden ocurrir accidentes como el que se muestra en la cubierta de este libro. Incluso en campos menos científicos, tales como la fabricación de prendas de vestir, el análisis de errores en la forma de control de calidad juega un papel vital.

En las ciencias básicas, el análisis de errores tiene un papel aún más fundamental. Cuando se proponen nuevas teorías, éstas deben ser contrastadas con otras anteriores mediante uno o varios experimentos en los que las teorías nuevas y viejas predigan resultados diferentes. En principio, un investigador simplemente realiza experimentos y deja que los resultados decidan entre las teorías rivales. En la práctica, sin embargo, la situación se complica debido a las inevitables incertidumbres experimentales. Todas estas incertidumbres deben ser analizadas cuidadosamente, y sus efectos minimizados hasta que los experimentos permitan escoger una teoría aceptable. Es decir, los resultados experimentales, con sus incertidumbres, deben ser *coherentes* con las predicciones de una teoría e *incoherentes* con los de todas las alternativas razonables conocidas. Obviamente, el éxito de este procedimiento depende en gran medida de cómo interprete el científico el análisis de errores y la capacidad de esta interpretación para convencer a otros.

Un ejemplo famoso de validación de una teoría científica es la medición de la curvatura de la luz al pasar cerca del Sol. Cuando Einstein publicó su teoría de la relatividad general en 1916, señaló que la teoría predecía que un rayo de luz de una estrella se desviaría un ángulo $\alpha = 1,8''$ cuando pasara cerca del Sol. La predicción de la teoría clásica más simple sería que no hay ninguna desviación ($\alpha = 0$), y la de un análisis clásico más cuidadoso sería que (como el propio Einstein señaló en 1911) se desviaría un ángulo $\alpha = 0,9''$. En principio, todo lo que se necesitaba era observar una estrella cuando se alineara con el borde del Sol y medir el ángulo de inclinación α . Si el resultado fuera $\alpha = 1,8''$, la relatividad general sería justificada (por lo menos para este fenómeno); si α resultara ser 0 o $0,9''$, la relatividad general sería incorrecta y una de las teorías clásicas corroborada.

En la práctica, la curvatura de la luz causada por el Sol era muy difícil de medir y sólo fue posible durante un eclipse solar. Sin embargo, en 1919 se midió con éxito por Dyson, Eddington y Davidson, quienes establecieron su mejor es-

Capítulo 2

Cómo expresar y utilizar las incertidumbres

Después de leer el capítulo 1, el lector debería tener ahora una idea de la importancia de las incertidumbres experimentales y de cómo aparecen. También debería entender cómo pueden estimarse en unas pocas situaciones sencillas. En este capítulo el lector aprenderá la notación y las reglas básicas del análisis de errores y estudiará ejemplos de su utilización en los experimentos típicos de un laboratorio de física. El objetivo es que se familiarice con el vocabulario básico del análisis de errores y su uso en un laboratorio de nivel introductorio. En el capítulo 3 se comienza un estudio sistemático sobre cómo se evalúan realmente las incertidumbres.

En las secciones 2.1 a 2.3 se definen varios conceptos básicos del análisis de errores y se discuten las normas generales para establecer las incertidumbres. En las secciones 2.4 a 2.6 se analizan cómo se podrían utilizar estas ideas en los experimentos típicos de un laboratorio de física de nivel introductorio. Por último, en las secciones 2.7 a 2.9 se introduce la incertidumbre relativa y se discute su significado.

2.1 Mejor estimación \pm incertidumbre

Hemos visto que la forma correcta de expresar el resultado de una medida es dar la mejor estimación de la cantidad y el intervalo o rango dentro del cual se tiene la convicción en que se halla dicha cantidad. Por ejemplo, el resultado de los intervalos de tiempo descritos en la sección 1.6 se expresó como

$$\begin{aligned}\text{mejor estimación del tiempo} &= 2,4 \text{ s}, \\ \text{rango probable} &: 2,3 \text{ a } 2,5 \text{ s}.\end{aligned}\tag{2.1}$$

En este caso, la mejor estimación, 2,4 s, se encuentra en el punto medio del rango estimado de los valores probables, 2,3 a 2,5 s, como se ha hecho en todos los ejemplos. Esta relación es obviamente natural y aparece en la mayoría de las mediciones. Esto permite que los resultados de la medición se expresen de forma compacta. Por ejemplo, la medición del tiempo establecida en (2.1) se indica generalmente de la siguiente manera:

$$\text{medición del tiempo} = 2,4 \pm 0,1 \text{ s}.\tag{2.2}$$

Esta única ecuación es equivalente a las dos declaraciones efectuadas en (2.1).

En general, el resultado de cualquier medición de una magnitud x se expresa como

$$(\text{valor medido de } x) = x_{\text{mejor}} \pm \delta x. \quad (2.3)$$

Esta afirmación significa, en primer lugar, que la mejor estimación del experimentador de la magnitud en cuestión es el número indicado como x_{mejor} , y segundo, que el experimentador está razonablemente seguro de que dicho valor se halla entre $x_{\text{mejor}} - \delta x$ y $x_{\text{mejor}} + \delta x$. Al número δx se le denomina *incertidumbre*, *error* o *margen de error* de la medición de x . Por conveniencia, la incertidumbre δx se define siempre como positiva, de modo que $x_{\text{mejor}} + \delta x$ es siempre el *mayor* valor probable de la cantidad medida y $x_{\text{mejor}} - \delta x$ el *menor*.

He dejado intencionadamente el sentido del intervalo desde $x_{\text{mejor}} - \delta x$ hasta $x_{\text{mejor}} + \delta x$ un tanto ambiguo, aunque a veces podamos darle una interpretación más precisa. En una medición sencilla, como la de la altura de una puerta, es fácil indicar un intervalo de $x_{\text{mejor}} - \delta x$ a $x_{\text{mejor}} + \delta x$ dentro del cual podemos estar *absolutamente* seguros de que se halle el valor real de la misma. Desafortunadamente, en la mayoría de las mediciones científicas tal afirmación es difícil de hacer. En particular, para estar *completamente* seguros de que la magnitud medida se encuentra entre $x_{\text{mejor}} - \delta x$ y $x_{\text{mejor}} + \delta x$, por lo general hay que elegir un valor de δx demasiado grande para ser útil. Para evitar esta situación, a veces podemos elegir un valor de δx que nos permita afirmar, con un cierto porcentaje de confianza, que la cantidad real se encuentra dentro del intervalo $x_{\text{mejor}} \pm \delta x$. Por ejemplo, los sondeos de opinión que se realizan durante las elecciones se presentan tradicionalmente con márgenes de error que representan el 95% de los límites de confianza. La afirmación de que el 60% del electorado está a favor del candidato A, con un margen de error de 3 puntos porcentuales (60 ± 3), significa que los encuestadores están un 95% seguros de que el porcentaje real de los votantes a favor del candidato A está entre 57 y 63, en otras palabras, después de muchas elecciones, debemos esperar que la respuesta correcta haya estado *dentro* de los márgenes establecidos de error un 95% de las veces y *fuera* de dichos márgenes sólo el 5% de las veces.

Obviamente, no podemos establecer un porcentaje de confianza en nuestros márgenes de error hasta que comprendamos las leyes estadísticas que gobiernan el proceso de medición. Volveremos a este punto en el capítulo 4. Por ahora, vamos a contentarnos con la definición de la incertidumbre δx como aquel valor que nos hace estar “razonablemente seguros” de que la magnitud que pretendemos medir se encuentra entre $x_{\text{mejor}} - \delta x$ y $x_{\text{mejor}} + \delta x$.

Test rápido¹ 2.1. (a) Un estudiante mide la longitud de un péndulo simple y considera que su mejor estimación es 110 mm y el intervalo

1. Estos “test rápidos” aparecen a lo largo del texto para dar al lector la oportunidad de comprobar su comprensión de los conceptos que se acaban de presentar. Son ejercicios sencillos, y muchos se pueden hacer mentalmente. Se recomienda tomarse el tiempo suficiente para asegurarse de que se saben hacer; en caso contrario, el lector debe releer los párrafos anteriores.

en el que probablemente se encuentra la longitud es de 108 a 112 mm. Vuelva a escribir este resultado según la forma estándar (2.3). **(b)** Si otro estudiante informa que su medida de una corriente es $I = 3,05 \pm 0,03$ amperios, ¿cuál es el intervalo dentro del cual probablemente se encuentra el verdadero valor de I ?

2.2 Cifras significativas

Merece la pena destacar varias reglas básicas para expresar incertidumbres. En primer lugar, puesto que la cantidad δx es una estimación de una incertidumbre, obviamente no debería indicarse con demasiada precisión. Si medimos la aceleración de la gravedad g , sería absurdo presentar el siguiente resultado

$$(\text{medida de } g) = 9,82 \pm 0,02385 \text{ m/s}^2. \quad (2.4)$$

La incertidumbre en la medición no es concebible que se pueda conocer a cuatro cifras significativas. En un trabajo de alta precisión, las incertidumbres se indican a veces con dos cifras significativas, pero para nuestros propósitos podemos establecer la siguiente regla:

<p>Regla para indicar las incertidumbres</p> <p>Las incertidumbres experimentales deben casi siempre redondearse a una cifra significativa.</p>	(2.5)
--	-------

Por lo tanto, si algún cálculo arroja la incertidumbre $\delta g = 0,02385 \text{ m/s}^2$, este resultado debe redondearse a $\delta g = 0,02 \text{ m/s}^2$, y la conclusión (2.4) debe reescribirse como

$$(\text{medida de } g) = 9,82 \pm 0,02 \text{ m/s}^2. \quad (2.6)$$

Una consecuencia práctica importante de esta regla es que muchos cálculos de error pueden realizarse mentalmente sin utilizar una calculadora o incluso sin lápiz ni papel.

La regla (2.5) tiene una sola excepción importante. Si el primer dígito de la incertidumbre δx es un 1, entonces puede ser mejor mantener dos cifras significativas en δx . Por ejemplo, supongamos que algún cálculo dio una incertidumbre $\delta x = 0,14$. Redondear este número a $\delta x = 0,1$ sería, en proporción, una reducción sustancial, y por tanto se puede argumentar que mantener dos cifras podría ser menos engañoso, y nos quedaríamos con $\delta x = 0,14$. Tal vez se podría aplicar el mismo argumento si el primer dígito fuera un 2, pero ciertamente no si se tratara de uno mayor.

Una vez que se ha estimado la incertidumbre en la medición, se deben considerar las cifras significativas en el valor medido. Un resultado tal como

$$\text{velocidad medida} = 6051,78 \pm 30 \text{ m/s} \quad (2.7)$$

es obviamente ridículo. Una incertidumbre de 30 significa que el dígito 5 en realidad podría ser tan pequeño como 2 o tan grande como 8. Es evidente que los

dígitos finales 1, 7 y 8 no tienen significado en absoluto y deben ser redondeados. Esto es, el resultado correcto (2.7) es

$$\text{velocidad medida} = 6050 \pm 30 \text{ m/s.} \quad (2.8)$$

La regla general es la siguiente:

Regla para expresar resultados

La última cifra significativa de cualquier resultado por lo general debe ser del mismo orden de magnitud (en la misma posición decimal) que la incertidumbre. (2.9)

Por ejemplo, el resultado 92,81 con una incertidumbre de 0,3 debe redondearse a

$$92,8 \pm 0,3.$$

Si la incertidumbre fuera 3, entonces la respuesta debería redondearse como

$$93 \pm 3,$$

y si la incertidumbre fuese 30, entonces la respuesta debería ser

$$90 \pm 30.$$

Un requisito importante para reglas (2.5) y (2.9) es el siguiente: para reducir los errores causados por el redondeo, *cualquier número que haya de utilizarse en cálculos posteriores normalmente debería retener al menos una cifra significativa más que la que finalmente queremos justificar*. Al terminar los cálculos, el resultado final debe redondearse para eliminar estas cifras adicionales no significativas. Una calculadora electrónica utilizará con facilidad bastantes más dígitos significativos de los que probablemente son significativos en cualquier cálculo que se realice en un laboratorio. Obviamente, estos números no necesitan ser redondeados en cálculos intermedios, pero ciertamente deben redondearse apropiadamente en la presentación del resultado final.²

Observe que la incertidumbre en la medida de cualquier magnitud tiene las mismas dimensiones que la propia magnitud. Por lo tanto, la escritura de las unidades (m/s², cm³, etc.) *solamente* después de la incertidumbre es más clara y concisa, como en las ecuaciones (2.6) y (2.8). Por la misma razón, si el número medido es tan grande o pequeño que requiere notación científica (el uso de la forma 3×10^3 en lugar de 3000, por ejemplo), entonces es más sencillo y claro expresar la respuesta y la incertidumbre en la misma notación. Por ejemplo, el resultado

$$\text{carga medida} = (1,61 \pm 0,05) \times 10^{-19} \text{ culombios}$$

es mucho más fácil de leer y entender en esta forma que en la siguiente

$$\text{carga medida} = 1,61 \times 10^{-19} \pm 5 \times 10^{-21} \text{ culombios.}$$

2. La regla (2.9) tiene una pequeña excepción. Si el primer dígito de la incertidumbre es pequeño (un 1 o, tal vez, un 2), mantener un dígito adicional en la respuesta final puede ser apropiado. Por ejemplo, una respuesta tal como $3,6 \pm 1$ es bastante aceptable, ya que se podría argumentar que el redondeo a 4 ± 1 podría suponer una pérdida apreciable de precisión.

Test rápido 2.2. Reescriba cada una de las siguientes mediciones en su forma más apropiada:

- (a) $v = 8,123456 \pm 0,0312$ m/s
 (b) $x = 3,1234 \times 10^4 \pm 2$ m
 (c) $m = 5,6789 \times 10^{-7} \pm 3 \times 10^{-9}$ kg.
-

2.3 Discrepancia

Antes de abordar la cuestión de cómo utilizar las incertidumbres en los informes experimentales, debo introducir y definir algunos términos importantes. En primer lugar, si dos mediciones de una misma magnitud no coinciden, decimos que hay una *discrepancia*. Numéricamente, la discrepancia entre dos mediciones se define como su diferencia:

$$\text{discrepancia} = \text{diferencia entre dos valores medidos de la misma magnitud.} \quad (2.10)$$

Más específicamente, cada una de las dos mediciones se compone de la mejor estimación y de una incertidumbre, y la discrepancia se define como la diferencia entre las dos mejores estimaciones. Por ejemplo, si dos estudiantes miden la misma resistencia eléctrica y obtienen los siguientes resultados

estudiante A: 15 ± 1 ohmios

y

estudiante B: 25 ± 2 ohmios,

su discrepancia es

$$\text{discrepancia} = 25 - 15 = 10 \text{ ohmios.}$$

Admitamos que una discrepancia puede o no ser *significativa*. Las dos medidas que acabamos de mencionar se ilustran en la figura 2.1(a), que muestra claramente que la discrepancia de 10 ohmios es *significativa* porque ningún valor individual de la resistencia es compatible con ambas mediciones. Obviamente, al menos una medición no es correcta, y se necesitarán algunas comprobaciones cuidadosas para averiguar lo que salió mal.

Supongamos, por otra parte, que otros dos estudiantes han obtenido los siguientes resultados:

estudiante C: 16 ± 8 ohmios

y

estudiante D: 26 ± 9 ohmios.

Aquí de nuevo, la discrepancia es de 10 ohmios, pero en este caso la discrepancia *no es significativa* ya que, tal como se muestra en la figura 2.1(b), los márgenes de error de los dos estudiantes se superponen ampliamente y ambas mediciones podrían muy bien ser correctas. La discrepancia entre las dos mediciones de la

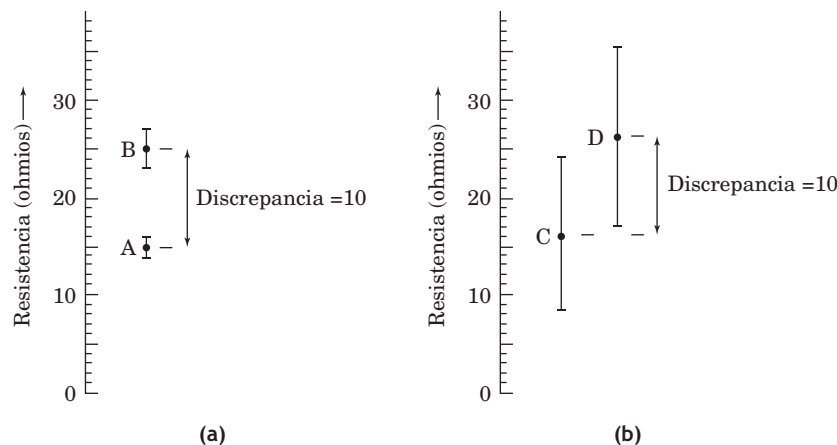


Figura 2.1. (a) Dos mediciones de la misma resistencia. Cada medición incluye la mejor estimación, indicada mediante un punto negro, y un intervalo de valores probables, indicado por medio de una barra de error vertical. La discrepancia (diferencia entre las dos mejores estimaciones) es de 10 ohmios y es *significativa* porque es mucho mayor que la incertidumbre combinada de las dos mediciones. Es casi seguro que al menos uno de los experimentadores haya cometido un error. (b) Dos mediciones diferentes de la misma resistencia. La discrepancia es de nuevo de 10 ohmios, pero en este caso *no es significativa* debido a que los márgenes de error se superponen. No hay ninguna razón para dudar de cualquiera de las mediciones (aunque se podrían criticar por su falta de precisión).

misma magnitud debe ser juzgada no sólo por su tamaño, sino, lo que es más importante, por su tamaño *comparado con las incertidumbres de las mediciones*.

En prácticas de laboratorio se le puede solicitar al estudiante que mida una cantidad que ya se ha determinado cuidadosamente con anterioridad muchas otras veces, y para la cual se conoce un *valor aceptado* exacto publicado, por ejemplo, la carga del electrón o la constante universal de los gases. Estos valores aceptados no son exactos, por supuesto, son el resultado de mediciones y, al igual que todas las mediciones, tienen cierto grado de incertidumbre. Sin embargo, en muchos casos el valor aceptado es mucho más preciso que lo que el estudiante puede lograr. Por ejemplo, el valor actualmente aceptado de la constante universal de los gases, R , es

$$(\text{valor aceptado } R) = 8,31451 \pm 0,00007 \text{ J/(mol}\cdot\text{K)}. \quad (2.11)$$

Como era de esperar, este valor es incierto, pero la incertidumbre es muy pequeña para los estándares de la mayoría de los laboratorios de enseñanza. Por lo tanto, cuando se compara el valor medido de una constante con su valor aceptado, por lo general este último se puede considerar como exacto.³

3. Esto no siempre es así. Por ejemplo, si se mira el índice de refracción del vidrio, se encuentran valores que van desde 1,5 a 1,9, dependiendo de la composición del vidrio. En un experimento para medir el índice de refracción de un trozo de vidrio cuya composición es desconocida, el valor aceptado es, por lo tanto, no más que una indicación aproximada del valor real.

Capítulo 3

Propagación de incertidumbres

Con frecuencia, la mayoría de las magnitudes físicas no se pueden medir mediante una sola observación directa, sino que se determinan en dos pasos sucesivos. En primer lugar, se miden una o varias magnitudes que *son* directamente observables y que intervendrán en el cálculo de la variable de interés. En segundo lugar, estos valores observados se utilizan para el cálculo efectivo de dicha variable. Por ejemplo, para hallar el área de un rectángulo, en realidad se mide su longitud l y la altura h para después calcular su área A como $A = lh$. Del mismo modo, la manera más obvia de encontrar la velocidad v de un objeto es medir la distancia recorrida, d , y el tiempo empleado en hacerlo, t , y luego calcular v a partir de $v = d/t$. Cualquier lector con experiencia básica de laboratorio puede pensar en más ejemplos. En realidad, si reflexionamos un poco, nos daremos cuenta de que casi todas las mediciones de interés implican estos dos pasos de medición directa a los que le siguen ciertos cálculos.

Cuando una medición implica estos dos pasos, la estimación de la incertidumbre implica también dos pasos. En primer lugar debemos estimar las incertidumbres en las magnitudes medidas directamente y luego determinar cómo estas incertidumbres “se propagan” a lo largo de los cálculos para producir la incertidumbre de la respuesta final.¹ Esta propagación de errores es el tema principal de este capítulo.

De hecho, en el capítulo 2 ya se presentaron algunos ejemplos de propagación de errores. En la sección 2.5 estudiamos lo que sucede cuando determinamos dos variables x e y y los resultados se utilizan para calcular la diferencia $q = x - y$. Allí vimos que la incertidumbre de q es simplemente la *suma* $\delta q \approx \delta x + \delta y$ de las incertidumbres de x e y . En la sección 2.9 se expone el producto $q = xy$, y el problema 2.13 trata de la suma $q = x + y$. Revisaremos estos casos en la sección 3.3; el resto del presente capítulo se dedica a casos más generales de la propagación de las incertidumbres e incluye varios ejemplos.

1. En el capítulo 4 se discute otra forma de estimar la incertidumbre final. Si todas las mediciones se pueden repetir varias veces, y si todas las incertidumbres conocidas son de naturaleza aleatoria, entonces la incertidumbre en la variable de interés se puede estimar examinando la dispersión de los resultados. Aun cuando este método es posible, por lo general es mejor utilizarlo para la comprobación del procedimiento en dos etapas que describimos en el presente capítulo.

Antes de abordar la propagación de errores en la sección 3.3, en las secciones 3.1 y 3.2 analizaremos brevemente la estimación de las incertidumbres de las cantidades medidas directamente. Repasaremos los métodos presentados en el capítulo 1 y daremos más ejemplos de la estimación de errores en las mediciones directas.

A partir de la sección 3.3 nos ocuparemos de la propagación de errores. El lector aprenderá que casi todos los problemas de propagación de errores pueden ser resueltos mediante tres reglas sencillas.

También presentaremos una única regla, más complicada, que abarca todos los casos y que permite deducir cada una de las tres reglas más sencillas vistas anteriormente.

El presente capítulo es largo, pero su longitud se debe simplemente a su gran importancia. La propagación de errores es una técnica que el lector utilizará con frecuencia en el laboratorio, y es necesario familiarizarse con los métodos descritos aquí. La única excepción es que el material de la sección 3.11 no se utiliza de nuevo hasta la sección 5.6; por lo tanto, si las ideas de este capítulo son nuevas para el lector, considere la posibilidad de omitir la sección 3.11 en una primera lectura.

3.1 Incertidumbre de las mediciones directas

Casi todas las mediciones directas implican la lectura de una escala (por ejemplo una regla, un reloj o un voltímetro) o una pantalla digital (por ejemplo en un reloj o un voltímetro digitales). Algunos problemas en la lectura de escalas se tratan en la sección 1.5. A veces, las principales fuentes de incertidumbre son la lectura de la escala y la necesidad de interpolar entre sus marcas. En tales situaciones, una estimación razonable de la incertidumbre se hace fácilmente. Por ejemplo, si el lector tiene que medir una longitud l claramente definida y dispone de una regla graduada en milímetros, de una forma razonable podría considerar que la longitud se puede obtener al milímetro, pero no con mayor precisión. Aquí, la incertidumbre δl sería $\delta l = 0,5 \text{ mm}$. Si las marcas de escala están más separadas (como en décimas de pulgada), el lector podría considerar razonable observar hasta una quinta parte de una división, por ejemplo. En cualquier caso, las incertidumbres asociadas a la lectura de una escala se pueden, obviamente, estimar fácilmente y de manera realista.

Desafortunadamente, con frecuencia hay otras fuentes de incertidumbre mucho más importantes que las dificultades de lectura de una escala graduada. Al medir la distancia entre dos puntos, el principal problema puede ser decidir dónde están en realidad esos dos puntos. Por ejemplo, en un experimento de óptica, el lector puede querer medir la distancia q desde el centro de una lente hasta una imagen enfocada, como en la figura 3.1. En la práctica, el objetivo tiene por lo general varios milímetros de espesor y localizar su centro es difícil; si el objetivo forma parte de un dispositivo voluminoso, como ocurre a menudo, la localización del centro de la lente es todavía más difícil. Además, la imagen puede parecer bien enfocada a lo largo de un intervalo de varios milímetros. A pesar de que el aparato estuviese montado en un banco óptico claramente graduado en

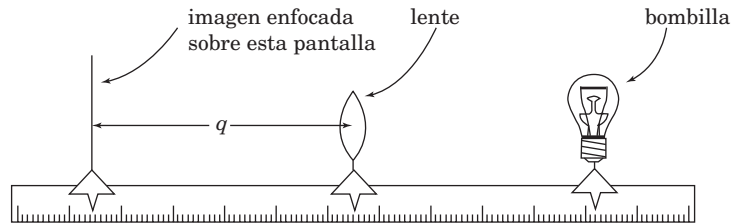


Figura 3.1. La imagen de la bombilla situada a la derecha es enfocada por la lente sobre la pantalla situada a la izquierda.

milímetros, la incertidumbre en la distancia de la lente a la imagen podría ser fácilmente de un centímetro aproximadamente. Puesto que esta incertidumbre se debe a que los dos puntos en cuestión no están claramente definidos, a este tipo de problema lo llamaremos *problema de definición*.

Este ejemplo ilustra un grave peligro en la estimación de errores. Si nos fijamos sólo en las escalas y olvidamos otras fuentes de incertidumbre, podemos subestimar muchísimo la incertidumbre total. De hecho, el error más común del estudiante novel es el pasar por alto algunas fuentes de incertidumbre y por lo tanto *subestimar* la incertidumbre, a menudo en factores de 10 o más. Por supuesto, también debe evitar los errores de *sobrestimación*. Los experimentadores que deciden ir sobre seguro y proporcionan incertidumbres generosas en todas las mediciones pueden evitar incoherencias embarazosas, pero sus mediciones tal vez no sean de mucha utilidad. Evidentemente, lo ideal es encontrar todas las posibles causas de incertidumbre y estimar con precisión sus efectos, lo que a menudo no es tan difícil como parece.

Al menos a primera vista, la lectura en un medidor digital es mucho más fácil que en un medidor analógico convencional. A menos que un medidor digital sea defectuoso, sólo debe mostrar las cifras significativas. Así, por lo general es seguro decir que el número de cifras significativas en una lectura digital es precisamente el número de cifras mostradas. Desafortunadamente, tal como se discute en la sección 2.8, el significado exacto de lo que entendemos por cifras significativas no siempre es claro. Así, un voltímetro digital que nos indica una tensión $V = 81$ microvoltios puede significar que la incertidumbre en el dígito correspondiente a las unidades es un valor cualquiera entre $\delta V = 0,5$ a $\delta V = 1$ o incluso más. Sin un manual que indique la incertidumbre de un medidor digital, un supuesto razonable es que la incertidumbre correspondiente al último dígito es ± 1 (de modo que la tensión que acabamos de mencionar sería $V = 81 \pm 1$).

Los medidores digitales, incluso más que los analógicos, pueden dar una impresión equivocada de precisión. Por ejemplo, un estudiante puede utilizar un cronómetro digital para cronometrar la caída de un peso en una máquina de Atwood o un dispositivo similar. Si el cronómetro muestra 8,01 segundos, el tiempo de caída es aparentemente

$$t = 8,01 \pm 0,01 \text{ s.} \quad (3.1)$$

Sin embargo, el estudiante cuidadoso que repite el experimento en condiciones casi idénticas podría realizar una segunda medición y obtener 8,41 s, es decir,

$$t = 8,41 \pm 0,01 \text{ s.}$$

Una explicación probable de esta gran discrepancia es que las incertidumbres en la puesta en marcha del cronómetro influyen en la determinación del momento exacto en que empieza la caída y por tanto en la duración de la misma; es decir, los tiempos realmente medidos son diferentes. En cualquier caso, la precisión que se reivindica en la ecuación (3.1) es evidente y absurdamente demasiado buena. Teniendo en cuenta las dos mediciones realizadas, una respuesta más realista sería

$$t = 8,2 \pm 0,2 \text{ s.}$$

En particular, la incertidumbre es unas 20 veces mayor que la sugerida en la ecuación (3.1) basada en una única lectura original.

Este ejemplo nos lleva a otro punto mencionado en el capítulo 1: siempre que una medición se pueda repetir, por lo general conviene hacerlo varias veces. La dispersión de los valores resultantes a menudo proporciona una buena indicación de las incertidumbres, y el promedio de los valores es casi siempre más fiable que cualquier medición individual. En los capítulos 4 y 5 se discuten el tratamiento estadístico de las mediciones múltiples. Aquí, sólo destaco que si una medición es repetible, debe repetirse, tanto para obtener una respuesta más fiable (usando el promedio) y, más importante, para obtener una estimación de la incertidumbre. Lamentablemente, como también se menciona en el capítulo 1, la repetición de una medición no siempre revela todas las incertidumbres. Si la medición está sujeta a errores sistemáticos, que distorsionan todos los resultados en un mismo sentido (por ejemplo, un reloj que atrasa), la dispersión de los resultados no reflejará este error sistemático. La eliminación de estos errores sistemáticos requiere un cuidadoso control de la calibración y de los procedimientos.

3.2 La regla de la raíz cuadrada para experimentos de recuento

Otro tipo diferente de observación directa tiene una incertidumbre que se puede estimar fácilmente. Algunos experimentos requieren que el investigador cuente sucesos que ocurren al azar pero con una determinada tasa media. Por ejemplo, el nacimiento de bebés en un hospital se da de una manera bastante aleatoria, pero a largo plazo el número de nacimientos en cualquier hospital probablemente se da con una cierta frecuencia media a lo largo de tiempo. Imaginemos que un demógrafo quiere conocer esta tasa y contabiliza 14 nacimientos en un determinado período de dos semanas en un hospital local. Basándose en este resultado, razonablemente dirá que su mejor estimación del número esperado de nacimientos en dos semanas es de 14. A menos que haya cometido un error, 14 es *exactamente* el número de nacimientos en el período de dos semanas que eligió para observar. Sin embargo, como los nacimientos se producen de forma aleatoria, el 14 obviamente puede *no* ser igual al promedio real de nacimientos en todos los períodos de dos semanas. Tal vez este número sea 13, 15 o incluso un número con decimales tal como 13,5 o 14,7.

Evidentemente, la incertidumbre en este tipo de experimento no está en el número de sucesos observados (14 en nuestro ejemplo). En cambio, la incerti-

dumbre está en lo bien que dicha cifra observada se aproxima al promedio real. El problema consiste en estimar cuan grande es esta incertidumbre. Aunque discutiremos la teoría de estos experimentos de recuento en el capítulo 11, la respuesta es muy simple y la podemos enunciar fácilmente aquí: en el recuento de sucesos aleatorios para estimar el verdadero valor medio, la incertidumbre es igual a *la raíz cuadrada del número de sucesos contabilizados*. En nuestro ejemplo, el demógrafo contó 14 nacimientos en un determinado período de dos semanas. Por lo tanto, su incertidumbre es igual a $\sqrt{14} \approx 4$, y su conclusión final sería

(promedio de nacimientos en un período de dos semanas) = 14 ± 4 .

Para generalizar, supongamos que contamos las veces que ocurre un determinado suceso (como los nacimientos de bebés en un hospital) que se da al azar, pero con una tasa media determinada. Concretamente contabilizamos dichos sucesos durante un determinado intervalo de tiempo T (por ejemplo, dos semanas), y denotamos el número de sucesos observados por la letra griega v (se pronuncia “nu”; es la forma griega de la letra n y representa *número*). Basándonos en este experimento, la mejor estimación para el número medio de sucesos durante el tiempo T es, por supuesto, el número observado v , y la incertidumbre correspondiente a esta estimación es igual la raíz cuadrada de dicho número, es decir, \sqrt{v} . Por lo tanto, nuestra respuesta para el número promedio de sucesos durante el tiempo T es

$$(\text{promedio de sucesos durante el tiempo } T) = v \pm \sqrt{v}. \quad (3.2)$$

Nos referiremos a este importante resultado como la *regla de la raíz cuadrada para experimentos de recuento*.

Los experimentos de recuento de este tipo son frecuentes en un laboratorio de física. El ejemplo más destacado se da en el estudio de la radiactividad. Dado un material radiactivo, el tiempo que tarda cada núcleo en desintegrarse es un tiempo aleatorio, pero en una muestra grande la desintegración se produce a una determinada tasa media. Para encontrar esa tasa, sólo tenemos que contar el número v de desintegraciones en algún intervalo de tiempo T conveniente; el número esperado de desintegraciones durante el tiempo T , con su incertidumbre, viene dado por la regla de la raíz cuadrada, (3.2).

Test rápido 3.1. (a) Para comprobar la actividad de una muestra radiactiva, un investigador coloca la muestra en un contador de centelleo líquido con el fin de contar el número de desintegraciones que ocurren durante un intervalo de dos minutos y el resultado es 33 desintegraciones. En su informe, ¿cuál será el número de desintegraciones producidas en la muestra durante dos minutos? (b) Supongamos, en cambio, que hubiera estudiado la misma muestra durante 50 minutos y hubiera obtenido 907 desintegraciones. ¿Cuál debería ser su respuesta para el número de desintegraciones en 50 minutos? (c) Halle las incertidumbres porcentuales en estas dos mediciones, y comente la utilidad de contar durante períodos largos, como en la parte (b).

3.3 Sumas y restas; productos y cocientes

En lo que queda de este capítulo supondremos que hemos medido una o varias variables x, y, \dots , con las incertidumbres correspondientes $\delta x, \delta y, \dots$, y que ahora deseamos utilizar dichos valores de x, y, \dots , para calcular un valor q , que es el que realmente nos interesa. El cálculo de q suele ser sencillo, pero el problema es cómo las incertidumbres, $\delta x, \delta y, \dots$, se propagan a lo largo del cálculo y conducen a una incertidumbre δq del valor final de q .

SUMAS Y RESTAS

En el capítulo 2 vimos lo que sucede cuando se miden dos cantidades x e y y se calcula su suma, $x + y$, o su diferencia, $x - y$. Para estimar la incertidumbre en la suma o la diferencia, sólo tenemos que decidir sobre sus valores probables máximos y mínimos. Los valores más altos y más bajos posibles de x son $x_{\text{mejor}} \pm \delta x$, y los de y son $y_{\text{mejor}} \pm \delta y$. Por lo tanto, el valor probable más alto de $x + y$ es

$$x_{\text{mejor}} + y_{\text{mejor}} + (\delta x + \delta y),$$

y el valor probable más bajo es

$$x_{\text{mejor}} + y_{\text{mejor}} - (\delta x + \delta y).$$

Por lo tanto, la mejor estimación de $q = x + y$ es

$$q_{\text{mejor}} = x_{\text{mejor}} + y_{\text{mejor}},$$

y su incertidumbre es

$$\delta q \approx \delta x + \delta y. \quad (3.3)$$

Un argumento similar (asegúrese de que usted puede razonarlo) muestra que la incertidumbre en la *diferencia* $x - y$ viene dada por la misma fórmula (3.3). Es decir, la incertidumbre, ya sea en la suma $x + y$ o la diferencia $x - y$ es igual a la suma $\delta x + \delta y$ de las incertidumbres de x e y .

Si tenemos varios números x, \dots, w para sumar o restar, entonces la aplicación repetida de (3.3) da lugar a la siguiente regla provisional.

Incertidumbre de sumas y restas (regla provisional)

Si hemos medido varias variables x, \dots, w con incertidumbres $\delta x, \dots, \delta w$ y los valores obtenidos se utilizan para calcular

$$q = x + \dots + z - (u + \dots + w), \quad (3.4)$$

entonces la incertidumbre en el valor calculado de q es igual a la suma,

$$\delta q \approx \delta x + \dots + \delta z + \delta u + \dots + \delta w,$$

de todas las incertidumbres originales.

Capítulo 4

Análisis estadístico de incertidumbres aleatorias

Hemos visto que una de las mejores formas de evaluar la fiabilidad de una medición es repetirla varias veces y examinar los diferentes valores obtenidos. En este capítulo y en el capítulo 5 se describen los métodos estadísticos para el análisis de mediciones efectuadas de este modo.

Como ya hemos señalado, no todos los tipos de incertidumbres experimentales se pueden evaluar mediante un análisis estadístico basado en mediciones repetidas. Por esta razón, las incertidumbres se clasifican en dos grupos: las incertidumbres *aleatorias*, que *pueden* ser tratadas estadísticamente, y las incertidumbres *sistemáticas*, que *no pueden* ser tratadas de ese modo. Esta distinción se describe en la sección 4.1. La mayor parte del resto de este capítulo está dedicado a las incertidumbres aleatorias. La sección 4.2 introduce, sin justificación formal, dos definiciones importantes relacionadas con una serie de valores observados x_1, \dots, x_N , todos ellos correspondientes a una cierta variable x a determinar. En primer lugar, definiremos \bar{x} como el *promedio* o la *media* de x_1, \dots, x_N . En condiciones adecuadas, \bar{x} es la mejor estimación de x basada en los valores medidos x_1, \dots, x_N . Luego definiremos la *desviación estándar* de x_1, \dots, x_N , que denotaremos por σ_x y que caracteriza la incertidumbre media en los valores individuales observados x_1, \dots, x_N . En la sección 4.3 daremos un ejemplo de utilización de la desviación estándar.

En la sección 4.4 se introduce la importante noción de la *desviación estándar de la media*. Este parámetro se denota $\sigma_{\bar{x}}$ y caracteriza la incertidumbre de la media \bar{x} como la mejor estimación de x . En la sección 4.5 daremos ejemplos de la desviación estándar de la media. Por último, en la sección 4.6, volveremos al molesto problema de los errores sistemáticos.

En ninguna parte del presente capítulo trataremos de efectuar una justificación completa de los métodos descritos. El objetivo principal es introducir las fórmulas básicas y describir la forma en que se utilizan. En el capítulo 5 daremos las justificaciones apropiadas, basadas en la importancia de la distribución normal.

La relación del material del presente capítulo (análisis estadístico) con el material del capítulo 3 (propagación de errores) merece un comentario. Desde un punto de vista práctico, ambos temas pueden ser vistos como dos ramas del análisis de errores, independientes, aunque relacionadas (un poco como el álgebra y la geometría, que son especialidades matemáticas independientes, pero relacio-

nadas). Ambos temas se deben dominar, porque la mayoría de los experimentos requieren el uso de ambos.

En unos pocos tipos de experimentos, los papeles de la propagación de errores y del análisis estadístico son complementarios. Es decir, el experimento puede ser analizado utilizando tanto la propagación de errores como los métodos estadísticos.

Consideremos un ejemplo: supongamos que decidimos medir la aceleración de la gravedad, g , midiendo el periodo, T , y la longitud, l , de un péndulo simple. Puesto que $T = 2\pi\sqrt{l/g}$, se puede expresar g como $g = 4\pi^2 l/T^2$. Es posible que decidamos repetir este experimento usando varios valores diferentes de l y midiendo el periodo T correspondiente para cada uno de ellos. De esta manera, obtendríamos varios valores de g . Para encontrar la incertidumbre correspondiente a estos valores de g , podríamos proceder de cualquiera de las dos maneras. Si podemos estimar de manera realista las incertidumbres en las mediciones de l y T , podemos estudiar la propagación de estas incertidumbres para encontrar las incertidumbres de los valores resultantes de g . Por otra parte, teniendo en cuenta los diversos valores de g , podemos analizarlos estadísticamente; en particular, su *desviación estándar* será una buena medida de su incertidumbre. Desafortunadamente, no existe en realidad la opción de elegir cómo hallar la incertidumbre. Si la incertidumbre se puede determinar de estas dos formas, entonces debería hacerse *ambas* para comprobar que proporcionan la misma respuesta, al menos aproximadamente.

4.1 Errores aleatorios y sistemáticos

Las incertidumbres experimentales que pueden evidenciarse mediante la repetición de las mediciones se denominan errores *aleatorios*; mientras que aquellas que no pueden evidenciarse de esta forma se llaman errores *sistemáticos*. Para ilustrar esta distinción, vamos a considerar algunos ejemplos. Supongamos en primer lugar que cronometramos una revolución de un tocadiscos que gira de manera constante. Una fuente de error será nuestro tiempo de reacción en arrancar y parar el reloj. Si nuestro tiempo de reacción fuera siempre exactamente el mismo, estos dos retardos se anularían entre sí. En la práctica, sin embargo, nuestro tiempo de reacción variará. Podemos demorarnos más en el inicio, y así subestimar el tiempo de una revolución, o podemos retrasarnos más en la detención, y así sobrestimarlo. Dado que cualquiera de las posibilidades es igualmente probable, el signo del efecto es *aleatorio*. Si repetimos la medición varias veces, a veces sobrestimaremos y a veces subestimaremos dicho tiempo. Por lo tanto, nuestro tiempo de reacción variable se manifestará como una variación de las respuestas encontradas. Mediante el análisis estadístico de la variabilidad de los resultados, podemos obtener una estimación muy fiable de este tipo de error.

Por otro lado, si nuestro cronómetro funciona mal y retrasa sistemáticamente, todos nuestros tiempos serán subestimaciones, y ninguna repetición de la observación (con el mismo reloj) revelará esta fuente de error. Este tipo de error se denomina *sistemático*, ya que siempre altera nuestro resultado en un mismo sentido. (Si el reloj retrasa, siempre subestimaremos el tiempo; si el reloj adelan-

ta, siempre lo sobrestimaremos). Los errores sistemáticos no se pueden descubrir mediante el tipo de análisis estadístico contemplado aquí.

Como un segundo ejemplo de los errores aleatorios en comparación con los sistemáticos, suponga que tenemos que medir con una regla alguna longitud bien definida. Una fuente de incertidumbre será la necesidad de interpolar entre marcas de escala; y esta incertidumbre es probablemente aleatoria. (Cuando interpolamos, tenemos probablemente las mismas probabilidades de sobrestimar que de subestimar). Pero también existe la posibilidad de que nuestra regla se haya distorsionado; y esta fuente de incertidumbre sería probablemente sistemática. (Si la regla se ha dilatado, siempre subestimaremos las longitudes medidas; si se ha contraído, siempre las sobrestimaremos).

Al igual que en estos dos ejemplos, casi todas las mediciones están sujetas tanto a incertidumbres aleatorias como sistemáticas. El lector no debería tener problemas en encontrar más ejemplos. En particular, debe tenerse en cuenta que las fuentes comunes de incertidumbres aleatorias pueden ser pequeños errores de juicio por parte del observador (como al interpolar), alteraciones pequeñas del aparato (por ejemplo, las vibraciones mecánicas), problemas de definición y algunos otros. Quizá la causa más evidente de error sistemático es la mala calibración de los instrumentos, tales como el reloj que atrasa, la regla que se ha dilatado, o un medidor que no está bien puesto a cero.

Para tener una mejor idea de la diferencia entre los errores aleatorios y sistemáticos, considere la analogía mostrada en la figura 4.1. Aquí el “experimento” es una serie de disparos a un blanco; las “medidas” precisas son los disparos que dan cerca del centro. Los errores aleatorios están causados por cualquier cosa que haga que los disparos den en puntos diferentes de forma aleatoria. Por ejemplo, el tirador puede tener una mano poco firme, o puede haber condiciones atmosféricas fluctuantes entre el tirador y el blanco que distorsionen la vista del objetivo de una forma aleatoria. Los errores sistemáticos surgen si algo hace que los disparos no den en el centro y se alejen de él en una dirección “sistemática”, por ejemplo, si la mira del arma está mal alineada. Observe cómo en la figura 4.1 los resultados cambian según las diferentes combinaciones de errores aleatorios o sistemáticos pequeños o grandes.

Aunque la figura 4.1 es una ilustración excelente de los efectos de los errores aleatorios y sistemáticos, es, sin embargo, engañosa en un aspecto importante. Puesto que cada una de las cuatro imágenes muestra la posición del centro del blanco de tiro, podemos decir a simple vista si un disparo en particular fue cercano al centro del blanco o no. En particular, la diferencia entre las dos imágenes es evidente de inmediato. Los disparos en la imagen de la izquierda están agrupados alrededor del centro del blanco, mientras que los de la imagen de la derecha se agrupan en torno a un punto claramente descentrado; por tanto, el tirador responsable de la imagen izquierda tenía poco error sistemático, mientras que el responsable de la imagen derecha tenía mucho más. El conocer la posición del centro del blanco en la figura 4.1 es equivalente, en una medición de laboratorio, a conocer el verdadero valor de la magnitud medida, y en la gran mayoría de las mediciones reales, *no* conocemos el verdadero valor. (Si conociéramos el verdadero valor, por lo general no nos molestaríamos en medirlo).

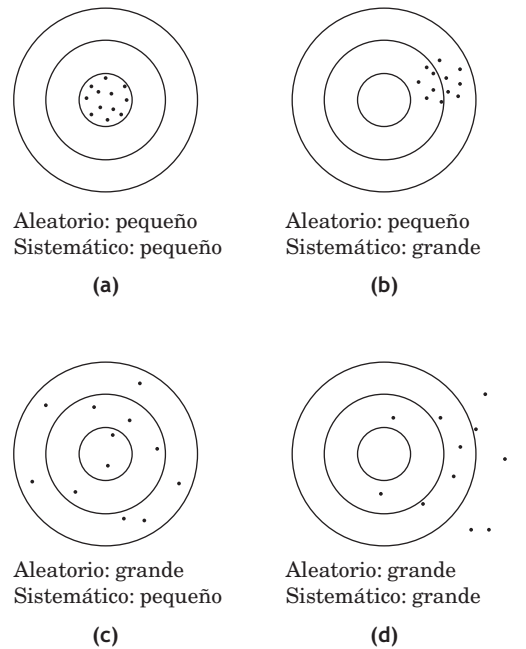


Figura 4.1. Errores aleatorios y sistemáticos en las prácticas de tiro. **(a)** Puesto que todos los disparos impactaron cerca unos de otros, podemos decir que los errores aleatorios son pequeños. Como la distribución de los disparos está situada en el centro de la diana, los errores sistemáticos son también pequeños. **(b)** Los errores aleatorios son todavía pequeños, pero los sistemáticos son mucho más grandes, pues los disparos están “sistemáticamente” situados hacia la derecha. **(c)** Aquí los errores aleatorios son grandes, pero los sistemáticos son pequeños ya que los tiros están muy dispersos pero no sistemáticamente descentrados. **(d)** En este caso, tanto los errores aleatorios como los sistemáticos son grandes.

Para mejorar la analogía de la figura 4.1 con experimentos más reales, tenemos que volver a dibujarla sin los anillos que muestran la posición del centro del blanco, como en la figura 4.2. En estas imágenes, la identificación de los errores aleatorios es todavía fácil. (Las dos primeras imágenes tienen claramente errores aleatorios menores que los de las dos inferiores). Sin embargo, a partir de la figura 4.2 es *imposible* determinar qué tirador tuvo los mayores errores sistemáticos. Esta situación es exactamente la que se da en la mayoría de los experimentos reales; examinando la distribución de los valores medidos podemos fácilmente evaluar los errores aleatorios, pero no obtenemos ninguna orientación sobre los errores sistemáticos.

La distinción entre errores aleatorios y sistemáticos no siempre es clara; así, un problema que provoca errores aleatorios en un experimento puede producir errores sistemáticos en otro. Por ejemplo, si colocamos nuestra cabeza primero a un lado y luego a otro para leer un medidor típico (tal como un reloj ordinario), la lectura en el medidor variará. Este efecto, llamado *paralaje*, significa que un dispositivo de medida sólo puede ser leído correctamente si nos colocamos directamente en frente de él. No importa lo cuidadoso que seamos, no siempre

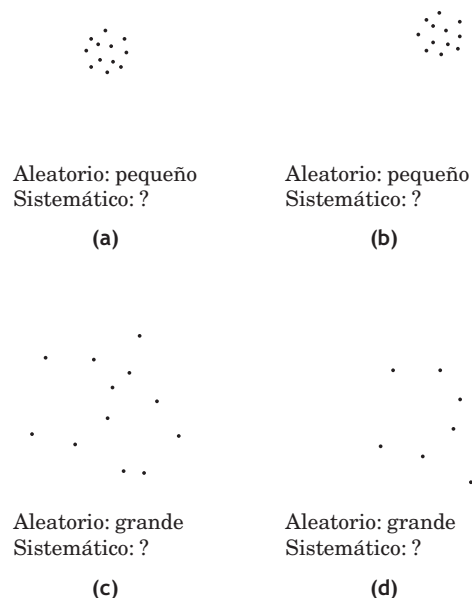


Figura 4.2. El mismo experimento que en la figura 4.1 pero dibujado sin mostrar la posición del centro del blanco. Esta situación se corresponde estrechamente con la de la mayoría de los experimentos reales, en los que no conocemos el verdadero valor de la magnitud medida. Aquí podemos evaluar los errores aleatorios con facilidad, pero no podemos decir nada sobre los sistemáticos.

podremos colocar el ojo *exactamente* en frente del medidor; en consecuencia, sus mediciones tendrán una pequeña incertidumbre debido al paralaje, y esta incertidumbre probablemente será aleatoria. Por otro lado, un experimentador poco cuidadoso que coloque un medidor a un lado de su asiento y olvide preocuparse del paralaje, introducirá un error sistemático en todas sus lecturas. Por lo tanto, el mismo efecto de paralaje puede producir incertidumbres aleatorias en un caso e incertidumbres sistemáticas en otro.

El tratamiento de los errores aleatorios es distinto del de los errores sistemáticos. Los métodos estadísticos descritos en las secciones siguientes aportan una estimación fiable de las incertidumbres aleatorias y, como veremos, proporcionan un procedimiento bien definido para reducirlas. Por las razones que acabamos de mencionar, las incertidumbres sistemáticas suelen ser difíciles de evaluar e incluso de detectar. El científico experto tiene que aprender a anticipar las posibles fuentes de error sistemático y asegurarse de que todos los errores sistemáticos son mucho menores que la precisión requerida. Hacerlo implicará, por ejemplo, comprobar los medidores en situaciones estándar y corregirlos o comprar otros mejores si es necesario. Desafortunadamente, en laboratorios de un primer año de física, tales controles son raramente posibles, por lo que el tratamiento de los errores sistemáticos es a menudo difícil. Todo ello se discutirá en la sección 4.6. De momento, consideraremos experimentos en los que todas las fuentes de error sistemático han sido identificadas y son mucho más pequeñas que la precisión requerida.

4.2 La media y la desviación estándar

Supongamos que necesitamos determinar el valor de cierta variable x , y que hemos identificado todas las fuentes de error sistemático y las hemos reducido a un nivel insignificante. Como todas las demás fuentes de incertidumbre son aleatorias, debemos ser capaces de detectarlas repitiendo la medición varias veces. Podríamos, por ejemplo, repetirla cinco veces y encontrar los resultados

$$71, 72, 72, 73, 71 \quad (4.1)$$

(donde, por conveniencia, se han omitido todas las unidades).

La primera pregunta que nos hacemos es la siguiente: teniendo en cuenta los cinco valores obtenidos (4.1), ¿cuál debemos tomar como nuestra mejor estimación x_{mejor} de la variable x ? Parece razonable dar como mejor estimación el *promedio* o la *media* \bar{x} de los cinco valores obtenidos, y en el capítulo 5 demostraremos que esta elección es normalmente la mejor. Así,

$$\begin{aligned} x_{\text{mejor}} &= \bar{x} \\ &= \frac{71 + 72 + 72 + 73 + 71}{5} = 71,8 \end{aligned} \quad (4.2)$$

Aquí, la segunda línea es simplemente la definición de la media \bar{x} para los datos en concreto que disponemos.¹

De manera más general, suponga que hacemos N determinaciones de la variable x (todas ellas con el mismo equipo y procedimientos) y encontramos los N valores siguientes

$$x_1, x_2, \dots, x_N \quad (4.3)$$

Una vez más, la mejor estimación de x es normalmente la media de x_1, \dots, x_n . Esto es,

$$x_{\text{mejor}} = \bar{x}, \quad (4.4)$$

donde

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{\sum x_i}{N}. \quad (4.5)$$

En la última expresión he introducido el sumatorio como notación útil, según el cual

$$\sum_{i=1}^N x_i = \sum_i x_i = \sum x_i = x_1 + x_2 + \dots + x_N$$

las expresiones segunda y tercera anteriores son abreviaturas comunes, que usaremos cuando no exista ningún peligro de confusión.

1. En esta era de las calculadoras de bolsillo, vale la pena señalar que una media como la indicada en (4.2) se calcula fácilmente mentalmente. Debido a que todos los números están en los setenta, lo mismo debe ser cierto para la media. Todo lo que queda es hacer un promedio de los números 1, 2, 2, 3, 1 en el lugar de las unidades. Estas cifras obviamente promedian a $9/5 = 1,8$, y nuestra respuesta es $x = 71,8$.

Capítulo 5

La distribución normal

Este capítulo continúa nuestra discusión del análisis estadístico de los datos resultantes de mediciones repetidas. En el capítulo 4 se introdujeron las importantes ideas de la media, la desviación estándar y la desviación estándar de la media; vimos su significado y alguno de sus usos. Este capítulo proporciona el fundamento teórico de estas ideas estadísticas y da la justificación de varios resultados enunciados sin demostración en capítulos anteriores. Al tratar las mediciones que se repiten muchas veces, el primer problema es encontrar una manera adecuada de manejar y mostrar los valores obtenidos. Un método conveniente consiste en utilizar una *distribución* o *histograma*, como se describe en la sección 5.1. La sección 5.2 introduce la noción de *distribución límite*, la distribución de los resultados que se obtendría si el número de mediciones fuese infinitamente grande. En la sección 5.3 definiremos la *distribución normal*, o *distribución de Gauss*, que es la distribución límite de los resultados de mediciones sujetas a muchos errores aleatorios pequeños.

Una vez que las propiedades matemáticas de la distribución normal se entienden, podemos proceder a probar varios resultados importantes con bastante facilidad. La sección 5.4 proporciona pruebas de que, tal como se había anticipado en el capítulo 4, alrededor del 68% de todas las mediciones (de una misma magnitud y todas obtenidas a partir de una misma técnica) deberían estar dentro de un intervalo en el que no se alejaran en más de una desviación estándar del valor verdadero de dicha magnitud. La sección 5.5 demuestra el resultado, utilizado anteriormente en el capítulo 1, de que si hacemos N mediciones x_1, x_2, \dots, x_N de alguna variable x , entonces nuestra mejor estimación x_{mejor} basada en estos valores es la media $\bar{x} = \Sigma x_i / N$. La sección 5.6 justifica el uso de la suma cuadrática en la propagación de errores que son independientes y aleatorios. En la sección 5.7 demostraremos que la incertidumbre de la media \bar{x} , cuando se la utiliza como mejor estimación de x , viene dada por la desviación estándar de la media $\sigma_{\bar{x}} = \sigma_x / \sqrt{N}$, tal como se ha indicado en el capítulo 4. Por último, la sección 5.8 explica cómo asignar un nivel de confianza numérico a los resultados experimentales.

Las matemáticas de este capítulo son más avanzadas que las utilizadas hasta el momento. En particular, el lector tendrá que entender las ideas básicas de integración: la integral como el área bajo la gráfica, cambios de variables y (ocasionalmente) la integración por partes. Sin embargo, una vez que haya estudiado la sección 5.3 sobre la distribución normal (repasando los cálculos con lápiz y papel, si es necesario), debería ser capaz de seguir el resto del capítulo sin mucha dificultad.

5.1 Histogramas y distribuciones

Debe quedar claro que un análisis estadístico serio de un experimento nos obliga a realizar muchas determinaciones o mediciones del valor de una misma variable. Por tanto, primero tenemos que idear métodos para registrar y presentar los resultados de un gran número de determinaciones. Por ejemplo, supongamos que tuviéramos que hacer 10 mediciones de cierta longitud x . Por ejemplo, x puede ser la distancia de una lente a la imagen formada por ésta. Podríamos obtener los valores (todos en cm)

$$26, 24, 26, 28, 23, 24, 25, 24, 26, 25. \quad (5.1)$$

Escritos esta manera, estos 10 números transmiten muy poca información, y si tuviéramos que registrar un número mucho mayor de ellos, el resultado sería una selva confusa de números. Obviamente, hace falta un mejor sistema de presentación de resultados.

Como primer paso, podemos reorganizar los números (5.1) en orden ascendente,

$$23, 24, 24, 24, 25, 25, 26, 26, 26, 28. \quad (5.2)$$

A continuación, en lugar de registrar las tres lecturas de 24, 24, 24, simplemente podemos registrar que obtuvimos el valor 24 tres veces; en otras palabras, podemos registrar los *diferentes* valores de x obtenidos, junto con el *número* de veces (frecuencia) en que se obtuvo cada valor, como en la tabla 5.1.

Tabla 5.1. Resultados de longitudes de x distintas y su correspondiente frecuencia.

Valores distintos obtenidos, x_k	23	24	25	26	27	28
Frecuencia correspondiente, n_k	1	3	2	3	0	1

Aquí hemos introducido la notación x_k ($k = 1, 2, \dots$) para indicar los valores diferentes que se han obtenido: $x_1 = 23$, $x_2 = 24$, $x_3 = 25$, y así sucesivamente. Y n_k ($k = 1, 2, \dots$) para indicar la frecuencia o número de veces que se encontró dicho valor x_k : $n_1 = 1$, $n_2 = 3$, y así sucesivamente.

Si registramos los resultados como en la tabla 5.1, podemos reescribir la definición de la media \bar{x} en una fórmula más conveniente. De la definición anterior, sabemos que

$$\bar{x} = \frac{\sum_i x_i}{N} = \frac{23 + 24 + 24 + 24 + 25 + \dots + 28}{10}. \quad (5.3)$$

Esta ecuación es la misma que

$$\bar{x} = \frac{23 + (24 \times 3) + (25 \times 2) + \dots + 28}{10}$$

o en general

$$\bar{x} = \frac{\sum_k x_k n_k}{N}. \quad (5.4)$$

En la forma original (5.3), sumamos *todos* los resultados obtenidos; en (5.4) sumamos todos los valores *diferentes* obtenidos, multiplicados cada uno de ellos por el número de veces que apareció. Estas dos sumas son, evidentemente, la misma, pero la forma (5.4) es más útil cuando tenemos muchos resultados. Una suma como la expresada en (5.4) se denomina a veces *suma ponderada*; cada valor x_k se *pondera* por el número de veces que apareció, n_k . Para futuras referencias, observemos que si sumamos todos los números n_k , obtenemos el número total de resultados obtenidos, N . Esto es,

$$\sum_k n_k = N. \quad (5.5)$$

(Por ejemplo, en la tabla 5.1 esta ecuación afirma que la suma de los números en la última línea es 10).

Test rápido 5.1. En sus dos primeros años de universidad, José cursó 20 asignaturas (todas con el mismo número de créditos) y obtuvo 7 A, 4 B, 7 C y 2 E. A los efectos de calcular el promedio de calificaciones, a cada letra se le asigna como de costumbre la puntuación numérica siguiente:

Letra:	F	D	C	B	A
Puntuación, s_k :	$s_1 = 0$	$s_2 = 1$	$s_3 = 2$	$s_4 = 3$	$s_5 = 4$

Cree una tabla como la tabla 5.1 mostrando las posibles diferentes calificaciones s_k y el número de veces n_k que se obtuvieron. Utilice la ecuación (5.4) para calcular el promedio de calificaciones de José.

Las ideas de los dos últimos párrafos pueden reformularse de una manera que a menudo es más conveniente. En lugar de decir que el resultado $x = 24$ se obtuvo en tres ocasiones, podemos decir que $x = 24$ se obtuvo en 3/10 de todas nuestras mediciones. En otras palabras, en lugar de utilizar n_k , la frecuencia o el *número* de veces que el resultado x_k se repitió, introducimos la fracción

$$F_k = \frac{n_k}{N}, \quad (5.6)$$

que es la *frecuencia relativa*, o *tanto por uno*, con que se presenta el resultado x_k entre el total N . Se dice que las frecuencias relativas F_k especifican la *distribución* de nuestros resultados, ya que describen cómo se *distribuyeron* entre los distintos valores posibles.

En términos de las frecuencias relativas F_k , podemos volver a escribir la fórmula (5.4) para la media \bar{x} en la forma compacta

$$\bar{x} = \sum_k x_k F_k.$$
(5.7)

Es decir, la media \bar{x} es la suma ponderada de todos los diferentes valores x_k obtenidos, con cada x_k ponderado por su frecuencia relativa, F_k .

El resultado (5.5) implica que

$$\sum_k F_k = 1. \quad (5.8)$$

Es decir, si sumamos las frecuencias relativas F_k para todos los resultados posibles x_k , debemos obtener 1. Cualquier conjunto de números cuya suma sea 1 se dice que está *normalizado*, y es por ello que la relación (5.8) se llama *condición de normalización*.

La distribución de nuestras mediciones se puede mostrar gráficamente en un *histograma*, como en la figura 5.1. Esta figura es sólo un gráfico de F_k respecto a x_k , en la que los diferentes resultados x_k se representan en el eje horizontal, mientras que la frecuencia relativa de cada uno de ellos se indica por la altura de la barra vertical dibujada encima de x_k . (También podemos representar n_k respecto a x_k , pero para nuestros propósitos el gráfico de F_k respecto a x_k es más conveniente). Los datos que se muestran en histogramas como éste pueden ser comprendidos de forma rápida y sencilla, como saben muchos periodistas que escriben en periódicos y revistas.

Un histograma como el de la figura 5.1 puede llamarse *histograma de barras* porque la distribución de los resultados está indicada por la altura de las barras verticales situadas encima de cada x_k . Este tipo de histograma es adecuado cuando los valores x_k están perfectamente espaciados, con valores enteros. (Por ejemplo, las puntuaciones de los estudiantes en un examen son generalmente números enteros y se visualizan convenientemente usando un histograma de barras). Sin embargo, la mayoría de las mediciones no proporcionan resultados ordenados de valores enteros, porque la mayoría de las variables físicas tienen un recorrido continuo de valores posibles. Por ejemplo, en lugar de las 10 longitudes indicadas en la ecuación (5.1), es mucho más probable obtener 10 valores como

$$26,4, 23,9, 25,1, 24,6, 22,7, 23,8, 25,1, 23,9, 25,3, 25,4. \quad (5.9)$$

Un histograma de barras de estos 10 valores consistiría en 10 barras separadas, todas de la misma altura, y aportarían relativamente poca información. Tenien-

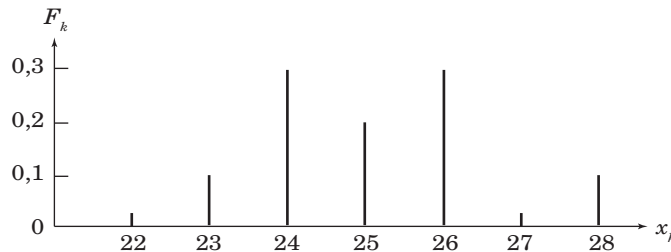


Figura 5.1. Histograma de las 10 mediciones de la longitud x . El eje vertical muestra la frecuencia relativa F_k con que cada valor x_k fue observado.

do en cuenta mediciones como las indicadas en (5.9), el mejor camino es dividir el rango de valores en un número conveniente de *intervalos* o “clases”, y contar cuántos valores caen dentro de cada “clase”. Por ejemplo, podríamos contar el número de resultados (5.9) entre $x = 22$ y 23 , entre $x = 23$ y 24 , y así sucesivamente. Los resultados de contar de esta manera se muestran en la tabla 5.2. (Si un determinado valor cae exactamente en el límite entre dos clases, hay que decidir dónde colocarlo. Una manera sencilla y razonable es asignar medio resultado a cada una de las clases).

Tabla 5.2. Los 10 resultados (5.9) agrupados en clases.

Clases	22 a 23	23 a 24	24 a 25	25 a 26	26 a 27	27 a 28
Frecuencia relativa de cada clase	1	3	1	4	1	0

Los resultados de la tabla 5.2 se pueden representar en una forma que podemos llamar *histograma de clases* como se muestra en la figura 5.2. En este gráfico, la frecuencia relativa de los resultados que pertenecen a cada clase se indica por el área del rectángulo dibujado encima de dicha clase. Por lo tanto, el rectángulo sombreado encima del intervalo de $x = 23$ a $x = 24$ tiene un área de $0,3 \times 1 = 0,3$, lo que indica que $3/10$ de todos los resultados pertenecen a este intervalo. En general, denotamos la anchura de la k -ésima clase por Δ_k . (Estas anchuras son generalmente todas iguales, aunque ciertamente no tienen por qué serlo). La altura f_k del rectángulo dibujado sobre esta clase se elige de manera que el área $f_k \Delta_k$ sea

$$f_k \Delta_k = \text{frecuencia relativa de resultados en la } k\text{-ésima clase.}$$

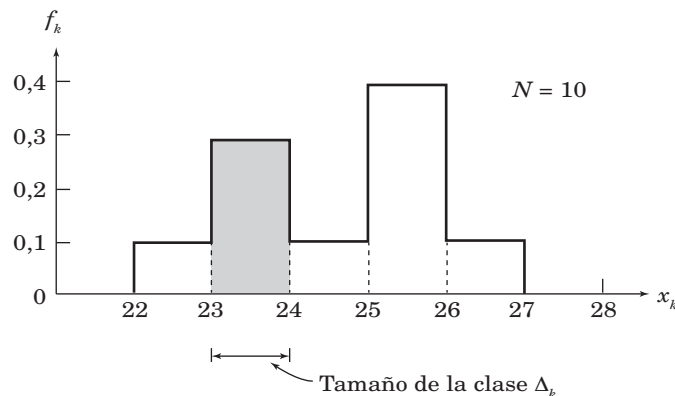


Figura 5.2. Histograma de clases que muestra la frecuencia relativa de los 10 resultados (5.9) de x que están en las “clases” 22 a 23, 23 a 24, y así sucesivamente. El área del rectángulo situado encima de cada intervalo es igual a la frecuencia relativa de los resultados que están en dicho intervalo. Así, el área del rectángulo sombreado es $0,3$, lo que indica que $3/10$ del total de los resultados se encuentran entre 23 y 24.

En otras palabras, en un histograma de clases el área $f_k \Delta_k$ del rectángulo k -ésimo tiene el mismo significado que la altura F_k de la barra k -ésima en un histograma de barras.

Se necesita tener algún cuidado en la elección de la anchura Δ_k de las clases en un histograma. Si las clases se hacen demasiado anchas, entonces todas las lecturas (o casi todas) caerán en una misma clase, y el histograma consistirá en un único rectángulo sin interés. Si las clases se hacen demasiado estrechas, entonces pocas de ellas contendrán más de un resultado, y el histograma consistirá en numerosos rectángulos estrechos, casi todos de la misma altura.

Claramente, la anchura de la clase debe elegirse de modo que cada una de las diferentes clases incluya varios resultados. Así, cuando el número total de resultados N es pequeño, tenemos que elegir nuestras clases relativamente amplias, pero si aumentamos N , entonces podemos elegir generalmente clases más estrechas.

Test rápido 5.2. Veinte estudiantes realiza un examen que se califica sobre 50 puntos. Los estudiantes obtienen las siguientes calificaciones:

26, 33, 38, 41, 49, 28, 36, 38, 47, 41,

32, 37, 48, 44, 27, 32, 34, 44, 37, 30

(Estos resultados se tomaron de un listado de alumnos ordenado alfabéticamente). En una hoja de papel cuadriculado, dibuje un histograma de clases a partir de las puntuaciones, tomando los límites de las clases en 25, 30, 35, 40, 45 y 50. Etiquete la escala vertical para que el área de cada rectángulo coincida con la frecuencia relativa de estudiantes cuya nota está en dicha clase.

5.2 Distribuciones límite

En la mayoría de los experimentos, a medida que se aumenta el número de resultados, el histograma empieza a tomar una forma definida y simple. Esta evolución es claramente visible en las figuras 5.3 y 5.4, que muestran 100 y 1000 mediciones de la misma magnitud que en la figura 5.2. Después de 100 resultados, el histograma muestra un solo pico, aproximadamente simétrico. Después de 1000 resultados, hemos podido reducir a la mitad la anchura de la clase, y el histograma se ha vuelto bastante suave y regular. Estos tres gráficos ilustran una propiedad importante de la mayoría de las mediciones. A medida que el número de resultados se aproxima a infinito, su distribución se asemeja cada vez más a una determinada curva continua. Cuando esto sucede, dicha curva continua define lo que denominaremos la *distribución límite*.¹ Así, para las mediciones de las figuras 5.2 a 5.4, la distribución límite parece estar próxima a la curva acampanada simétrica superpuesta en la figura 5.4.

1. Algunos sinónimos comunes (o sinónimos aproximados) para la distribución límite son: distribución poblacional, distribución poblacional infinita, distribución asintótica.

Parte II

1. Exclusión de datos
 2. Medias ponderadas
 3. Ajuste mínimo cuadrático
 4. Covarianza y correlación
 5. La distribución binomial
 6. La distribución de Poisson
 7. La prueba χ^2 -cuadrado de ajuste a una distribución
-

Si el lector ha leído y comprendido el capítulo 5, ahora está preparado, sorprendentemente con pocas dificultades adicionales, para estudiar una serie de temas más avanzados. En los capítulos de esta segunda parte se estudian siete temas, algunos de los cuales son aplicaciones de la teoría estadística ya desarrollada, mientras que otros son extensiones adicionales de dicha teoría. Todos son importantes, y es probable que el lector los necesite estudiar tarde o temprano. Sin embargo, es posible que no quiera aprenderlos todos a la vez, y por este motivo se han organizado en capítulos cortos e independientes, que pueden estudiarse en cualquier orden, conforme a sus necesidades e intereses.

Capítulo 6

Exclusión de datos

Este capítulo trata la incómoda pregunta de si se debe descartar un resultado que parece tan poco razonable que parece un error. Este es un tema controvertido; algunos científicos argumentan que desechar un resultado sólo porque parece ilógico *nunca* está justificado. Sin embargo, hay un sencillo test que el lector podría considerar aplicar si se encuentra frente a esta disyuntiva. La prueba se llama *criterio de Chauvenet* y es una buena aplicación de las ideas estadísticas desarrolladas en los capítulos 4 y 5.

6.1 El problema de excluir datos

A veces, en una serie de resultados un dato parece estar en total desacuerdo con los demás. Cuando esto sucede, el experimentador debe decidir si dicho resultado anómalo fue consecuencia de un error, y debe ser rechazado, o fue obtenido de forma *fiable*, y debe ser utilizado con todos los demás. Por ejemplo, imagine que realizamos seis mediciones del período de un péndulo y obtenemos los resultados siguientes (todos en segundos)

$$3,8; 3,5; 3,9; 3,9; 3,4; 1,8 \quad (6.1)$$

En este ejemplo, el valor 1,8 es sorprendentemente diferente de todos los demás, y tenemos que decidir qué hacer con él.

Sabemos por el capítulo 5 que un resultado legítimo *puede* desviarse significativamente de otras mediciones de la misma cantidad. Sin embargo, una discrepancia legítima tan grande como la que corresponde a la última medición en (6.1) es *muy improbable*, por lo que nos inclinamos a sospechar que el tiempo de 1,8 s se dio como resultado de un error no detectado u otra causa externa. Tal vez, por ejemplo, simplemente leímos mal el último tiempo o nuestro cronómetro eléctrico se detuvo brevemente durante la última medición a causa de un fallo momentáneo de alimentación.

Si hemos detallado con mucho cuidado el proceso de observación, a veces podremos ser capaces de establecer una causa definitiva para dicha medición anómala. Por ejemplo, nuestros datos podrían indicar que se utilizó un cronómetro distinto para la última medición (6.1), y un control posterior podría mostrar que este reloj atrasa. En este caso, la medición anómala debe ser rechazada definitivamente.

Desafortunadamente, establecer una causa externa que explique un resultado anómalo no es generalmente posible. Debemos decidir entonces si rechazar o no dicho resultado anómalo simplemente mediante el examen de los resultados por sí solos, y aquí nuestro conocimiento de la distribución de Gauss resulta muy útil.

El rechazo de los datos es una cuestión controvertida en la que los expertos no están de acuerdo. También es una cuestión *importante*. En nuestro ejemplo, la mejor estimación para el período del péndulo se ve significativamente afectada si rechazamos el valor sospechoso de 1,8 s. El promedio de las seis mediciones es 3,4 s, mientras que el de las primeras cinco es 3,7 s, una diferencia apreciable.

Por otra parte, la decisión de rechazar los datos es, en última instancia, subjetiva, y el científico que toma esta decisión puede hasta cierto punto ser acusado por otros científicos de “arreglar” los datos. La situación se ve agravada al considerar la posibilidad de que el resultado anómalo puede reflejar algún aspecto importante del fenómeno estudiado. De hecho, muchos descubrimientos científicos importantes aparecieron por primera vez como mediciones anómalas que parecían errores. Al desechar el valor de 1,8 s en el ejemplo (6.1), *podríamos* estar desechando el dato más interesante.

En realidad, frente a datos como los presentados en (6.1), el único camino realmente honesto es repetir la medición muchas, muchas veces. Si la anomalía aparece de nuevo, presumiblemente seremos capaces de rastrear su causa, ya sea esta debida a un error o a un efecto físico real. Si no aparece, entonces en el momento en que hayamos hecho, por ejemplo, 100 mediciones, no habrá ninguna diferencia significativa en nuestra respuesta final, incluyamos o no el dato anómalo.

Sin embargo, repetir una medición 100 veces cada vez que un resultado parece sospechoso suele ser poco práctico (especialmente en un laboratorio con fines docentes). Por tanto, necesitamos algún criterio para rechazar un resultado sospechoso. Existen diversos criterios, algunos bastante complicados. El criterio de Chauvenet proporciona una aplicación simple e instructiva de la distribución de Gauss.

6.2 El criterio de Chauvenet

Volvamos a las seis mediciones del ejemplo (6.1):

$$3,8; 3,5; 3,9; 3,9; 3,4; 1,8.$$

Si aceptamos por el momento que se trata de seis determinaciones válidas de una cierta cantidad x , podemos calcular la media \bar{x} y la desviación estándar σ_x ,

$$\bar{x} = 3,4 \text{ s} \tag{6.2}$$

y

$$\sigma_x = 0,8 \text{ s} \tag{6.3}$$

Ahora podemos cuantificar el grado en que la medición sospechosa, 1,8, resulta anómala. Difiere en 1,6 de la media 3,4, es decir, en dos desviaciones estándar. Si suponemos que las mediciones se rigen por una distribución de Gauss con centro y anchura dados por (6.2) y (6.3), podemos calcular la probabilidad de obtener

mediciones que difieran de la media al menos en esta cantidad. De acuerdo con las probabilidades que se muestran en el apéndice A, esta probabilidad es

$$\begin{aligned}\text{Prob}(\text{desviación mayor que } 2\sigma) &= 1 - \text{Prob}(\text{desviación menor que } 2\sigma) \\ &= 1 - 0,95 \\ &= 0,05.\end{aligned}$$

En otras palabras, suponiendo que los valores (6.2) y (6.3) para \bar{x} y σ_x son correctos, esperaríamos que una de cada 20 mediciones difiriera de la media por lo menos tanto como lo hace el sospechoso resultado de 1,8 s. Si hubiéramos hecho 20 o más mediciones, en realidad deberíamos *esperar* obtener una o dos mediciones tan desviadas como ésta, y no tendríamos ninguna razón para rechazarla. Pero sólo hemos efectuado seis mediciones, por lo que el número esperado de mediciones tan desviadas como la anterior es en realidad

$$\begin{aligned}(\text{número esperado de resultados tan desviado al menos como 1,8 s}) \\ &= (\text{número de mediciones}) \times \text{Prob}(\text{fuera de } 2\sigma) \\ &= 6 \times 0,05 = 0,3.\end{aligned}$$

Es decir, en seis mediciones deberíamos esperar (en promedio) sólo la tercera parte de una medición tan desviada como el valor sospechoso 1,8 s.

Este resultado nos proporciona una medida cuantitativa de la “razonabilidad” de nuestro resultado sospechoso. Si optamos por considerar a la tercera parte de una medición como “absurdamente improbable”, entonces llegaremos a la conclusión de que el valor 1,8 s no es un resultado válido y debe ser rechazado.

La decisión de dónde establecer el límite de la “improbabilidad absurda” corresponde al experimentador. El criterio de Chauvenet, en su versión habitual, establece que si el número esperado de mediciones que sean por lo menos tan desviadas como la medición sospechosa es menor que media unidad, entonces la medición sospechosa debe ser rechazada. Obviamente, la elección de media unidad es arbitraria, pero también es razonable y puede ser defendida.

La aplicación del criterio Chauvenet a un problema general se puede ahora describir fácilmente. Suponga que ha realizado N determinaciones

$$x_1, \dots, x_N$$

de una cierta magnitud x . De todas las N determinaciones, calcule \bar{x} y σ_x . Si una de las mediciones (llamémosla x_{sosp}) difiere de \bar{x} tanto que parece sospechosa, encuentre entonces

$$t_{\text{sosp}} = \frac{|x_{\text{sosp}} - \bar{x}|}{\sigma_x}, \quad (6.4)$$

el número de desviaciones estándar para las cuales x_{sosp} difiere de \bar{x} . A continuación, del apéndice A se puede calcular la probabilidad

$$\text{Prob}(\text{desviación mayor que } t_{\text{sosp}}\sigma)$$

de que una medición válida diferiría de \bar{x} en t_{sosp} o más desviaciones estándar. Finalmente, multiplicando por N , el número total de mediciones da

$$\begin{aligned}n &= (\text{número esperado de resultados tan o más desviados que } x_{\text{sosp}}) \\ &= N \times \text{Prob}(\text{desviación mayor que } t_{\text{sosp}}\sigma).\end{aligned}$$

Si este número esperado n es menor que un medio, entonces, según el criterio de Chauvenet, se puede rechazar x_{sosp} .

Si se decide rechazar x_{sosp} , naturalmente se debe volver a calcular \bar{x} y σ_x utilizando sólo los datos restantes; en particular, su respuesta final para x sería esta nueva media, con una incertidumbre igual a la nueva desviación estándar de la media.

Ejemplo: Diez mediciones de la longitud de un objeto

Un estudiante hace 10 mediciones de la longitud x de un objeto y obtiene los siguientes resultados (todos en mm):

46, 48, 44, 38, 45, 47, 58, 44, 45, 43.

Al darse cuenta de que el valor 58 parece anormalmente grande, revisa sus registros, pero no puede encontrar ninguna evidencia de que el resultado se deba a un error. Por lo tanto, aplica el criterio de Chauvenet. ¿Cuál será su conclusión?

Aceptando provisionalmente las 10 mediciones, calcula

$$\bar{x} = 45,8 \quad \text{y} \quad \sigma_x = 5,1.$$

La diferencia entre el valor sospechoso $x_{\text{sosp}} = 58$ y la media $\bar{x} = 45,8$ es 12,2, o 2,4 desviaciones estándar; esto es,

$$t_{\text{sosp}} = \frac{x_{\text{sosp}} - \bar{x}}{\sigma_x} = \frac{58 - 45,8}{5,1} = 2,4.$$

En referencia a la tabla del apéndice A, puede verse que la probabilidad de que una medición difiera de \bar{x} en $2,4\sigma_x$ o más es

$$\begin{aligned} \text{Prob}(\text{desviación mayor que } 2,4\sigma) &= 1 - \text{Prob}(\text{desviación menor que } 2,4\sigma) \\ &= 1 - 0,984 \\ &= 0,016. \end{aligned}$$

En 10 mediciones, por lo tanto, el número esperado de mediciones tan desviadas como su resultado sospechoso es 0,16. Como 0,16 es menor que el número 0,5 establecido por el criterio de Chauvenet, al menos debería considerar la posibilidad de rechazar el resultado.

Si decide rechazar el resultado sospechoso de valor igual a 58, entonces debe recalcular \bar{x} y σ_x como

$$\bar{x} = 44,4 \quad \text{y} \quad \sigma_x = 2,9.$$

Como era de esperar, su media cambia un poco y su desviación estándar disminuye sensiblemente.

Test rápido 6.1. Un estudiante mide 20 veces un voltaje V y calcula su media y su desviación estándar obteniendo $\bar{V} = 51$ y $\sigma_V = 2$ (ambas en microvoltios). La noche anterior al día que debía presentar el trabajo, comenzó a escribir su informe y se dio cuenta de que uno de sus valores medidos fue $V_{\text{sosp}} = 56$. ¿Cuál es la probabilidad de obtener una medición tan desviada de \bar{V} ? Si decide utilizar el criterio de Chauvenet, ¿debería rechazar el valor sospechoso?

6.3 Discusión

El lector debe ser consciente de que algunos científicos creen que los datos nunca deben ser rechazados sin una prueba *externa* de que la medición en cuestión no es correcta. Un compromiso razonable es utilizar el criterio de Chauvenet para identificar los datos que se podrían *considerar* rechazables; una vez hecha esta identificación, todos los cálculos posteriores se pueden hacer dos veces, una vez incluyendo los datos sospechosos y otra excluyéndolos, para ver hasta qué punto dichos valores cuestionables afectan a la conclusión final.

Una razón por la que muchos científicos se sienten incómodos con el criterio de Chauvenet es que la elección de un medio como el valor límite de rechazo (en la condición de que $n < \frac{1}{2}$ es arbitraria. Quizás aún más importante, a menos que haya realizado un gran número de mediciones ($N \approx 50$, por ejemplo), el valor de σ_x es extremadamente incierto como estimación de la verdadera desviación estándar de las mediciones. (Vea el problema 6.7 para un ejemplo). Esto significa a su vez que el número t_{sosp} en (6.4) es muy poco preciso. Debido a que la probabilidad de que una medición difiera en más de t desviaciones estándar del verdadero valor es muy sensible a t , un gran error en t_{sosp} causa un error muy grande en la estimación de dicha probabilidad y plantea serias dudas sobre todo el procedimiento. Por ambas razones, el criterio de Chauvenet sólo debe utilizarse como último recurso, cuando no podemos comprobar nuestras mediciones repitiéndolas.

Hasta ahora, hemos supuesto que sólo una medición es sospechosa. ¿Qué se debe hacer si son varias? Dado que el uso del criterio de Chauvenet de rechazar *una* medición es cuestionable, claramente su uso para eliminar *varias* mediciones es aún más problemático. Sin embargo, si no hay absolutamente ninguna manera de que el lector pueda repetir sus mediciones (porque ha desmantelado temerariamente su equipo antes de escribir su informe, por ejemplo), puede que tenga que hacer frente a esta cuestión.

Suponga primero que tiene dos mediciones que se desvían de la media en una misma cantidad grande. En este caso, podemos calcular el número esperado de mediciones con desviación mayor o igual que la anterior, y si este número es menor que uno (es decir, *dos veces* un medio), entonces ambas mediciones podrían ser consideradas candidatas para el rechazo. Si tiene dos mediciones sospechosas, x_1 y x_2 , con x_2 más desviada que x_1 , primero debe aplicar el criterio de Chauvenet utilizando el valor x_1 . Si el número esperado desviado es menor que uno, podría rechazar *ambos* valores. Si este número esperado es mayor que uno, sin duda no debería rechazar los dos, sino que debería volver a aplicar el criterio de Chauvenet con x_2 y, si el número esperado desviados es inferior a un medio, puede rechazar sólo x_2 .

Después de haber rechazado cualquier medición que no cumpla el criterio de Chauvenet, naturalmente deberíamos volver a calcular \bar{x} y σ_x utilizando sólo los datos restantes. El valor resultante de σ_x será menor que el original, y con el nuevo σ_x , algunas mediciones más pueden fallar según el criterio de Chauvenet. Sin embargo, hay acuerdo generalizado en que el criterio de Chauvenet *no* debe ser aplicado por segunda vez con el nuevo cálculo de los valores de \bar{x} y σ_x .

Capítulo 7

Medias ponderadas

En este capítulo se aborda el problema de combinar dos o más determinaciones distintas e independientes de una misma magnitud física. Veremos que la mejor estimación de dicha magnitud, a partir de varias determinaciones, es una *media ponderada* apropiada de las mismas.

7.1 El problema de combinar mediciones diferentes

A menudo, una magnitud física se mide varias veces, tal vez en varios laboratorios independientes, y surge la cuestión de cómo se pueden combinar estas mediciones para dar una mejor estimación única. Suponga, por ejemplo, que dos estudiantes, A y B , miden una magnitud x cuidadosamente y obtienen estos resultados:

$$\text{estudiante } A: \quad x = x_A \pm \sigma_A \quad (7.1)$$

y

$$\text{estudiante } B: \quad x = x_B \pm \sigma_B. \quad (7.2)$$

Cada resultado probablemente será en sí mismo el resultado de varias mediciones, en cuyo caso x_A será la media de todas las mediciones de A y σ_A la desviación estándar de la media (y de manera similar para x_B y σ_B). La pregunta es ¿cuál es la mejor manera de combinar x_A y x_B para tener una nueva y mejor estimación de x ?

Antes de examinar esta cuestión, hemos de tener en cuenta que si la discrepancia $|x_A - x_B|$ entre ambas mediciones es mucho mayor que las dos incertidumbres σ_A y σ_B , deberíamos sospechar que algo ha ido mal en al menos una de aquellas. En tal caso, podríamos decir que las dos mediciones son *incompatibles*, y deberíamos examinar ambas con cuidado para ver si alguna (o las dos) fueron objeto de errores sistemáticos inadvertidos.

Supongamos, sin embargo, que las dos mediciones (7.1) y (7.2) son *compatibles*; es decir, la discrepancia $|x_A - x_B|$ no es significativamente mayor que σ_A y σ_B . Podemos entonces preguntarnos de forma sensata cuál es la mejor estimación x_{mejor} del valor verdadero X , sobre la base de ambas mediciones. El primer impulso podría ser el uso de la media $(x_A + x_B)/2$ de las dos estimaciones. Un análisis más profundo debería sugerir, sin embargo, que esta media no es ade-

cuada si las dos incertidumbres σ_A y σ_B son desiguales. El simple promedio $(x_A + x_B)/2$ da igual importancia a ambas mediciones, mientras que a la lectura más precisa debería dársele de alguna manera más peso.

A lo largo de este capítulo voy a suponer que todos los errores sistemáticos han sido identificados y reducidos a un nivel insignificante. Por lo tanto, todos los errores restantes son aleatorios y las mediciones de x se distribuyen normalmente alrededor del valor verdadero X .

7.2 La media ponderada

Podemos resolver nuestro problema fácilmente usando el principio de máxima verosimilitud, tal como lo hicimos en la sección 5.5. Estamos suponiendo que ambas mediciones se rigen por la distribución de Gauss y denotaremos el valor verdadero desconocido de x por X . Por lo tanto, la probabilidad de que el estudiante A obtenga su valor particular x_A es

$$\text{Prob}_X(x_A) \propto \frac{1}{\sigma_A} e^{-(x_A - X)^2 / 2\sigma_A^2}, \quad (7.3)$$

y la de B de obtener su observado x_B es

$$\text{Prob}_X(x_B) \propto \frac{1}{\sigma_B} e^{-(x_B - X)^2 / 2\sigma_B^2}. \quad (7.4)$$

El subíndice X indica explícitamente que estas probabilidades dependen del valor real desconocido.

La probabilidad de que A encuentre el valor x_A y B el valor x_B es simplemente el producto de las dos probabilidades (7.3) y (7.4). De una manera que ahora debe estar familiarizado, este producto implicará una función exponencial cuyo exponente es la suma de los dos exponentes en (7.3) y (7.4). Escribimos esto como

$$\begin{aligned} \text{Prob}_X(x_A, x_B) &= \text{Prob}_X(x_A) \text{Prob}_X(x_B) \\ &\propto \frac{1}{\sigma_A \sigma_B} e^{-\chi^2 / 2}, \end{aligned} \quad (7.5)$$

donde hemos introducido la conveniente notación abreviada χ^2 (ji-cuadrado) para el exponente

$$\chi^2 = \left(\frac{x_A - X}{\sigma_A} \right)^2 + \left(\frac{x_B - X}{\sigma_B} \right)^2 \quad (7.6)$$

Esta cantidad importante es la suma de los cuadrados de las desviaciones de X de las dos mediciones, cada uno dividido por su incertidumbre correspondiente.

El principio de máxima verosimilitud afirma, igual que antes, que nuestra mejor estimación del valor real desconocido X es aquel valor para el que las observaciones reales x_A, x_B son más probables. Esto es, la mejor estimación de X es el valor para el que la probabilidad (7.5) es máxima o, equivalentemente, el exponente χ^2 es mínimo. (Debido a que maximizar la probabilidad implica minimizar la “suma de los cuadrados” χ^2 , a este método para la estimación de X se le llama a veces el “método de los mínimos cuadrados”). Por lo tanto, para encontrar la mejor estimación,

simplemente derivamos (7.6) con respecto a X y formamos una ecuación igualando a cero la derivada,

$$2 \frac{x_A - X}{\sigma_A} + 2 \frac{x_B - X}{\sigma_B} = 0.$$

La solución de esta ecuación para X es nuestra mejor estimación y se ve fácilmente que es

$$(\text{mejor estimación de } X) = \left(\frac{x_A}{\sigma_A^2} + \frac{x_B}{\sigma_B^2} \right) / \left(\frac{1}{\sigma_A^2} + \frac{1}{\sigma_B^2} \right). \quad (7.7)$$

Este resultado bastante feo se puede hacer más atractivo si definimos los *pesos*

$$w_A = \frac{1}{\sigma_A^2} \quad \text{y} \quad w_B = \frac{1}{\sigma_B^2}. \quad (7.8)$$

Con esta notación podemos reescribir (7.7) como la *media ponderada* (denotada x_{mp})

$$(\text{mejor estimación de } X) = x_{\text{mp}} = \frac{w_A x_A + w_B x_B}{w_A + w_B}. \quad (7.9)$$

Si las dos mediciones originales tienen la misma incertidumbre ($\sigma_A = \sigma_B$ y por tanto $w_A = w_B$), este resultado se reduce a la simple media aritmética $(x_A + x_B)/2$. En general, cuando $w_A \neq w_B$, la media ponderada (7.9) *no* es igual a la media aritmética; es parecida a la fórmula de la posición del centro de gravedad de dos cuerpos, donde w_A y w_B son los pesos, y x_A y x_B sus posiciones. En (7.9), los “pesos” son el inverso de los cuadrados de las incertidumbres en las mediciones originales, como en (7.8). Si la medición de A es más precisa que B , entonces $\sigma_A < \sigma_B$ y por tanto $w_A > w_B$, entonces la mejor estimación x_{mejor} es más cercana a x_A que a x_B , tal como es razonable que sea.

Test rápido 7.1. Científicos de dos laboratorios determinaron independientemente la vida media de cierta partícula y obtuvieron $10,0 \pm 0,5$ y 12 ± 1 , ambas en nanosegundos. Si deciden combinar los resultados, ¿cuáles serán los respectivos pesos en la expresión (7.8) y la correspondiente media ponderada (7.9)?

Nuestro análisis basado en dos mediciones puede ser generalizado hasta considerar cualquier número de mediciones. Supongamos que tenemos N mediciones distintas de una misma magnitud x ,

$$x_1 \pm \sigma_1, \quad x_2 \pm \sigma_2, \dots, x_N \pm \sigma_N,$$

con sus correspondientes incertidumbres $\sigma_1, \sigma_2, \dots, \sigma_N$. Extendiendo el razonamiento anterior, encontraremos que la mejor estimación basada en estas mediciones viene dada por la media ponderada

$$x_{\text{mp}} = \frac{\sum w_i x_i}{\sum w_i}, \quad (7.10)$$

donde las sumas se extienden a las N mediciones, $i = 1, \dots, N$, y el *peso* w_i de cada medición es el inverso del cuadrado de su incertidumbre correspondiente,

$$w_i = \frac{1}{\sigma_i^2}, \quad (7.11)$$

para $i = 1, 2, \dots, N$.

Debido a que el peso $w_i = 1/\sigma_i^2$ asociado con cada medición implica el *cuadrado* de la incertidumbre correspondiente σ_i , cualquier medición *mucho* menos precisa que las demás contribuirá mucho menos al resultado final (7.10). Por ejemplo, si una medida es cuatro veces menos precisa que las demás, su peso correspondiente será 16 veces menor que los otros pesos, y para muchos propósitos esta medida podría simplemente ignorarse.

Puesto que el promedio ponderado x_{mp} es una función de los valores originales medidos x_1, x_2, \dots, x_N , la incertidumbre en x_{mp} se puede calcular utilizando la propagación de errores. Como se puede comprobar fácilmente (problema 7.8), la incertidumbre en x_{mp} es

$$\sigma_{\text{mp}} = \frac{1}{\sqrt{\sum w_i}}, \quad (7.12)$$

Este resultado bastante feo tal vez sea un poco más fácil de recordar si reescribimos (7.11) como

$$\sigma_i = \frac{1}{\sqrt{w_i}}. \quad (7.13)$$

Interpretando la ecuación (7.13), se puede decir que la incertidumbre en cada medición es igual a la raíz cuadrada recíproca de su correspondiente peso. Volviendo a la ecuación (7.12), podemos decir de manera similar que la incertidumbre en el resultado final x_{mp} es el recíproco de la raíz cuadrada de la *suma de todos los pesos individuales*, es decir, el peso total de la respuesta final es la suma de cada uno los pesos individuales w_i .

Test rápido 7.2. ¿Cuál es la incertidumbre de la respuesta final para Test rápido 7.1?

7.3 Un ejemplo

Aquí está un ejemplo que implica tres mediciones distintas de la misma resistencia.

Ejemplo: Tres mediciones de resistencia

Tres estudiantes miden cada uno de ellos una misma resistencia varias veces, y sus tres respuestas finales son (todas ellas en ohmios)

Estudiante 1	$R = 11 \pm 1$
Estudiante 2	$R = 12 \pm 1$
Estudiante 3	$R = 10 \pm 3$

Teniendo en cuenta estos tres resultados, ¿cuál es la mejor estimación de la resistencia R ?

Las tres incertidumbres $\sigma_1, \sigma_2, \sigma_3$ son 1, 1 y 3. Por lo tanto, los pesos correspondientes $w_i = 1/\sigma_i^2$ son

$$w_1 = 1, \quad w_2 = 1, \quad w_3 = \frac{1}{9},$$

La mejor estimación de R es la media ponderada, que según (7.10) es igual a

$$\begin{aligned} R_{\text{mp}} &= \frac{\sum w_i R_i}{\sum w_i} \\ &= \frac{(1 \times 11) + (1 \times 12) + (\frac{1}{9} \times 10)}{1 + 1 + \frac{1}{9}} = 11,42 \text{ ohmios.} \end{aligned}$$

La incertidumbre de este resultado viene dada por (7.12) y es igual a

$$\sigma_{\text{mp}} = \frac{1}{\sqrt{\sum w_i}} = \frac{1}{\sqrt{1 + 1 + \frac{1}{9}}} = 0,69.$$

Por consiguiente, nuestro resultado final (convenientemente redondeado) es

$$R = 11,4 \pm 0,7 \text{ ohmios.}$$

Resulta de interés examinar cuál es la respuesta que obtendríamos si hiciéramos caso omiso del resultado del tercer estudiante, que es tres veces menos preciso y, por tanto, nueve veces menos importante. Aquí, un cálculo simple proporciona $R_{\text{mejor}} = 11,50$ (en comparación con 11,42) con una incertidumbre de 0,71 (en comparación con 0,69). Vemos pues que el resultado del tercer estudiante no afecta demasiado a la respuesta final.

Principales definiciones y ecuaciones del capítulo 7

Si x_1, x_2, \dots, x_N son los resultados de mediciones de una determinada magnitud x , con incertidumbres conocidas $\sigma_1, \sigma_2, \dots, \sigma_N$, entonces la mejor estimación para el verdadero valor de x es la *media ponderada*

$$x_{\text{mp}} = \frac{\sum w_i x_i}{\sum w_i}, \quad [\text{Vea (7.10)}]$$

Capítulo 8

Ajuste mínimo cuadrático

Nuestra discusión sobre el análisis de datos estadístico se ha centrado exclusivamente hasta ahora en la medición repetida de una misma magnitud, no porque dicho análisis sea el problema estadístico más interesante, sino porque este sencillo problema debe estar bien entendido antes de que podamos discutir problemas más generales. Ahora estamos listos para abordar nuestro primer, y muy importante, problema más general.

8.1 Datos que deberían ajustarse a una línea recta

Uno de los tipos más comunes e interesantes de experimentos consiste en la medición de varios valores de dos variables físicas diferentes para investigar la relación matemática que hay entre ellas. Por ejemplo, un investigador puede dejar caer una piedra desde diferentes alturas h_1, \dots, h_n y medir los correspondientes tiempos de caída t_1, \dots, t_n para ver si las alturas y los tiempos están vinculados por la relación esperada $h = \frac{1}{2}gt^2$.

Probablemente los experimentos más importantes de este tipo son aquellos en los que la relación esperada es *lineal*. Por ejemplo, si consideramos que un cuerpo cae con aceleración constante g , entonces su velocidad v debe ser una función lineal del tiempo t ,

$$v = v_0 + gt.$$

De manera más general, consideraremos cualquier par de variables físicas x e y que se sospecha están vinculadas por una relación lineal de la forma

$$y = A + Bx, \tag{8.1}$$

donde A y B son constantes. Desafortunadamente se utilizan muchas notaciones diferentes para una relación lineal; hay que tener cuidado para evitar la confusión entre la forma (8.1) con la igualmente popular $y = ax + b$.

Si las dos variables x e y están relacionadas linealmente como en (8.1), entonces una gráfica de y respecto de x debería ser una línea recta de pendiente B y cuya intersección con el eje de ordenadas ocurre en $y = A$. Si tuviéramos que medir N valores diferentes de x_1, \dots, x_N y los valores correspondientes y_1, \dots, y_N , entonces, si nuestras mediciones no estuviesen sujetas a ninguna incertidumbre, cada uno de los puntos (x_i, y_i) se situaría exactamente en la recta $y = A + Bx$, como en la figura 8.1(a). Sin embargo, en la práctica *hay* incertidumbres, y todo lo que

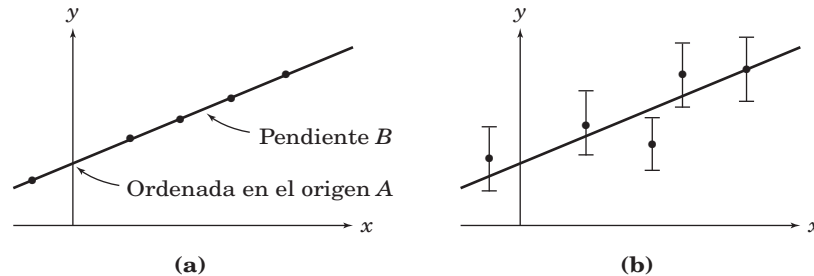


Figura 8.1. (a) Si las dos variables x e y están relacionadas linealmente como en la ecuación (8.1), y si no hay incertidumbres experimentales, entonces los puntos medidos (x_i, y_i) se encuentran todos exactamente en la recta $y = A + Bx$. (b) En la práctica siempre hay incertidumbres que se pueden mostrar mediante barras de error, y solamente se puede esperar que los puntos (x_i, y_i) estén situados razonablemente cerca de la recta. Aquí se muestra que solamente y está sujeta a incertidumbres apreciables.

podemos esperar es que la distancia de cada punto (x_i, y_i) a la línea sea compatible con las incertidumbres de las variables, como en la figura 8.1(b).

Cuando realizamos una serie de medidas del tipo que acabamos de describir, podemos hacernos dos preguntas. En primer lugar podemos preguntarnos si damos por sentado que y y x están relacionadas linealmente; entonces el problema interesante es encontrar la línea recta $y = A + Bx$ que mejor se ajuste a las mediciones obtenidas, es decir, encontrar las mejores estimaciones de las constantes A y B a partir de los datos $(x_1, y_1), \dots, (x_N, y_N)$. Este problema se puede abordar gráficamente, como se discutió brevemente en la sección 2.6. Pero también se puede tratar analíticamente mediante el principio de máxima verosimilitud. Este método analítico de encontrar la mejor línea recta para adaptarse a una serie de puntos experimentales se denomina *regresión lineal*, o *ajuste a una recta por mínimos cuadrados*, y es el tema principal de este capítulo.

La segunda pregunta que nos podemos hacer es si los valores medidos $(x_1, y_1), \dots, (x_N, y_N)$ realmente confirman nuestra expectativa de que y depende linealmente de x . Para responder a esta pregunta, primero tenemos que encontrar la línea recta que mejor se ajuste a los datos, pero para ello debemos idear alguna medida de lo *bien* que se ajustan los datos a la línea recta. Si ya conocemos las incertidumbres de nuestras mediciones, se puede dibujar un gráfico, como el de la figura 8.1(b), que muestra la recta de mejor ajuste y los datos experimentales con sus barras de error. Entonces podemos juzgar visualmente si la línea de mejor ajuste pasa suficientemente cerca de todas las barras de error. Si no conocemos las incertidumbres de manera fiable, debemos juzgar el ajuste de nuestros puntos a una línea recta examinando cómo se distribuyen los propios puntos. Volveremos a esta cuestión en el capítulo 9.

8.2 Cálculo de las constantes A y B

Volvamos ahora a la cuestión de encontrar la línea recta $y = A + Bx$ que mejor se ajusta a un conjunto de puntos obtenidos $(x_1, y_1), \dots, (x_N, y_N)$. Para simplificar nuestro análisis, vamos a suponer que nuestras mediciones de y tienen una incertidumbre apreciable, mientras que la incertidumbre de las mediciones de x

es despreciable. Esta suposición es con frecuencia razonable, ya que a menudo las incertidumbres de una variable son mucho mayores que las correspondientes a la otra, y ésta, por tanto, se puede ignorar sin perjuicios. A continuación supondremos que todos los valores de y tienen la misma incertidumbre. (Esta suposición también es razonable en muchos experimentos, pero si las incertidumbres son diferentes, entonces nuestro análisis puede ser generalizado ponderando las mediciones apropiadamente; vea el problema 8.9). Más concretamente, supondremos que la medición de cada y_i sigue una distribución de Gauss, con idéntico parámetro de anchura σ_y para todas las mediciones. Si conociéramos las constantes A y B , entonces, para cualquier valor dado x_i (que suponemos conocido sin ninguna incertidumbre), podríamos calcular el verdadero valor correspondiente y_i

$$(\text{verdadero valor de } y_i) = A + Bx_i. \quad (8.2)$$

La medición de y_i seguirá una distribución normal centrada en el valor verdadero, con parámetro de anchura σ_y . Por lo tanto, la probabilidad de obtener el valor observado es y_i es

$$\text{Prob}_{A,B}(y_i) \propto \frac{1}{\sigma_y} e^{-(y_i - A - Bx_i)^2 / 2\sigma_y^2}, \quad (8.3)$$

donde los subíndices A y B indican que esta probabilidad depende de los valores (desconocidos) A y B . La probabilidad de obtener nuestro conjunto completo de mediciones y_1, \dots, y_N es el producto

$$\begin{aligned} \text{Prob}_{A,B}(y_1, \dots, y_N) &= \text{Prob}_{A,B}(y_1) \dots \text{Prob}_{A,B}(y_N) \\ &\propto \frac{1}{\sigma_y^N} e^{-\chi^2/2}, \end{aligned} \quad (8.4)$$

donde el exponente viene dado por

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - A - Bx_i)^2}{\sigma_y^2}. \quad (8.5)$$

Siguiendo un razonamiento, ya familiar, supondremos que las mejores estimaciones de las constantes desconocidas A y B , a partir de las mediciones obtenidas, son los valores de A y B para los cuales la probabilidad $\text{Prob}_{A,B}(y_1, \dots, y_N)$ es máxima, o para los que la suma de cuadrados χ^2 en (8.5) es mínima. (Esta es la razón por la que este método se conoce como ajuste mínimo cuadrático). Para encontrar estos valores, derivaremos χ^2 con respecto a A y B e igualaremos dichas derivadas parciales a cero:

$$\frac{\partial \chi^2}{\partial A} = \frac{-2}{\sigma_y^2} \sum_{i=1}^N (y_i - A - Bx_i) = 0 \quad (8.6)$$

y

$$\frac{\partial \chi^2}{\partial B} = \frac{-2}{\sigma_y^2} \sum_{i=1}^N x_i (y_i - A - Bx_i) = 0. \quad (8.7)$$

Estas dos ecuaciones pueden reescribirse como ecuaciones simultáneas para A y B :

$$AN + B\sum x_i = \sum y_i \quad (8.8)$$

y

$$A\sum x_i + B\sum x_i^2 = \sum x_i y_i \quad (8.9)$$

Aquí hemos omitido los límites $i = 1$ a N en los sumatorios Σ . En la siguiente discusión también se omiten los subíndices i cuando no hay peligro de confusión, por lo que $\sum x_i y_i$, se abrevia como $\sum xy$ y así sucesivamente.

Las dos ecuaciones (8.8) y (8.9), a veces llamadas *ecuaciones normales*, se resuelven fácilmente y proporcionan las estimaciones mínimo cuadráticas de A y B ,

$$A = \frac{\sum x^2 \sum y - \sum x \sum xy}{\Delta} \quad (8.10)$$

y

$$B = \frac{N \sum xy - \sum x \sum y}{\Delta}, \quad (8.11)$$

donde hemos introducido una abreviación conveniente para el denominador

$$\Delta = N \sum x^2 - (\sum x)^2. \quad (8.12)$$

Los resultados (8.10) y (8.11) proporcionan las mejores estimaciones para las constantes A y B de la línea recta $y = A + Bx$, a partir de los N puntos obtenidos a partir de las mediciones, $(x_1, y_1), \dots, (x_N, y_N)$. La línea recta resultante es la que presenta un *mejor ajuste mínimo-cuadrático* a los datos, o *recta de regresión* de y sobre x .

Ejemplo: Longitud de un muelle respecto a la masa suspendida en un dinamómetro
Una estudiante construye una escala para medir masas mediante un muelle. Para ello fija el extremo superior del muelle a un soporte rígido, cuelga una bandeja en su parte inferior con una aguja indicadora para leer con comodidad la longitud del muelle sobre la escala. Antes de que pueda usar el dispositivo, debe calibrarlo, es decir, tiene que encontrar la relación que hay entre la masa situada en la bandeja y la longitud del muelle. Para esta calibración dispone de cinco masas de exactamente 2 kg cada una, y que añade a la bandeja de una en una y registra las correspondientes longitudes del muelle tal como se muestra en las tres primeras columnas de la tabla 8.1. Suponiendo que el muelle obedece la ley de Hooke, la estudiante anticipa que l debe ser una función lineal de m ,

$$l = A + Bm. \quad (8.13)$$

Tabla 8.1. Masas m_i (en kg) y longitudes l_i (en cm) de un dinamómetro. Las “x” y las “y” entre comillas indican que variables desempeñan el papel de x (abscisas) e y (ordenadas) en este ejemplo.

Nº. de ensayo i	“x” Carga, m_i	“y” Longitud, l_i	m_i^2	$m_i l_i$
1	2	42,0	4	84
2	4	48,4	16	194
3	6	51,3	36	308
4	8	56,3	64	450
5	10	58,6	100	586
$N = 5$	$\Sigma m_i = 30$	$\Sigma l_i = 256,6$	$\Sigma m_i^2 = 220$	$\Sigma m_i l_i = 1622$

(Aquí, la constante A es la longitud sin carga del muelle, y B es g/k , donde k es la usual constante de elasticidad del muelle). La ecuación de calibración (8.13) le permite determinar cualquier masa m desconocida a partir de la correspondiente longitud l , una vez determinadas las constantes A y B . Para encontrar estas constantes, utiliza el método de los mínimos cuadrados. ¿Cuáles serán sus respuestas para A y B ? Dibuje sus datos de calibración y la recta obtenida mediante su mejor ajuste (8.13). Si coloca una masa m desconocida en la bandeja y observa que la longitud del muelle es igual a $l = 53,2$ cm, ¿cuál es el valor de m ?

Como suele suceder en este tipo de problemas, las dos variables no se llaman x e y , y hay que tener cuidado al identificar cuál es cuál. Comparando (8.13) con la forma estándar, $y = A + Bx$, vemos que la longitud l juega el papel de la variable dependiente y , mientras que la masa m juega el papel de la variable independiente x . Las constantes A y B se estiman a partir de (8.10) a (8.12), con las sustituciones

$$x_i \leftrightarrow m_i \quad \text{y} \quad y_i \leftrightarrow l_i$$

(Esta correspondencia se indica por los encabezados “x” e “y” en el cuadro 8.1). Para hallar A y B , tenemos que calcular las sumas Σm_i , Σl_i , Σm_i^2 y $\Sigma m_i l_i$; por tanto, las últimas dos columnas de la tabla 8.1 muestran las cantidades m_i^2 y $m_i l_i$, y la suma correspondiente se muestra en la parte inferior de cada columna.

El cálculo de las constantes A y B es ahora directo. De acuerdo con (8.12), el denominador Δ es

$$\begin{aligned} \Delta &= N \Sigma m^2 - (\Sigma m)^2 \\ &= 5 \times 220 - 30^2 = 200. \end{aligned}$$

A continuación, a partir de (8.10) encontramos la ordenada en el origen (la longitud del muelle sin carga que lo estire)

$$\begin{aligned} A &= \frac{\Sigma m^2 \Sigma l - \Sigma m \Sigma m l}{\Delta} \\ &= \frac{220 \times 256,6 - 30 \times 1622}{200} = 39,0 \text{ cm.} \end{aligned}$$

Por último, a partir de (8.11) encontramos la pendiente

$$B = \frac{N\sum ml - \sum m \sum l}{\Delta} \\ = \frac{5 \times 1622 - 30 \times 256,6}{200} = 2,06 \text{ cm/kg.}$$

En la figura 8.2 se muestra la gráfica de los datos y la recta (8.13) con estos valores de A y B . Si la masa m estira el muelle a 53,2 cm, entonces, según (8.13), la masa es

$$m = \frac{l - A}{B} = \frac{(53,2 - 39,0) \text{ cm}}{2,06 \text{ cm/kg}} = 6,9 \text{ kg.}$$

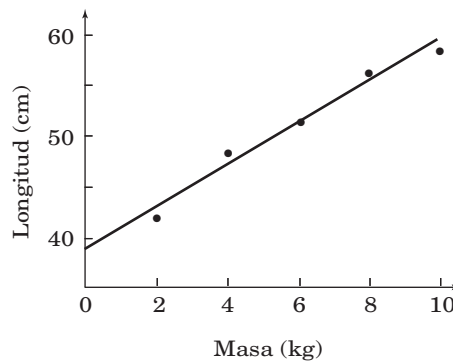


Figura 8.2. Gráfico de los datos de la tabla 8.1 y la recta con mejor ajuste (8.13).

Test rápido 8.1. Encuentre la recta $y = A + Bx$ con mejor ajuste mínimo cuadrático a los tres puntos $(x, y) = (-1, 0)$, $(0, 6)$ y $(1, 6)$. Utilizando papel cuadriculado, represente los puntos y la recta. [Observación: puesto que los tres valores de x (-1 , 0 y 1) son simétricos con respecto a cero, $\sum x = 0$, lo que simplifica el cálculo de A y B . En algunos experimentos, los valores de x pueden estar simétricamente espaciados, ahorrando de este modo alguna dificultad.]

Ahora que sabemos cómo encontrar las mejores estimaciones de las constantes A y B , es natural preguntar por las incertidumbres de estas estimaciones. Sin embargo, antes de que podamos encontrarlas, debemos discutir la incertidumbre σ_y de las mediciones originales y_1, y_2, \dots, y_N .

8.3 Incertidumbre en la medición de y

En el transcurso de la medición de los valores y_1, \dots, y_N , probablemente nos habremos hecho una idea de su incertidumbre. Sin embargo, saber cómo calcular la incertidumbre mediante el análisis de los datos en sí es importante. Recuerde que los números y_1, \dots, y_N no son N mediciones de la misma magnitud. (Podrían

Capítulo 9

Covarianza y correlación

Este capítulo introduce el importante concepto de covarianza. Como este concepto surge de forma natural en la propagación de errores, la sección 9.1 empieza con un repaso rápido de la misma. Esta revisión prepara el escenario para la sección 9.2, que define la covarianza y discute su papel en dicha propagación. Entonces, la sección 9.3 utiliza la covarianza para definir el coeficiente de correlación lineal para un conjunto de puntos $(x_1, y_1), \dots, (x_N, y_N)$ correspondientes a mediciones. Este coeficiente, denotado por r , proporciona una medida de lo bien que dichos puntos se ajustan a una línea recta de la forma $y = A + Bx$; su uso se describe en las secciones 9.4 y 9.5.

9.1 Revisión de la propagación de errores

Esta y la siguiente sección proporcionan una última mirada a la importante cuestión de la propagación de errores. Este tema lo analizamos por primera vez en el capítulo 3, donde llegamos a varias conclusiones. Nos imaginamos la medición de dos cantidades x e y para calcular alguna función $q(x, y)$, tal como $q = x + y$ o $q = x^2 \sin y$. [De hecho, hablamos de una función $q(x, \dots, z)$ de un número arbitrario de variables x, \dots, z ; para simplificar, vamos a considerar ahora sólo dos variables]. Un simple argumento sugiere que la incertidumbre de q es

$$\delta q \approx \left| \frac{\partial q}{\partial x} \right| \delta x + \left| \frac{\partial q}{\partial y} \right| \delta y. \quad (9.1)$$

En primer lugar habíamos deducido esta aproximación para los casos especiales y sencillos de sumas, diferencias, productos y cocientes. Por ejemplo, si q es igual a la suma $q = x + y$, entonces (9.1) se reduce al familiar $\delta q \approx \delta x + \delta y$. El resultado general (9.1) se dedujo de la ecuación (3.43).

A continuación reconocimos que a menudo (9.1) es probablemente una estimación por exceso de δq , porque puede haber cancelaciones parciales de los errores de x e y . Afirmamos, sin demostrarlo, que cuando los errores en x e y son independientes y aleatorios, un mejor valor para la incertidumbre del valor calculado de $q(x, y)$ es la suma cuadrática

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x} \delta x \right)^2 + \left(\frac{\partial q}{\partial y} \delta y \right)^2}. \quad (9.2)$$

También afirmamos, sin demostración, que tanto para los errores independientes como para los aleatorios, la fórmula más sencilla (9.1) siempre da un límite superior de δq ; es decir, la incertidumbre δq nunca es mayor que la proporcionada por (9.1).

En el capítulo 5 se da una definición y una demostración adecuadas de (9.2). En primer lugar, vimos que una buena medida de la incertidumbre δx de una medición está dada por la desviación estándar σ_x ; en particular, vimos que si las mediciones de x se distribuyen normalmente, podemos estar seguros de que en el 68% de los casos el valor medido diferirá en menos de σ_x del valor real. En segundo lugar, vimos que si las mediciones de x e y se rigen por distribuciones normales independientes, con desviaciones estándar σ_x y σ_y , los valores de $q(x, y)$ siguen también una distribución normal, con una desviación estándar

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x} \sigma_x\right)^2 + \left(\frac{\partial q}{\partial y} \sigma_y\right)^2}. \quad (9.3)$$

Este resultado proporciona la justificación de la afirmación (9.2).

En la sección 9.2 vamos a deducir una fórmula precisa para la incertidumbre de q que se aplica tanto si los errores de x e y son independientes y se distribuyen normalmente como si no lo son. En particular, demostraremos que (9.1) siempre proporciona un límite superior a la incertidumbre de q .

Antes de deducir estos resultados, vamos a revisar primero la definición de la desviación estándar. La desviación estándar σ_x de N mediciones x_1, \dots, x_N se definió originalmente a través de la ecuación

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2. \quad (9.4)$$

Si las mediciones de x se distribuyen normalmente, entonces en el límite, donde N es grande, la definición (9.4) equivaldría a definir σ_x como el parámetro de anchura que aparece en la función de Gauss

$$\frac{1}{\sigma_x \sqrt{2\pi}} e^{-(x - \bar{x})^2 / 2\sigma_x^2}$$

que cuantifica la expectación de las mediciones de x . Puesto que ahora vamos a considerar la posibilidad de que los errores de x puedan no seguir una distribución normal, esta segunda definición ya no está disponible para nosotros. Sin embargo, aún podemos, y lo haremos, definir σ_x mediante (9.4). Sea o no normal la distribución de los errores, dicha definición de σ_x da una medida razonable de las incertidumbres aleatorias de nuestra medición de x . (Al igual que en el capítulo 5, supondremos que todos los errores sistemáticos se han identificado y reducido a un nivel insignificante, por lo que todos los errores restantes *son* aleatorios).

La ambigüedad habitual sobre si usar la definición (9.4) para σ_x o la definición “mejorada” con el factor N en el denominador sustituido por $(N - 1)$. Afortunadamente, la discusión que sigue se aplica a cualquiera de ambas definiciones, con tal de que seamos coherentes en el uso de una u otra. Para mayor comodidad, voy a utilizar la definición (9.4), con N en el denominador a lo largo del presente capítulo.

9.2 La covarianza en la propagación de errores

Suponga que para encontrar un valor de la función $q(x, y)$ medimos varias veces las dos magnitudes x e y , y obtenemos N pares de datos, $(x_1, y_1), \dots, (x_N, y_N)$. A partir de las N mediciones x_1, \dots, x_N , podemos calcular su media muestral \bar{x} y la desviación estándar σ_x de la forma habitual; de manera similar, a partir de y_1, \dots, y_N , podemos calcular \bar{y} y σ_y . A continuación, utilizando los N pares de mediciones, se pueden calcular N valores de la cantidad de interés

$$q_i = q(x_i, y_i), \quad (i = 1, \dots, N).$$

Dados q_1, \dots, q_N , podemos calcular su media \bar{q} , que suponemos proporciona nuestra mejor estimación de q , y su desviación estándar σ_q , que es nuestra medida de la incertidumbre aleatoria de los valores q_i .

Supondremos, como es habitual, que todas nuestras incertidumbres son pequeñas y por lo tanto que todos los números x_1, \dots, x_N están cerca de \bar{x} , y que todos los y_1, \dots, y_N están cerca de \bar{y} . Podemos entonces hacer la aproximación

$$\begin{aligned} q_i &= q(x_i, y_i) \\ &\approx q(\bar{x}, \bar{y}) + \frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}). \end{aligned} \quad (9.5)$$

En esta expresión, las derivadas parciales $\partial q/\partial x$ y $\partial q/\partial y$ se toman en el punto $x = \bar{x}, y = \bar{y}$, y son por lo tanto las mismas para todos los $i = 1, \dots, N$. Con esta aproximación, la media se convierte en

$$\begin{aligned} \bar{q} &= \frac{1}{N} \sum_{i=1}^N q_i \\ &= \frac{1}{N} \sum_{i=1}^N \left[q(\bar{x}, \bar{y}) + \frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}) \right]. \end{aligned}$$

Esta ecuación expresa \bar{q} como la suma de tres términos. El primer término es $q(\bar{x}, \bar{y})$, y los otros dos son exactamente iguales a cero. [Observemos que de la definición de \bar{x} que $\sum (x_i - \bar{x}) = 0$.] Por lo tanto, tenemos el resultado notablemente simple

$$\bar{q} = q(\bar{x}, \bar{y}); \quad (9.6)$$

es decir, para encontrar la media \bar{q} solo tenemos que calcular la función $q(x, y)$ en el punto $x = \bar{x}$ e $y = \bar{y}$.

La desviación estándar de los N valores q_1, \dots, q_N viene dada por

$$\sigma_q^2 = \frac{1}{N} \sum (q_i - \bar{q})^2.$$

Sustituyendo (9.5) y (9.6), encontramos que

$$\begin{aligned} \sigma_q^2 &= \frac{1}{N} \sum \left[\frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}) \right]^2 \\ &= \left(\frac{\partial q}{\partial x} \right)^2 \frac{1}{N} \sum (x_i - \bar{x})^2 + \left(\frac{\partial q}{\partial y} \right)^2 \frac{1}{N} \sum (y_i - \bar{y})^2 \\ &\quad + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} \frac{1}{N} \sum (x_i - \bar{x})(y_i - \bar{y}). \end{aligned} \quad (9.7)$$

Las sumas de los dos primeros términos son los que aparecen en la definición de las desviaciones estándar σ_x y σ_y . La suma final es una que no ha aparecido hasta ahora. Se llama *covarianza*¹ de x y y y se define como

$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}). \quad (9.8)$$

Con esta definición, la ecuación (9.7) para la desviación estándar σ_q se convierte en

$$\sigma_q^2 = \left(\frac{\partial q}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial q}{\partial y} \right)^2 \sigma_y^2 + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} \sigma_{xy}. \quad (9.9)$$

Esta ecuación proporciona la desviación estándar σ_q , sin requerir que las mediciones de x y y sean independientes o se distribuyan normalmente.

Si las mediciones de x y y son independientes, podemos ver fácilmente que, después de muchas mediciones, la covarianza σ_{xy} debe aproximarse a cero: cualquiera que sea el valor de y_i , la cantidad $x_i - \bar{x}$ es igual de probable que sea negativa como que sea positiva. Así, después de muchas mediciones, los términos positivos y negativos en (9.8) deben casi equilibrarse; en el límite de un número infinito de mediciones, el factor $1/N$ en (9.8) garantiza que σ_{xy} sea cero. (Después de un número finito de mediciones, σ_{xy} no será exactamente igual a cero, pero deberá ser *pequeña* si los errores en x y y son realmente independientes y aleatorios). Con σ_{xy} igual a cero, la ecuación (9.9) para σ_q se reduce a

$$\sigma_q^2 = \left(\frac{\partial q}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial q}{\partial y} \right)^2 \sigma_y^2, \quad (9.10)$$

el resultado familiar correspondiente a las incertidumbres independientes y aleatorias.

Si las mediciones de x y y no son independientes, la covarianza σ_{xy} no necesita ser cero. Por ejemplo, es fácil imaginar una situación en la que una sobreestimación de x siempre esté acompañada por una sobreestimación de y , y viceversa. Los números $(x_i - \bar{x})$ y $(y_i - \bar{y})$ tendrán entonces siempre el mismo signo (ambos positivos o ambos negativos), y su producto siempre será positivo. Debido a que todos los términos en la suma (9.8) son positivos, σ_{xy} será positivo (y distinto de cero), incluso en el límite al tender el número de mediciones a infinito. A la inversa, es posible imaginar situaciones en las que una sobreestimación de x siempre se acompaña de una subestimación de y , y viceversa, en este caso $(x_i - \bar{x})$ y $(y_i - \bar{y})$

1. El nombre *covarianza* para σ_{xy} (para dos variables x, y) es análogo al nombre *varianza* para σ_x^2 (de una variable x). Para enfatizar este paralelismo, la covarianza (9.8) a veces se designa como σ_{xy}^2 , una notación no demasiado conveniente debido a que la covarianza puede ser negativa. Una característica adecuada de la definición (9.8) es que σ_{xy} tiene las dimensiones de xy , tal como σ_x tiene las dimensiones de x .

siempre tendrán signos opuestos, y σ_{xy} será negativa. Este caso se ilustra en el ejemplo siguiente.

Cuando la covarianza σ_{xy} no es cero (incluso en el límite de un número infinito de mediciones), se dice que los errores de x e y están *correlacionados*. En este caso, la incertidumbre σ_q de $q(x, y)$ dada por (9.9) *no* es la misma que la que se obtiene a partir de la fórmula (9.10) de los errores aleatorios independientes.

Ejemplo: Dos ángulos con covarianza negativa

Cinco estudiantes miden cada uno los mismos dos ángulos α y β y obtienen los resultados que se muestran en las tres primeras columnas de la tabla 9.1.

Tabla 9.1. Cinco mediciones de dos ángulos α y β (en grados).

Estudiante	α	β	$(\alpha - \bar{\alpha})$	$(\beta - \bar{\beta})$	$(\alpha - \bar{\alpha})(\beta - \bar{\beta})$
A	35	50	2	-2	-4
B	31	55	-2	3	-6
C	33	51	0	-1	0
D	32	53	-1	1	-1
E	34	51	1	-1	-1

Encuentre la media y la desviación estándar para cada uno de los dos ángulos y, a continuación, busque la covarianza $\sigma_{\alpha\beta}$ según la expresión definida en (9.8). Los estudiantes a continuación calculan la suma $q = \alpha + \beta$. Encuentre su mejor estimación para q según (9.6) y la desviación estándar σ_q según (9.9). Compare la desviación estándar que se obtendría si supusiéramos (erróneamente) que los errores en α y β son independientes y que σ_q viniese dada por (9.10).

Se observa de inmediato que las medias son $\bar{\alpha} = 33$ y $\bar{\beta} = 52$. Con estos valores podemos encontrar las diferencias $(\alpha - \bar{\alpha})$ y $(\beta - \bar{\beta})$, como se muestra en la tabla 9.1, y de estas desviaciones es fácil encontrar

$$\sigma_{\alpha}^2 = 2,0 \quad \text{y} \quad \sigma_{\beta}^2 = 3,2.$$

[Aquí se ha utilizado la definición (9.4), con N en el denominador].

Puede ver en la tabla 9.1 que los valores altos de α parecen estar correlacionados con los valores bajos de β y viceversa, debido a que $(\alpha - \bar{\alpha})$ y $(\beta - \bar{\beta})$ siempre tienen signos opuestos. (Para un experimento en el que se produce este tipo de correlación, vea el problema 9.6). Esta correlación significa que los productos $(\alpha - \bar{\alpha})$ y $(\beta - \bar{\beta})$ que se muestran en la última columna de la tabla son todos negativos (o cero). Por lo tanto, la covarianza $\sigma_{\alpha\beta}$ tal como se define en (9.8) es negativa,

$$\sigma_{\alpha\beta} = \frac{1}{N} \Sigma (\alpha - \bar{\alpha})(\beta - \bar{\beta}) = \frac{1}{5} \times (-12) = -2,4.$$

La mejor estimación de la suma $q = \alpha + \beta$ viene dada por (9.6) como

$$q_{\text{mejor}} = \bar{q} = \bar{\alpha} + \bar{\beta} = 33 + 52 = 85.$$

Capítulo 10

La distribución binomial

Hasta ahora, el único ejemplo de distribución que hemos estudiado es la distribución normal o de Gauss. Ahora vamos a ver otros dos ejemplos importantes, la distribución binomial (en el presente capítulo) y la distribución de Poisson (en el capítulo 11).

10.1 Distribuciones

En el capítulo 5 se introdujo la idea de *distribución*, la función que describe la proporción de cada uno de los diferentes resultados posibles al repetir un determinado experimento. Por ejemplo, podríamos hacer N mediciones del periodo T de un péndulo y encontrar la distribución de los distintos valores de T , o podríamos medir la altura h de N personas y encontrar la distribución de las distintas alturas h obtenidas.

Después se introdujo la noción de *distribución límite*, la distribución que se obtendría en el límite a medida que el número de mediciones N se hace muy grande. La distribución límite puede ser vista como una herramienta que nos dice la *probabilidad* de que una medición tenga cualquiera de los valores posibles: la probabilidad de que al medir el periodo obtengamos un determinado valor de T , la probabilidad de que la altura de una persona (elegida al azar) tenga un determinado valor h . Por esta razón, a la distribución límite a veces también se le llama *distribución de probabilidad*.

De las muchas posibles distribuciones límite, la única que hemos discutido es la distribución normal o de Gauss, que describe la distribución de resultados de cualquier medición influida por muchas fuentes de error, todas ellas aleatorias y pequeñas. Esta distribución es para los científicos la más importante de todas las distribuciones límite, y la importancia que se le ha dado aquí está plenamente justificada. No obstante, otras varias distribuciones tienen una gran importancia teórica o práctica, y en el presente capítulo y en el siguiente se presentan sendos ejemplos.

En este capítulo se describe la distribución binomial, que no es de gran utilidad práctica para el físico experimental. Su sencillez, sin embargo, la convierte en una excelente introducción a muchas propiedades de las distribuciones, y tiene importancia desde un punto de vista teórico ya que permite obtener a partir de ella la siempre importante distribución de Gauss.

10.2 Probabilidades en un lanzamiento de dados

La distribución binomial se puede describir mejor con un ejemplo. Supongamos que realizamos el “experimento” consistente en lanzar tres dados y anotar el número de ases (o unos) aparecidos. Los posibles resultados del experimento son 0, 1, 2 o 3 ases. Si repetimos el experimento un número grande de veces, obtendremos la distribución límite, que nos indicará la probabilidad de que en cualquier lanzamiento (de los tres dados) obtengamos v ases, donde $v = 0, 1, 2$ o 3 .

Este experimento es lo suficientemente simple para poder calcular fácilmente la probabilidad de los cuatro resultados posibles. Observamos en primer lugar que, suponiendo que los dados sean perfectos, la probabilidad de sacar un as al lanzar *un* dado es $1/6$. Vamos ahora a lanzar los tres dados y hallar primero la probabilidad de obtener tres ases ($v = 3$). Debido a que cada dado por separado tiene la probabilidad $1/6$ de mostrar un as, y porque los lanzamientos de los tres dados se efectúan de forma independiente, la probabilidad de de obtener tres ases es

$$\text{Prob}(3 \text{ ases en } 3 \text{ lanzamientos}) = \left(\frac{1}{6}\right)^3 \approx 0,5\%.$$

El cálculo de la probabilidad de obtener dos ases ($v = 2$) es un poco más difícil, ya que podemos obtener dos ases en los lanzamientos de varias formas distintas. El primer dado y el segundo podrían mostrar sendos ases y no así el tercero (A, A, no A), o el primero y el tercero podría mostrar ases, pero no el segundo (A, no A, A), y así sucesivamente. Ahora se argumenta en dos pasos. En primer lugar, consideramos la probabilidad de obtener dos ases en un orden definido, tal como (A, A, no A). La probabilidad de que el primer dado muestre un as es $1/6$, y lo mismo ocurre con el segundo. Por otra parte, la probabilidad de que el último dado no muestre un as es $5/6$. Por lo tanto, la probabilidad de obtener dos ases en este orden particular es

$$\text{Prob}(A, A, \text{no } A) = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right).$$

La probabilidad de obtener dos ases en cualquier otro orden definido es la misma. Por último, hay tres ordenaciones diferentes en las cuales se pueden obtener dos ases: (A, A, no A) o (A, no A, A), o (no A, A, A). Por lo tanto, la probabilidad total para conseguir dos ases (en cualquier orden) es

$$\text{Prob}(2 \text{ ases en } 3 \text{ tiradas}) = 3 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right) \approx 6,9\%. \quad (10.1)$$

Cálculos similares permiten obtener las probabilidades de obtener un as en tres tiradas (34,7%) y ningún as los tres lanzamientos (57,9%). Nuestras conclusiones numéricas pueden ser resumidas representando gráficamente la distribución de probabilidad de la variable número de ases obtenidos al lanzar tres dados, como en la figura 10.1. Esta distribución es un ejemplo de la distribución binomial, cuya forma general describiremos a continuación.

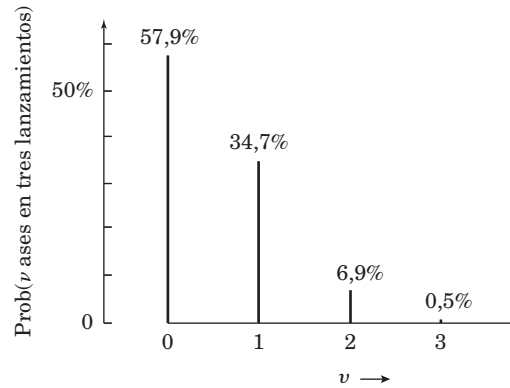


Figura 10.1. Probabilidad de obtener v ases cuando lanzamos tres dados. Esta función es la distribución binomial $B_{n,p}(v)$, con $n = 3$ y $p = 1/6$.

10.3 Definición de la distribución binomial

Para describir la distribución binomial general es necesario introducir algo de terminología. En primer lugar, imaginemos que estamos realizando n ensayos independientes, como lanzar n dados, lanzar n monedas o probar n petardos. Cada ensayo puede tener varios resultados: un dado puede mostrar cualquier cara de la 1 a la 6, una moneda puede mostrar cara o cruz, un petardo puede explotar o no explotar. Al resultado que nos interesa lo llamaremos *éxito*. Así, el “éxito” podría ser obtener un as en el lanzamiento de un dado, una cara al tirar una moneda o que un petardo explote al encenderlo. Denotaremos por p la probabilidad de éxito en cualquier ensayo, y por $q = 1 - p$ la de “fracaso” (es decir, de tener otro resultado diferente del que nos interesa). Por lo tanto, $p = 1/6$ para obtener un as al lanzar un dado, $p = 1/2$ para obtener cara al lanzar una moneda y, para una determinada marca de petardo, la probabilidad p de que explote correctamente podría ser igual al 95%.

Provistos con estas definiciones, podemos preguntarnos por la probabilidad de obtener v éxitos en n ensayos. A continuación esbozaremos un cálculo que muestra que esta probabilidad viene dada por la llamada *distribución binomial*:

$$\begin{aligned} \text{Prob}(v \text{ éxitos en } n \text{ ensayos}) &= B_{n,p}(v) \\ &= \frac{n(n-1)\cdots(n-v+1)}{1 \times 2 \times \cdots \times v} p^v q^{n-v}. \end{aligned} \quad (10.2)$$

Aquí la letra B significa “binomio”; los subíndices n y p en $B_{n,p}(v)$ indican que la distribución depende de n , el número de ensayos efectuados, y p , la probabilidad de éxito en cualquier ensayo individual.

La distribución (10.2) se denomina distribución binomial porque está estrechamente relacionada con el bien conocido desarrollo del binomio. En con-

creto, el cociente en (10.2) es el *coeficiente binomial*, a menudo indicado por

$$\binom{n}{v} = \frac{n(n-1) \cdots (n-v+1)}{1 \times 2 \times \cdots \times v} \quad (10.3)$$

$$= \frac{n!}{v!(n-v)!}, \quad (10.4)$$

donde hemos introducido la útil notación *factorial*,

$$n! = 1 \times 2 \times \cdots \times n.$$

[Por convenio, $0! = 1$, y por tanto $\binom{n}{0} = 1$]. Los coeficientes binomiales aparecen en el desarrollo del binomio

$$\begin{aligned} (p+q)^n &= p^n + np^{n-1}q + \cdots + q^n \\ &= \sum_{v=0}^n \binom{n}{v} p^v q^{n-v}, \end{aligned} \quad (10.5)$$

que se cumple para cualquier par de números p y q y cualquier entero positivo n (consulte los problemas 10.5 y 10.6).

Con la notación (10.3), podemos reescribir la distribución binomial en la forma más compacta

Distribución binomial

$$\begin{aligned} \text{Prob}(v \text{ éxitos en } n \text{ ensayos}) &= B_{n,p}(v) \\ &= \binom{n}{v} p^v q^{n-v}, \end{aligned} \quad (10.6)$$

donde, como de costumbre, p indica la probabilidad de éxito en un ensayo cualquiera y $q = 1 - p$.

La deducción del resultado (10.6) es similar al ejemplo del lanzamiento de tres dados en (10.1)

$$\text{Prob}(2 \text{ ases en } 3 \text{ lanzamientos}) = 3 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right). \quad (10.7)$$

De hecho, si fijamos $v = 2$, $n = 3$, $p = 1/6$ y $q = 5/6$ en (10.6), se obtiene precisamente (10.7), como el lector debería comprobar. Además, el significado de cada factor en (10.6) es el mismo que el del factor correspondiente en (10.7). El factor p^v es la probabilidad de obtener solo éxitos en v ensayos cualesquiera previamente determinados, y q^{n-v} es la probabilidad de obtener solo fallos en los restantes $n - v$ ensayos. El coeficiente binomial $\binom{n}{v}$ se demuestra fácilmente que es igual al número de ordenaciones diferentes que podemos obtener v éxitos en n ensayos. Este razonamiento establece que la distribución binomial (10.6) es efectivamente la probabilidad buscada.

Ejemplo: Lanzamiento de cuatro monedas

Supongamos que tiramos cuatro monedas ($n = 4$) y contamos el número de caras obtenidas, v . ¿Cuál es la probabilidad de obtener los distintos valores posibles $v = 0, 1, 2, 3, 4$?

Debido a que la probabilidad de obtener cara en una lanzamiento es $p = 1/2$, la probabilidad requerida es simplemente la distribución binomial $B_{n,p}(v)$, con $n = 4$ y $p = q = 1/2$.

$$\text{Prob}(v \text{ caras en 4 lanzamientos}) = \binom{4}{v} \left(\frac{1}{2}\right)^4.$$

Estas probabilidades se calculan fácilmente. Por ejemplo,

$$\text{Prob}(0 \text{ caras en 4 lanzamientos}) = 1 \times \left(\frac{1}{2}\right)^4 = 0,0625.$$

Las cinco probabilidades se muestran en la figura 10.2.

Vemos que el número más probable de caras es $v = 2$, como era de esperar. Aquí, las probabilidades son simétricas alrededor de este valor más probable. Es decir, la probabilidad de obtener tres caras es la misma que la de obtener una, y la probabilidad de salgan cuatro caras es la misma que la de que no salga ninguna. Como veremos, esta simetría se produce sólo cuando $p = 1/2$.

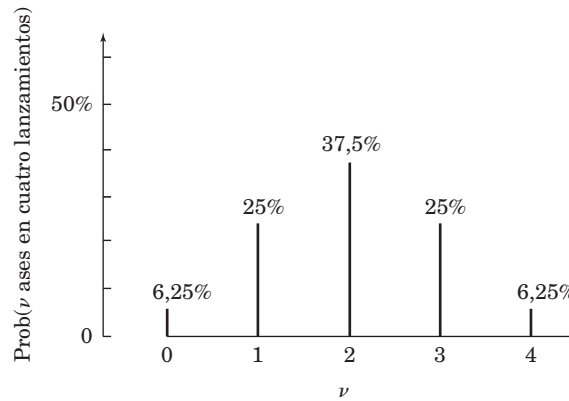


Figura 10.2. La distribución binomial $B_{n,p}(v)$ con $n = 4$, $p = 1/2$. Esto proporciona la probabilidad de obtener v caras cuando lanzamos cuatro monedas.

Test rápido 10.1. Si escogemos una carta al azar de un mazo completo de 52 cartas, la probabilidad de sacar cualquier palo particular es $1/4$. Si elegimos tres cartas, reemplazando cada una de ellas después de cada sorteo, ¿cuál es la probabilidad de sacar tres corazones? ¿De sacar exactamente dos corazones? ¿De sacar dos o más corazones?

Capítulo 11

La distribución de Poisson

En este capítulo se presenta un tercer ejemplo de distribución límite, la distribución de Poisson, que describe los resultados de experimentos en los que contamos los sucesos que se producen al azar, pero con una determinada una tasa media de ocurrencia. Ejemplos de este tipo de experimentos de recuento aparecen en casi todos los ámbitos de la ciencia, por ejemplo, un sociólogo puede contar el número de bebés nacidos en un hospital en un plazo de tres días. Un ejemplo importante en física es el recuento de las desintegraciones de una muestra radiactiva; por ejemplo, un físico nuclear podría decidir contar el número de partículas alfa emitidas por una muestra de gas radón durante un intervalo de diez segundos.

Esta clase de experimentos de recuento se discute en la sección 3.2, donde se enunció, sin demostración, la “regla de la raíz cuadrada”: si se cuentan las ocurrencias de un determinado suceso, en las condiciones comentadas, durante un intervalo de tiempo T y se obtienen v ocurrencias, entonces la mejor estimación para el promedio real del número de ocurrencias de dicho suceso durante un intervalo de tiempo T es, por supuesto, v , y la incertidumbre de esta estimación es \sqrt{v} .

En las secciones 11.1 y 11.2 introduciremos la distribución de Poisson y exploraremos algunas de sus propiedades. En particular, en la sección 11.2 podremos demostrar que la desviación estándar de la distribución de Poisson es la raíz cuadrada del número esperado de sucesos. Este resultado justifica la regla de la raíz cuadrada de la sección 3.2. Las secciones 11.3 y 11.4 describen algunas aplicaciones de la distribución de Poisson.

11.1 Definición de la distribución de Poisson

Como ejemplo de la distribución de Poisson, supongamos que se nos da una muestra de material radiactivo y utilizamos un detector adecuado para encontrar el número v de las partículas emitidas por desintegración radiactiva en un intervalo de dos minutos. Si el contador es fiable, nuestro valor v no tendrá incertidumbre. Sin embargo, si repetimos el experimento, casi con total seguridad obtendremos un valor diferente de la variable v . Esta variación en el valor numérico de v no refleja la incertidumbre en el proceso de recuento, sino que refleja el carácter intrínsecamente aleatorio del proceso de desintegración radiactiva.

Cada núcleo radiactivo tiene una probabilidad definida para la desintegración en cualquier intervalo de tiempo de dos minutos. Si supiéramos dicha probabilidad y el número de núcleos en nuestra muestra, podríamos calcular el *promedio esperado de desintegraciones* en dos minutos. En cualquier caso, el tiempo que tarda en desintegrarse cada núcleo es aleatorio, y dado un intervalo de dos minutos determinado, el número de desintegraciones que registraremos durante el mismo será variable, aunque con un determinado valor medio o esperado.

Obviamente, la pregunta que debemos hacernos es la siguiente: si repetimos el experimento muchas veces (reemplazando la muestra si se agota de manera significativa), ¿qué distribución deberíamos esperar del número de desintegraciones v observado en intervalos de dos minutos de duración? Si el lector ha estudiado el capítulo 10, reconocerá que la distribución requerida es la distribución binomial. Si hay n núcleos y la probabilidad de que un núcleo cualquiera se desintegre en el intervalo es p , entonces la probabilidad de que haya v desintegraciones es sólo la probabilidad de v "éxitos" en n "ensayos", o $B_{n,p}(v)$. Para el tipo de experimento que estamos ahora discutiendo, sin embargo, existe una importante simplificación. El número de "ensayos" (es decir, núcleos) es enorme ($n \sim 10^{20}$, tal vez), y la probabilidad de "éxito" (desintegración) para cualquier núcleo es muy baja (a menudo tan baja como $p \sim 10^{-20}$). En estas condiciones (n grande y p pequeño), se puede demostrar que es posible aproximar la distribución binomial por una distribución más sencilla llamada *distribución de Poisson*. Concretamente, se puede demostrar que

$$\text{Prob}(v \text{ desintegraciones en cualquier intervalo determinado}) = P_{\mu}(v), \quad (11.1)$$

donde la *distribución de Poisson*, $P_{\mu}(v)$, viene dada por

La distribución de Poisson

$$P_{\mu}(v) = e^{-\mu} \frac{\mu^v}{v!}. \quad (11.2)$$

En esta definición, μ es un parámetro positivo ($\mu > 0$) que, como mostraremos directamente, es sólo el valor medio de las desintegraciones registradas en el intervalo de tiempo en cuestión, y $v!$ denota el factorial de v (con $0! = 1$).

EL SIGNIFICADO DE μ COMO TASA MEDIA DE RECuento

No vamos a deducir aquí la distribución de Poisson (11.2), sino que simplemente afirmaremos que *es* la distribución apropiada para el tipo de experimento de recuento que estamos considerando.¹ Para establecer el significado del parámetro μ en (11.2), sólo debemos calcular el número promedio de sucesos registrados, v , que esperamos si repetimos nuestro experimento de recuento muchas

1. Para una demostración, vea, por ejemplo, H. D. Young, *Statistical Treatment of Experimental Data* (McGraw-Hill, 1962), Section 8, o S. L. Meyer, *Data Analysis for Scientists and Engineers* (John Wiley, 1975), p. 207.

veces. Este promedio se calcula sumando todos los valores posibles de v multiplicado cada uno por su probabilidad:

$$\bar{v} = \sum_{v=0}^{\infty} v P_{\mu}(v) = \sum_{v=0}^{\infty} v e^{-\mu} \frac{\mu^v}{v!} \quad (11.3)$$

El primer término de esta suma se puede descartar (ya que es igual a cero), y $v/v!$ puede ser sustituido por $1/(v-1)!$. Si sacamos $\mu e^{-\mu}$ como factor común obtenemos

$$\bar{v} = \mu e^{-\mu} \sum_{v=1}^{\infty} \frac{\mu^{v-1}}{(v-1)!}. \quad (11.4)$$

La suma infinita resultante es

$$1 + \mu + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + \dots = e^{\mu}, \quad (11.5)$$

que es precisamente la función exponencial e^{μ} (como hemos indicado). Por tanto, la exponencial $e^{-\mu}$ en (11.4) se cancela exactamente por la suma, y llegamos a la simple conclusión de que

$$\bar{v} = \mu. \quad (11.6)$$

Es decir, el parámetro μ que caracteriza la distribución de Poisson $P_{\mu}(v)$ es *el valor medio de los recuentos si repetimos el experimento de recuento muchas veces*. A veces es posible conocer de antemano la tasa media R por unidad de tiempo en que deben producirse los acontecimientos que estamos contando. En este caso, el número medio esperado de los sucesos en un tiempo T es simplemente

$$\mu = \text{tasa} \times \text{tiempo} = RT.$$

Si, por el contrario, la tasa de R es desconocida, contando el número de sucesos en un tiempo T , podemos obtener una estimación de μ y por lo tanto estimar la tasa R a partir de $R_{\text{mejor}} = \mu_{\text{mejor}}/T$

Ejemplo: Recuento de desintegraciones radiactivas

Mediciones cuidadosas han establecido que una muestra de torio radiactivo emite partículas alfa a una tasa de 1,5 por minuto. Si cuento el número de partículas alfa emitidas en dos minutos, ¿cuál es el número medio esperado? ¿Cuál es la probabilidad de que el número de partículas alfa coincida realmente con dicho número? ¿Cuál es la probabilidad de observar v partículas alfa para $v = 0, 1, 2, 3, 4$ y para $v \geq 5$?

La media esperada es la tasa media de emisiones ($R = 1,5$ por minuto) multiplicada por el tiempo durante el cual realizo mis observaciones ($T = 2$ minutos)

$$(\text{valor medio esperado}) = \mu = 3.$$

Este resultado no significa, por supuesto, que debamos esperar observar exactamente tres partículas en ningún ensayo individual. Por el contrario, las probabi-

lidades para la observación de un número (v) de dichas partículas vienen dadas por la distribución de Poisson

$$\text{Prob}(v \text{ partículas}) = P_3(v) = e^{-3} \frac{3^v}{v!}.$$

En particular, la probabilidad de que observe exactamente tres partículas es

$$\text{Prob}(3 \text{ partículas}) = P_3(3) = e^{-3} \frac{3^3}{3!} = 0,22 = 22\%.$$

Observe que a pesar de que el número medio es $\mu = 3$, esperaríamos obtener exactamente este número una vez cada cinco ensayos.

Las probabilidades de cualquier valor de v pueden calcularse de la misma forma y (como el lector puede comprobar) son:

Número v :	0	1	2	3	4
Probabilidad:	5%	15%	22%	22%	17%

Estas probabilidades (hasta $v = 9$) se representan en la figura 11.1. La forma más sencilla de calcular la probabilidad de registrar 5 o más sucesos consiste en sumar las probabilidades de $v = 0, \dots, 4$ y luego restar dicha suma del 100% para dar²

$$\text{Prob}(v \geq 5) = 100\% - (5 + 15 + 22 + 22 + 17)\% = 19\%.$$

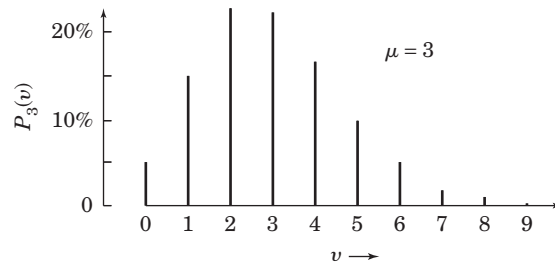


Figura 11.1. La distribución de Poisson $P_3(v)$ proporciona la probabilidad de observar v sucesos de un experimento de recuento cuyo valor medio es igual a 3.

Test rápido 11.1. En promedio, cada una de las 18 gallinas de un gallinero pone 1 huevo por día. Si reviso el gallinero una vez cada hora y retiro los huevos que han puesto durante esta hora, ¿cuál es el promedio, μ , de huevos encontrados en mis visitas? Utilice la distribución de Poisson $P_\mu(v)$ para calcular las probabilidades de encontrar v huevos, para $v = 0, 1, 2, 3$ y $v = 4$ o más. ¿Cuál es el número de huevos más probable? ¿Cuál es la probabilidad de encontrar exactamente μ huevos? Verifique las probabilidades que se muestran en la figura 11.2.

2. El valor correcto es en realidad 18,48%, como el lector puede comprobar si tiene en cuenta un par más de dígitos decimales en el cálculo de todas las probabilidades.

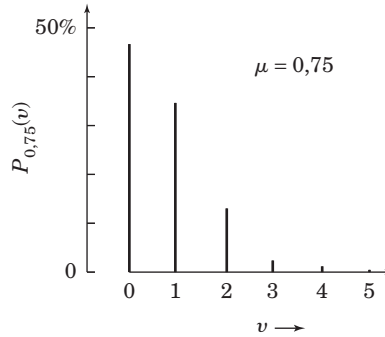


Figura 11.2. La distribución de Poisson $P_{0,75}(v)$ proporciona la probabilidad de observar v sucesos en un experimento de recuento cuyo valor esperado sea 0,75.

11.2 Propiedades de la distribución de Poisson

LA DESVIACIÓN ESTÁNDAR

La distribución de Poisson $P_{\mu}(v)$ proporciona la probabilidad de obtener el resultado v en un experimento en el que contamos sucesos que se producen al azar, pero con una determinada tasa media. Hemos visto que el parámetro μ es precisamente el número promedio esperado \bar{v} . La siguiente cuestión natural es preguntarse por la desviación estándar del número de sucesos contados v cuando se repite el experimento muchas veces. La desviación estándar de una distribución (después de un gran número de ensayos) es la raíz cuadrada de la desviación cuadrática media de la media. Es decir,

$$\sigma_v^2 = \overline{(v - \bar{v})^2},$$

o, usando los resultados de los problemas 10.15(a) o 4.5(a),

$$\sigma_v^2 = \overline{v^2} - (\bar{v})^2. \quad (11.7)$$

Para la distribución de Poisson, ya hemos encontrado que $\bar{v} = \mu$ y un cálculo parecido (problema 11.9) proporciona $\overline{v^2} = \mu^2 + \mu$. Por tanto, la ecuación (11.7) implica que $\sigma_v^2 = \mu$ o bien

$$\sigma_v = \sqrt{\mu}. \quad (11.8)$$

Es decir, la distribución de Poisson cuyo valor medio es μ tiene desviación estándar $\sqrt{\mu}$.

El resultado (11.8) justifica la regla de la raíz cuadrada de la sección 3.2. Si hacemos un experimento de recuento una vez y obtenemos la respuesta v , podemos ver fácilmente (utilizando el principio de máxima verosimilitud, como en el problema 11.11) que la mejor estimación del valor medio del recuento es $\mu_{\text{mejor}} = v$. De (11.8), se deduce inmediatamente que la mejor estimación de la desviación

Capítulo 12

El test ji-cuadrado para una distribución

A estas alturas, el lector debería estar razonablemente familiarizado con el concepto de distribución límite. Estas son las funciones que describen la distribución esperada de los resultados de un experimento si se repite muchas veces. Hay muchas distribuciones límite diferentes, correspondientes a los innumerables diferentes tipos de experimentos posibles. Tal vez las tres distribuciones límite más importantes para las ciencias físicas son las tres que ya hemos discutido: la función de Gauss (o normal), la distribución binomial y la distribución de Poisson.

Este último capítulo se centra en cómo comprobar si los resultados de un experimento real se rigen por la distribución límite prevista. En concreto, suponga que llevamos a cabo un experimento para el que creemos que conocemos la distribución esperada de los resultados. Suponga, además, que repetimos el experimento varias veces y registramos nuestras observaciones. La pregunta que ahora nos hacemos es la siguiente: ¿cómo podemos decidir si nuestra distribución observada es coherente con la distribución teórica esperada? Veremos que a esta pregunta se puede responder mediante un procedimiento simple llamado el test *ji-cuadrado* o χ^2 . (El nombre de letra griega χ se escribe y pronunciada como “ji”).

12.1 Introducción a la ji cuadrado

Comencemos con un ejemplo concreto. Suponga que hacemos 40 mediciones x_1, \dots, x_{40} del alcance x de un proyectil disparado por una cierta arma y obtenemos los resultados que se muestran en la tabla 12.1. Suponga también que tenemos razones para creer que estas mediciones se rigen por una distribución de Gauss $G_{x,\sigma}(x)$, como es por cierto muy natural. En este tipo de experimento, por lo general no sabemos de antemano cuál es el valor central X o el parámetro de anchura σ de la distribución esperada. Nuestro primer paso, por lo tanto, es el

Tabla 12.1. Valores medidos de x (en cm).

731	772	771	681	722	688	653	757	733	742
739	780	709	676	760	748	672	687	766	645
678	748	689	810	805	778	764	753	709	675
698	770	754	830	725	710	738	638	787	712

uso de nuestras 40 mediciones para calcular las mejores estimaciones de estas cantidades:

$$(\text{mejor estimación de } X) = \bar{x} = \frac{\sum x_i}{40} = 730,1 \text{ cm} \quad (12.1)$$

y

$$(\text{mejor estimación de } \sigma) = \sqrt{\frac{\sum (x_i - \bar{x})^2}{39}} = 46,8 \text{ cm.} \quad (12.2)$$

Ahora podemos preguntarnos si la distribución real de nuestros resultados x_1, \dots, x_{40} es coherente con nuestra hipótesis de que estas mediciones se rigen por la distribución de Gauss $G_{x,\sigma}(x)$ con X y σ según las estimaciones. Para responder a esta pregunta, tenemos que calcular cómo esperaríamos que estuvieran distribuidos nuestros 40 resultados si la hipótesis fuera verdadera y comparar esta distribución esperada con nuestra distribución observada realmente. La primera dificultad es que x es una variable continua, por lo que no podemos hablar del número esperado de mediciones iguales a cualquier valor de x . Más bien, debemos discutir el número esperado en algún intervalo $a < x < b$. Es decir, hay que dividir el rango de valores posibles en *clases*. Con 40 mediciones, podríamos elegir los límites de las clases en $\bar{X} - \sigma$, \bar{X} y $\bar{X} + \sigma$, lo que da las cuatro clases de la tabla 12.2.

Hablaremos más tarde de los criterios para la elección de clases. En particular, se deben elegir de modo que todas ellas contengan varios valores observados x_i . En general, indicaremos el número de clases por n ; para este ejemplo con cuatro clases, $n = 4$.

Después de haber dividido el rango de posibles valores medidos en las diversas clases, podemos formular nuestra pregunta con mayor precisión. En primer lugar, podemos contar el número de mediciones correspondientes a cada clase k .¹ Denotamos este número por O_k (donde O significa “número observado”). Para los datos de nuestro ejemplo, los números observados O_1, O_2, O_3, O_4 se muestran en la última columna de la tabla 12.2. A continuación, suponiendo que nuestras mediciones se distribuyen normalmente (con X y σ según las estimaciones obtenidas), se puede calcular el número *esperado* E_k de mediciones en cada clase k . Por tanto, debemos decidir si los números observados O_k son suficientemente parecidos a los números esperados E_k .

Tabla 12.2. Una posible elección de las clases para los datos de la tabla 12.1. La última columna muestra la frecuencia de las observaciones en cada clase.

Número de la clase, k	Valores de x en la clase		Observaciones O_k
1	$x < \bar{X} - \sigma$	(o $x < 683,3$)	8
2	$\bar{X} - \sigma < x < \bar{X}$	(o $683,3 < x < 730,1$)	10
3	$\bar{X} < x < \bar{X} + \sigma$	(o $730,1 < x < 776,9$)	16
4	$\bar{X} + \sigma < x$	(o $776,9 < x$)	6

1. Si una medición cae exactamente en el límite entre dos clases, podemos asignar la mitad de dicha medición a cada clase.

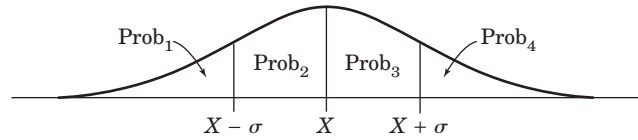


Figura 12.1. Las probabilidades Prob_k de que una medición caiga dentro de cada clases, $k = 1, 2, 3, 4$, de la tabla 12.2 son las cuatro áreas mostradas en la función de Gauss.

El cálculo del número esperado E_k es bastante sencillo. La probabilidad de que cualquier medición se encuentre en un intervalo $a < x < b$ es simplemente el área bajo la función de Gauss entre $x = a$ y $x = b$. En este ejemplo, las probabilidades $\text{Prob}_1, \text{Prob}_2, \text{Prob}_3, \text{Prob}_4$ de que una medición caiga en cada una de nuestras cuatro clases coinciden con las cuatro áreas indicadas en la figura 12.1. Las dos áreas iguales Prob_2 y Prob_3 juntas representan el bien conocido 68%, por lo que la probabilidad de que caiga en una de las dos clases centrales es del 34%; es decir, $\text{Prob}_2 = \text{Prob}_3 = 0,34$. Las dos áreas iguales más externas entre las dos comprenden el 32% restante, por lo que $\text{Prob}_1 = \text{Prob}_4 = 0,16$. Para encontrar el número esperado E_k , simplemente multiplicaremos estas probabilidades por el número total de mediciones, $N = 40$. Por tanto, nuestros números esperados son los que se muestran en la tercera columna de la tabla 12.3. Que los números E_k no son números enteros sirve para recordarnos que “el número esperado” no es el resultado que realmente esperamos en cualquier experimento, sino que es más bien el número promedio esperado después de repetir toda nuestra serie de mediciones en muchas ocasiones.

Nuestro problema ahora es decidir hasta qué punto el número esperado E_k se aproxima a los números observados correspondientes O_k (mostrados en la última columna de la tabla 12.3). Obviamente, no cabe esperar una coincidencia *perfecta* entre E_k y O_k después de cualquier número finito de mediciones. Por otro lado, si nuestra hipótesis de que las mediciones se distribuyen normalmente es correcta, cabría esperar que, en cierto sentido, las desviaciones

$$O_k - E_k \quad (12.3)$$

fueran *pequeñas*. Por el contrario, si las desviaciones $O_k - E_k$ demuestran ser *grandes*, sospecharíamos que nuestra hipótesis es incorrecta.

Tabla 12.3. El número esperado E_k y los números observados O_k para las 40 mediciones de la tabla 12.1, con clases elegidas como en la tabla 12.2.

Número de la clase k	Probabilidad Prob_k	Número esperado $E_k = N\text{Prob}_k$	Número observado O_k
1	16%	6,4	8
2	34%	13,6	10
3	34%	13,6	16
4	16%	6,4	6

Para precisar la afirmación de que la desviación $O_k - E_k$ es “pequeña” o “grande”, tenemos que determinar qué tan grande *esperaríamos* que fuera $O_k - E_k$ si las mediciones realmente siguieran una distribución normal. Afortunadamente, esta determinación es fácil de realizar. Si nos imaginamos repitiendo toda nuestra serie de 40 mediciones muchas veces, a continuación, el número de mediciones O_k en cualquier clase k puede ser considerada como el resultado de un experimento de recuento del tipo descrito en el capítulo 11. Los diferentes valores de O_k deben tener un valor medio E_k y se espera que fluctúen alrededor de E_k con una desviación estándar de orden $\sqrt{E_k}$. Por tanto, los dos números que deben ser comparados son la desviación $O_k - E_k$ y el tamaño esperado de sus fluctuaciones $\sqrt{E_k}$.

Estas consideraciones nos llevan a considerar el cociente

$$\frac{O_k - E_k}{\sqrt{E_k}}. \quad (12.4)$$

Para algunas clases k , este cociente será positivo, y para otras será negativo; para unos pocos valores de k , puede ser sensiblemente mayor que uno, pero la mayoría, en valor absoluto, debería ser de orden uno o menor. Para probar nuestra hipótesis (que las mediciones tienen una distribución normal), es natural elevar al cuadrado el número (12.4) para cada k y luego sumar el resultado de todas las clases $k = 1, \dots, n$ (en este caso $n = 4$). Este procedimiento define un número llamado *ji cuadrado*,

$$\chi^2 = \sum_{k=1}^n \frac{(O_k - E_k)^2}{E_k}. \quad (12.5)$$

Este número χ^2 es claramente un indicador razonable del grado de ajuste entre la distribución observada y la esperada. Si $\chi^2 = 0$, el acuerdo es perfecto; es decir, $O_k = E_k$ para todas las clases k , una situación sumamente improbable. En general, se espera que los términos individuales de la suma (12.5) sean de orden uno, y hay n términos en dicha suma. Por tanto, si

$$\chi^2 \leq n$$

(χ^2 de orden n o menor), las distribuciones observadas y esperadas son tan parecidas como podíamos esperar. En otras palabras, si $\chi^2 \leq n$, no tenemos ninguna razón para dudar de que nuestras mediciones se distribuyeron tal como esperábamos. Por otro lado, si

$$\chi^2 \gg n$$

(χ^2 significativamente mayor que el número de clases), los números observados y esperados difieren significativamente, y tenemos buenas razones para sospechar que nuestras mediciones no se regían por la distribución esperada.

En nuestro ejemplo, los números observados y esperados en las cuatro clases y sus diferencias se muestran en la tabla 12.4, y un cálculo simple utilizándolos da

$$\begin{aligned} \chi^2 &= \sum_{k=1}^4 \frac{(O_k - E_k)^2}{E_k} \\ &= \frac{(1,6)^2}{6,4} + \frac{(-3,6)^2}{13,6} + \frac{(2,4)^2}{13,6} + \frac{(-0,4)^2}{6,4} \\ &= 1,80. \end{aligned} \quad (12.6)$$

Tabla 12.4. Los datos de la tabla 12.1, mostrados aquí con las diferencias $O_k - E_k$.

Número de la clase k	Número observado O_k	Número esperado $E_k = N\text{Prob}_k$	Diferencia $O_k - E_k$
1	8	6,4	1,6
2	10	13,6	-3,6
3	16	13,6	2,4
4	6	6,4	-0,4

Debido a que el valor de 1,80 para χ^2 es menor que el número de términos de la suma (es decir, 4), no tenemos ninguna razón para dudar de la hipótesis de que nuestras mediciones se distribuyen normalmente.

Test rápido 12.1. Cada uno de los 100 estudiantes de una clase mide el tiempo que tarda una bola en caer desde una ventana del tercer piso. Seguidamente calculan sus medias \bar{t} y desviación estándar σ_t y luego agrupan sus mediciones en cuatro clases, elegidas como en el ejemplo que acabamos de discutir. Sus resultados son los siguientes:

menores que $(\bar{t} - \sigma_t)$: 19
entre $(\bar{t} - \sigma_t)$ y (\bar{t}) : 30
entre \bar{t} y $(\bar{t} + \sigma_t)$: 37
mayores que $(\bar{t} + \sigma_t)$: 14.

Suponiendo que sus mediciones se distribuyen normalmente, ¿cuáles son los números esperados de las mediciones en cada una de las cuatro clases? ¿Cuál es el valor de la χ^2 ? ¿Hay razones para dudar de que las mediciones se distribuyan normalmente?

12.2 Definición general de la ji cuadrado

La discusión hasta ahora se ha centrado en un ejemplo particular, 40 mediciones de una variable continua x , que representa el alcance de un proyectil disparado desde un arma determinada. Hemos definido el número χ^2 y vimos que como mínimo es una medida aproximada del acuerdo entre nuestra distribución observada de las mediciones y la distribución esperada de éstas, la distribución de Gauss. Ahora podemos definir y utilizar el valor de la χ^2 de la misma manera para muchos experimentos diferentes.

Consideremos cualquier experimento en el que podemos determinar un número x para el cual tenemos razones para esperar una cierta distribución de los resultados. Nos imaginamos repitiendo la medición varias veces (N) y, después de haber dividido el intervalo de los posibles resultados de x en n clases, $k = 1, \dots, n$, contamos el número O_k de observaciones que realmente caen en cada clase k . Suponiendo que las mediciones realmente se rigen por la distribución esperada,

Apéndice A

Integral de error normal, I

Si la medición de una variable x continua está sujeta a muchos errores pequeños, todos ellos al azar, la distribución esperada de los resultados es la distribución normal o de Gauss dada por,

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

donde X es el verdadero valor de x , y σ es la desviación estándar.

La integral de la función de Gauss $\int_a^b G_{X,\sigma}(x) dx$, se denomina *integral de error normal*, y es igual a la probabilidad de que una medición caiga entre $x = a$ y $x = b$,

$$\text{Prob}(a \leq x \leq b) = \int_a^b G_{X,\sigma}(x) dx.$$

La tabla A muestra los valores de dicha integral para $a = X - t\sigma$ y $b = X + t\sigma$. Esta es igual a la probabilidad de que una medición difiera de X en menos de t veces la desviación estándar,

$$\begin{aligned} \text{Prob}(\text{dentro de } t\sigma) &= \text{Prob}(X - t\sigma \leq x \leq X + t\sigma) \\ &= \int_{X-t\sigma}^{X+t\sigma} G_{X,\sigma}(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-z^2/2} dz. \end{aligned}$$

Esta función a veces se denota por $\text{erf}(t)$, aunque dicha notación se emplea también para indicar una función ligeramente distinta.

La probabilidad de que una medición caiga *fuera* del intervalo antes mencionado puede hallarse fácilmente por sustracción;

$$\text{Prob}(\text{fuera de } t\sigma) = 100\% - \text{Prob}(\text{dentro de } t\sigma).$$

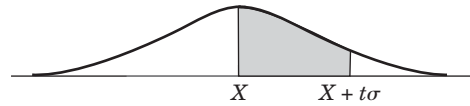
Para más detalles, vea la sección 5.4 y el apéndice B.

Apéndice B

Integral de error normal, II

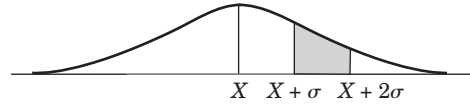
En ciertos cálculos, una forma conveniente de la integral de error normal es

$$\begin{aligned} Q(t) &= \int_X^{X+t\sigma} G_{X,\sigma}(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_0^t e^{-z^2/2} dz. \end{aligned}$$



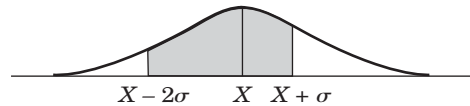
(Esta integral es, naturalmente, exactamente igual a la mitad de la integral tabulada en el apéndice A). La probabilidad $\text{Prob}(a \leq x \leq b)$ de que una medida esté en los límites del cualquier intervalo $a \leq x \leq b$ puede hallarse a partir de $Q(t)$ mediante una sola sustracción o adición. Por ejemplo,

$$\text{Prob}(X + \sigma \leq x \leq X + 2\sigma) = Q(2) - Q(1).$$



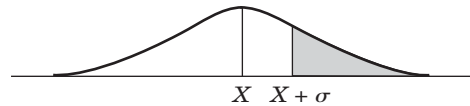
De forma parecida,

$$\text{Prob}(X - 2\sigma \leq x \leq X + \sigma) = Q(2) + Q(1).$$



La probabilidad de que una medida sea mayor que $X + t\sigma$ es simplemente $0,5 - Q(t)$. Por ejemplo,

$$\text{Prob}(x \geq X + \sigma) = 50\% - Q(1).$$



Apéndice C

Probabilidades para los coeficientes de correlación

El coeficiente de correlación lineal es una medida que indica el grado de ajuste de N puntos $(x_1, y_1), \dots, (x_N, y_N)$ a una línea recta

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}},$$

cuyo valor está siempre en el intervalo $-1 \leq r \leq 1$. Los valores de r cercanos a ± 1 indican una alta correlación lineal; valores cercanos a 0 indican poca o nula correlación.

Una medida más cuantitativa del ajuste puede hallarse usando la tabla C. Dado un valor observado cualquiera r_o , $\text{Prob}_N(|r| \geq |r_o|)$ es la probabilidad de que N pares de mediciones de dos variables no correlacionadas nos den un coeficiente r cuyo valor absoluto sea mayor que el de r_o . Por tanto, si obtenemos un coeficiente r_o para el que $\text{Prob}_N(|r| \geq |r_o|)$ sea pequeña, es en consecuencia muy improbable que nuestras variables no estén correlacionadas; y es razonable hablar de variables correlacionadas. En particular, si $\text{Prob}_N(|r| \geq |r_o|) \leq 5\%$, llamaremos a dicha correlación *significativa*; y si es menor que un 1%, llamaremos a dicha correlación *muy significativa*.

Por ejemplo, la probabilidad de que 20 mediciones ($N = 20$) de dos variables no correlacionadas proporcionen un valor $|r| \geq 0,5$ según la tabla es igual a 2,5%. Por tanto, si tras 20 mediciones de pares de valores de dos variables obtenemos $r = 0,5$, tendremos una evidencia *significativa* de correlación lineal entre las dos variables. Para más detalles, vea de las secciones 9.3 a la 9.5.

Los valores de la tabla C han sido calculados a partir de la integral

$$\text{Prob}_N(|r| \geq |r_o|) = \frac{2\Gamma(N-1)/2}{\sqrt{\pi}[\Gamma(N-1)/2]} \int_{|r_o|}^1 (1-r^2)^{(N-4)/2} dr.$$

Vea, por ejemplo, E. M. Pugh y G. H. Winslow. *The Analysis of Physical Measurements* (Addison-Wesley, 1966), sección 12 - 8.

Apéndice D

Probabilidades para la distribución ji cuadrado

Si una serie de mediciones está agrupada en categorías $k = 1, \dots, n$, denotamos por O_k el número de mediciones observadas en la categoría k . El número *esperado* (según alguna distribución de probabilidad supuesta o esperada) en la categoría k se denota por E_k . El grado en que las observaciones se ajustan a la distribución supuesta viene dado por la ji cuadrado reducida, $\tilde{\chi}^2$, definida como

$$\tilde{\chi}^2 = \frac{1}{d} \sum_{k=1}^n \frac{(O_k - E_k)^2}{E_k},$$

donde d es el número de grados de libertad, $d = n - c$, y c es el número de restricciones (vea la sección 12,3). El valor esperado de $\tilde{\chi}^2$ es igual a 1. Si $\tilde{\chi}^2 \gg 1$, los resultados observados no se ajustan a la distribución supuesta; si $\tilde{\chi}^2 \leq 1$ la concordancia es satisfactoria,

Este test se analiza cuantitativamente con las probabilidades mostradas en la tabla D. Sea $\tilde{\chi}_o^2$ el valor de $\tilde{\chi}^2$ realmente obtenido en un experimento con d grados de libertad. El número $\text{Prob}_d(\tilde{\chi}^2 \geq \tilde{\chi}_o^2)$ es la probabilidad de obtener un valor de $\tilde{\chi}^2$ tan grande como el observado $\tilde{\chi}_o^2$, si la probabilidad de los resultados está basada realmente en la distribución que estamos suponiendo verdadera. Por tanto, si $\text{Prob}_d(\tilde{\chi}^2 \geq \tilde{\chi}_o^2)$ es grande, los resultados observados son compatibles con la distribución supuesta como verdadera; si es pequeña, hay evidencias de lo contrario. En particular, si $\text{Prob}_d(\tilde{\chi}^2 \geq \tilde{\chi}_o^2)$ es menor que el 5%, diremos que el desacuerdo es *significativo* y rechazaremos la distribución supuesta al nivel del 5%. Si es menor que el 1%, diremos que el desacuerdo es *altamente significativo*, y rechazaremos la distribución supuesta al nivel del 1%.

Por ejemplo, supongamos que obtenemos un valor de una ji cuadrado reducida igual a 2,6 (esto es, $\tilde{\chi}^2 = 2,6$) en un experimento con seis grados de libertad ($d = 6$). Según la tabla D, la probabilidad de obtener $\tilde{\chi}^2 \geq 2,6$ es 1,6%, si los resultados estuvieran determinados según la distribución supuesta. Por tanto, al nivel del 5% rechazaríamos la distribución asumida (pero no hay suficiente discrepancia para hacerlo al nivel del 1%). Para más detalles, vea el capítulo 12.

Apéndice E

Dos demostraciones relativas a las desviaciones estándar muestrales

En el capítulo 5 se citaron sin demostración dos resultados importantes en relación con la medición de una magnitud x distribuida normalmente con un parámetro de anchura σ : (1) La mejor estimación de σ basada en N mediciones de x es la desviación estándar de la muestra, σ_x , de las N mediciones, tal como se define por (5.45) con el factor de $(N - 1)$ en el denominador. (2) La incertidumbre relativa de σ_x como estimación de σ es $1/\sqrt{2(N - 1)}$, como en (5.46). Las demostraciones de estos dos resultados son sorprendentemente pesadas y se omitieron en el capítulo 5. Para aquellos a quienes les gusta ver las demostraciones, las proporcionaremos aquí.

Consideremos un experimento que consiste en realizar N mediciones (todas usando el mismo método) de cierta cantidad x , con los resultados

$$x_1, x_2, \dots, x_N.$$

La desviación estándar muestral de estas N medidas se define como

$$\begin{aligned}\sigma_x^2 &= \frac{\sum (x_i - \bar{x})^2}{N - 1} \\ &= \frac{SC}{N - 1},\end{aligned}\tag{E.1}$$

donde hemos introducido la notación SC para referirnos a la *Suma de Cuadrados*,

$$\begin{aligned}SC &= \text{Suma de Cuadrados} \\ &= \sum_1^N (x_i - \bar{x})^2 \\ &= \sum_1^N (x_i)^2 - \frac{1}{N} \left(\sum_1^N x_i \right)^2.\end{aligned}\tag{E.2}$$

Al escribir la última línea de (E.2), hemos utilizado la identidad (4.28) del problema 4.5. La ecuación (E.1) en realidad define σ_x al *cuadrado*, llamada la *varianza muestral* de las N mediciones. Para evitar tener numerosos símbolos de raíces cuadradas, vamos a trabajar sobre todo con la varianza en lugar de la propia desviación estándar.

Nuestras demostraciones se simplificarán señalando que ni el verdadero parámetro de anchura σ ni la verdadera desviación estándar muestral σ_x cambian si restamos cualquier constante fija a nuestras mediciones x . En particular, podemos restar a x su verdadero valor X , que proporciona una cantidad normalmente

Respuestas a los tests rápidos y a los problemas de numeración impar

Capítulo 2

TESTS RÁPIDOS

TR2.1. (a) 110 ± 2 mm. (b) Entre 3,02 y 3,08 amperios.

TR2.2. (a) $8,12 \pm 0,03$ m/s. (b) $31\,234 \pm 2$ m o $(3,1234 \pm 0,0002) \times 10^4$ m. (c) $(5,68 \pm 0,03) \times 10^{-7}$ kg.

TR2.3. 43 ± 5 gramos.

TR2.4. (a) $0,04 = 4\%$. (b) $0,1 = 10\%$. (c) $4,58 \pm 0,09$ J.

TR2.5. Las incertidumbres porcentuales son 1% y 3%; área = $30\text{ cm}^2 \pm 4\% = 30 \pm 1\text{ cm}^2$.

PROBLEMAS

2.1. 210 ± 5 cm; $36,0 \pm 0,5$ mm; $5,3 \pm 0,1$ V; $2,4 \pm 0,1$ s.

2.3. (a) $5,03 \pm 0,04$ m. (b) Hay un buenas razones para retener un dígito extra, indicándolo $1,5 \pm 1$ s. (c) $(-3,2 \pm 0,3) \times 10^{-19}$ C. (d) $(5,6 \pm 0,7) \times 10^{-7}$ m. (e) $(3,27 \pm 0,04) \times 10^3$ gramos · cm/s.

2.5. Discrepancia = 2 mm, lo cual *no* es significativo.

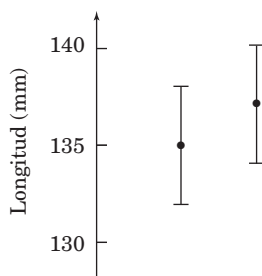


Figura A2.5

2.7. (a) Probablemente la única conclusión razonable en este punto es $1,9 \pm 0,1$ gramos/cm³. (b) La discrepancia es 0,05 gramos/cm³; puesto que este valor es menor que la incertidumbre, *no* es significativo.

2.9. Longitud, $l = 93,5 \pm 0,1$ cm.