



# Anormalidad

## PALABRAS CLAVE

Datos nominales	Constructos	Sensibilidad
Datos dicotómicos	Escalas	Interpretabilidad
Datos ordinales	Validez del contenido	Fracción de muestreo
Datos de intervalo	Validez del constructo	Distribución de frecuencia
Datos continuos	Validez de criterio	Tendencia central
Datos discretos	Fiabilidad	Dispersión
Validez	Reproducibilidad	Asimetría
Exactitud	Precisión	Distribución normal
Apartados	Amplitud	Regresión a la media

*... el significado médico de «normal» se ha perdido en la gran baraja de la estadística.*

*– Alvan Feinstein*

Los médicos dedican una gran parte del tiempo a distinguir entre lo que es «normal» y lo que es «anómalo». El tamaño de la glándula tiroidea, ¿es normal o está ligeramente aumentado? El soplo cardíaco, ¿es inocente (sin importancia para la salud) o es un signo de valvulopatía? Una concentración sérica de fosfatasa alcalina ligeramente elevada, ¿es un indicio de hepatopatía, de enfermedad de Paget asintomática, o es un hecho sin importancia?

Cuando los médicos se enfrentan a un caso claramente diferente de lo habitual, no tienen dificultades para distinguirlo de los demás. Todos hemos visto imágenes en los libros de texto de diagnósticos clínicos que muestran grandes hepatoesplenomegalias, enormes bocios o manos gravemente deformadas por la artritis reumatoide. No es necesaria

una gran habilidad para reconocer este grado de anomalía. Sin embargo, lo más frecuente es que los médicos tengan que realizar distinciones más sutiles entre lo normal y lo anómalo. Es entonces cuando adquieren importancia la habilidad y una base conceptual para decidir.

Las decisiones sobre lo que es anómalo son especialmente difíciles entre los pacientes relativamente no seleccionados, que generalmente suelen encontrarse fuera de los hospitales. Cuando los pacientes ya han sido seleccionados para recibir una atención especial, como ocurre en la mayoría de los centros de remisión, habitualmente, suele estar claro que algo va mal. En tal caso, las tareas consisten en precisar el diagnóstico y tratar el problema. No obstante, en los ámbitos de atención primaria y en los

servicios de urgencias, los pacientes con manifestaciones patológicas sutiles se mezclan con los pacientes que refieren las afecciones cotidianas de la gente que básicamente está sana. Es imposible hacer un seguimiento activo de todas estas afecciones. De los muchos pacientes con dolor abdominal, ¿quiénes sufren una gastroenteritis que se resolverá de forma espontánea, y quiénes una apendicitis inicial? De los pacientes con dolor de garganta y ronquera, ¿quiénes sufren una faringitis vírica y quiénes una rara, pero potencialmente mortal, epiglotitis por *Haemophilus influenzae*? Éstos son ejemplos de lo difícil e importante que puede llegar a ser la distinción entre los diferentes tipos de anomalías.

El objetivo de distinguir entre lo normal y lo anómalo es separar las observaciones clínicas que deben constituir la base para la actuación de las que sencillamente pueden quedar registradas. Las observaciones que se consideran normales suelen describirse como «dentro de los límites normales», «insignificantes» o «sin contribución», y quedan enterradas en la historia clínica del paciente. Los hallazgos anómalos se incluyen en una lista de problemas o bajo títulos como «Impresiones» o «Diagnósticos», y conforman la base para la actuación.

Calificar los hallazgos clínicos sencillamente como normales o anómalos denota una clara falta de precisión y da lugar a clasificaciones erróneas. El motivo de estos enfoques es que, a menudo, no resulta práctico o es innecesario considerar los datos brutos con todo detalle. Como afirmó Bertrand Russell: «Para ser perfectamente inteligible, uno tiene que ser inexacto, y para ser perfectamente exacto, uno tiene que ser ininteligible». Habitualmente, los médicos prefieren inclinarse por la inteligibilidad (para ellos mismos y para los demás médicos), incluso a expensas de cierta exactitud. Otra razón para simplificar los datos es que cada aspecto del cometido de un médico termina en una decisión: seguir con la evaluación o esperar, empezar el tratamiento o tranquilizar al paciente. En estas circunstancias, es necesario algún tipo de clasificación como «presente» o «ausente».

La tabla 2-1 es un ejemplo de cómo las expresiones relativamente sencillas de anormalidad proceden de datos clínicos más complejos. A la izquierda, figura una lista de problemas habituales, un enunciado de los problemas médicos importantes para el paciente. A la derecha, se exponen algunos

de los datos utilizados como base para tomar la decisión de etiquetarlos como problemas. Las conclusiones de los datos, representadas por la lista de problemas, son por supuesto objetables. Por ejemplo, la media de las cuatro determinaciones de la tensión arterial diastólica es de 92 mm Hg. Puede que algunos afirmen que tal nivel de tensión arterial no justifica la calificación de «hipertensión arterial», ya que no se trata de una cifra especialmente elevada, y el hecho de informar a los pacientes de que están enfermos, así como de recetar fármacos, presenta algunos inconvenientes. Puede que otros consideren que la calificación es adecuada, si se tiene en cuenta que este nivel de tensión arterial se asocia a un aumento del riesgo de enfermedad cardiovascular, y que dicho riesgo puede reducirse mediante tratamiento. Aunque poco precisa, la lista de problemas sirve como base para tomar decisiones (sobre el diagnóstico, el pronóstico y el tratamiento). Y deben tomarse decisiones clínicas, ya sea de forma activa (mediante pruebas diagnósticas adicionales y tratamiento) o de forma pasiva (sin intervención).

Este capítulo describe algunos de los métodos que utilizan los médicos para distinguir entre lo normal y lo anómalo. Para ello, ante todo, será necesario considerar cómo se miden, cómo varían y cómo se resumen los fenómenos biológicos. Posteriormente, se podrá considerar cómo se utilizan estos datos como base de los juicios de valor que permiten determinar qué merece la pena denominar anómalo.

**TABLA 2-1. ■ Resumen de datos clínicos: lista de problemas de un paciente y datos en los que éstos se basan**

<b>Lista de problemas</b>	<b>Datos brutos</b>
Hipertensión	Diferentes mediciones de la tensión arterial: 145/92, 149/93, 142/91
Diabetes mellitus	Glucemia en ayunas 135 mg/dl
Insuficiencia renal	Creatinina sérica, 2,7 mg/dl
Enfermedad pulmonar obstructiva	Volumen espiratorio forzado en un segundo (FEV <sub>1</sub> ) 2,3 l

## TIPOS DE DATOS

Las mediciones de los fenómenos clínicos dan lugar a tres tipos de datos: nominales, ordinales y de intervalo.

### Datos nominales

Los **datos nominales** se producen en categorías sin ningún orden natural. Las características determinadas por un pequeño grupo de genes (como el grupo sanguíneo ABO o el sexo) o que constituyen episodios aislados, pero espectaculares (como la muerte, la diálisis o una intervención quirúrgica), son ejemplos de datos nominales. Dichos datos pueden incluirse en diferentes categorías sin preocuparse demasiado por clasificarlos de forma errónea. Los datos nominales que se dividen en dos categorías (como presente/ausente, sí/no, vivo/muerto) se denominan **dicotómicos**.

### Datos ordinales

Los **datos ordinales** poseen cierto orden natural, como por ejemplo, todo lo situado entre pequeño y grande, o entre bueno y malo, si bien el tamaño de los intervalos entre categorías no tiene por qué ser definido o detallado. Algunos ejemplos clínicos son el edema de las extremidades inferiores entre 1+ y 4+, los soplos cardíacos entre los grados I (audible sólo tras un esfuerzo especial) y VI (audible sin necesidad de fonendoscopio), y la fuerza muscular entre los grados 1 (sin movimiento) y 5 (fuerza normal). La FDA (*Food and Drug Administration*, Agencia Estadounidense del Medicamento) ha clasificado el riesgo de sufrir anomalías congénitas a causa de la utilización de fármacos durante el embarazo en una escala que consta de cinco categorías: desde «los estudios clínicos comparativos muestran que no hay riesgo», pasando por «no hay indicios de riesgo para los seres humanos», «no puede descartarse el riesgo» e «indicios positivos de riesgo», hasta «contraindicados durante el embarazo».

### Datos de intervalo

Para los **datos de intervalo**, existe un orden natural y el intervalo entre valores sucesivos es idéntico, independientemente de su situación dentro de la

escala. Existen dos tipos de datos de intervalo. Los **datos continuos** pueden adquirir cualquier valor dentro de un intervalo continuo, independientemente de la forma en que se notifiquen. La mayor parte de la bioquímica sérica, el peso, la tensión arterial y la presión parcial del oxígeno en sangre arterial son algunos ejemplos de datos continuos. En la práctica, la medición y la descripción de las variables continuas pueden limitarse a un número determinado de puntos dentro del intervalo continuo, a menudo representados por números enteros, ya que la precisión de la medición o su utilización no justifican un mayor detalle. Por ejemplo, la lectura de una medición concreta de la glucemia puede ser de 193,2846573... mg/100 ml, pero se documenta sencillamente como 193 mg/100 ml. Los **datos discretos** sólo pueden adquirir valores determinados, y se expresan como recuentos. El número de embarazos y partos de una mujer, así como el número de crisis de migraña que sufre un paciente al mes, son ejemplos de datos discretos.

Precisamente, es para los datos ordinales y de intervalo cuando surge la pregunta: «¿Dónde termina lo normal y dónde empieza lo anómalo?» Por ejemplo, ¿en qué momento puede afirmarse que una próstata normal de tamaño grande ha crecido lo bastante como para dejar de considerarse normal? Los médicos son libres de elegir cualquier punto de corte. Algunas de las razones para estas elecciones se tratarán más adelante en este capítulo.

## RENDIMIENTO DE LAS MEDICIONES

Independientemente del tipo de medición, su rendimiento puede describirse de varias formas.

### Validez

La **validez** es el grado en que los datos miden aquello que pretendían medir, es decir, hasta qué punto los resultados de la medición se corresponden con el estado real del fenómeno medido. Un sinónimo de validez es **exactitud**.

Resulta relativamente fácil comprobar la validez de las observaciones clínicas que pueden medirse por medios físicos. La medición observada se compara con alguna referencia aceptada. Por ejemplo, el sodio sérico puede medirse mediante un instru-

mento calibrado recientemente con unas soluciones con concentraciones conocidas de sodio. Las mediciones analíticas suelen someterse a amplias y repetidas comprobaciones para determinar su validez. Por ejemplo, es una práctica habitual controlar la exactitud de las mediciones de la glucemia comparando las lecturas con referencias elevadas y bajas al principio de cada día, antes de que cada técnico empiece su jornada, y después de cualquier cambio en los procedimientos, como un nuevo frasco de reactivo o una nueva pila para el instrumento. De igual modo, puede medirse la exactitud de una gammagrafía pulmonar realizada para detectar un émbolo pulmonar, comparándola con una angiografía pulmonar en la que se visualiza directamente la anatomía de las arterias pulmonares. La validez de un hallazgo durante una exploración física puede comprobarse al compararlo con los resultados de la cirugía o de las exploraciones radiológicas.

Existen otras mediciones clínicas, como el dolor, las náuseas, la disnea, la depresión y el miedo, que no pueden comprobarse de forma física. En las consultas, la información relacionada con estos fenómenos suele obtenerse de modo informal a través de la anamnesis, aunque existen otros métodos más formales y normalizados que se utilizan en la investigación, como las entrevistas estructuradas y los cuestionarios. En este caso, se diseñan preguntas concretas (**apartados**) para medir fenómenos determinados (como síntomas, sensaciones, actitudes, conocimientos, creencias) denominados **constructos**, y dichos apartados se agrupan en **escalas**. La tabla 2-2 muestra una de estas escalas; se trata de un breve cuestionario utilizado para detectar el consumo excesivo o abusivo de alcohol.

Para comprobar la validez de las mediciones que no pueden verificarse directamente mediante los sentidos físicos, se utilizan tres estrategias generales:

### Validez del contenido

La **validez del contenido** es el grado en que un método concreto de medición incluye todas las dimensiones del constructo que se pretende medir y nada más. Por ejemplo, una escala para medir el dolor tendrá validez de contenido si incluye preguntas sobre dolor sordo, dolor pulsátil, dolor urente y escozor, pero no sobre tensión, prurito, náuseas y hormigueo.

**TABLA 2-2. ■ Cuestionario CAGE para la detección del consumo excesivo o abusivo de alcohol**

- ¿Ha sentido alguna vez la necesidad de dejar de beber?
  - ¿Le ha molestado que la gente haya criticado su forma de beber?
  - ¿Se ha sentido mal o culpable por su forma de beber?
  - ¿Ha bebido alguna vez a primera hora de la mañana para calmar los nervios o evitar una resaca?
- Una respuesta «afirmativa» sugiere la necesidad de una valoración más detenida.
- Un mínimo de dos respuestas «afirmativas» expresan una estrecha relación con el alcoholismo agudo, el alcoholismo crónico, o con ambos.

### Validez del constructo

Existe **validez del constructo** cuando la medición está relacionada de forma coherente con otras mediciones, que, al igual que ésta, no pueden comprobarse físicamente, y que se consideran parte del mismo fenómeno. Así, por ejemplo, será más fiable la validez del constructo de una escala que mide la depresión cuando ésta esté relacionada con la fatiga y la cefalea (otros constructos que se consideran diferentes pero relacionados con la depresión).

### Validez de criterio

Existe **validez de criterio** cuando las mediciones predicen un fenómeno que se puede observar directamente. Por ejemplo, es posible verificar si las respuestas en una escala que mide el dolor guardan una relación predecible con el dolor de una intensidad conocida: dolor leve producido por una abrasión sin importancia, dolor moderado causado por una cefalea común y una úlcera péptica, y dolor intenso debido a un cólico nefrítico. También es posible observar que las respuestas a una escala que mide el dolor están relacionadas con otras manifestaciones observables vinculadas a la intensidad del dolor, como sudoración, gemidos y el hecho de solicitar fármacos contra el dolor.

La validez de una escala no es algo que esté, como se afirma a menudo, presente o ausente. Por el contrario, con estas estrategias, es posible adoptar una posición a favor o en contra de la validez de una escala en las condiciones en las que se utiliza,

de forma que se pueda convencer a los demás de que la escala es más o menos válida.

Debido a su selección y formación, los médicos tienden a preferir el tipo de mediciones precisas que permiten las ciencias físicas y biológicas, y pueden evitar o descartar las otras, especialmente para la investigación. A pesar de todo, el alivio de los síntomas y el fomento de la satisfacción y de la sensación de bienestar figuran entre los resultados más importantes de la asistencia a los pacientes, y constituyen las preocupaciones centrales tanto de éstos como de los médicos. La investigación también debe incluir dichas preocupaciones, a fin de guiar las decisiones clínicas y evitar que la propia investigación distorsione la imagen que ella misma ha dibujado de la medicina.

Tal como Feinsein afirmó (1):

El término «sólido» suele aplicarse a los datos que son fiables y preferiblemente dimensionales (p. ej., datos analíticos, datos demográficos y costes económicos). Pero el rendimiento clínico, la conveniencia, la anticipación y los datos familiares son «débiles». Dependen de afirmaciones subjetivas, habitualmente expresadas con palabras y no con números, realizadas por las personas que actúan como observadoras y como observadas.

Para evitar este tipo de datos débiles, los resultados del tratamiento suelen restringirse a la información analítica que pueda ser objetiva, dimensional y fiable, aunque también quede deshumanizada. Si nos dicen que el colesterol sérico es de 230 mg/100 ml, que la radiografía torácica muestra una cardiomegalia y que el electrocardiograma presenta ondas Q, no sabremos si el objeto de que se trata es un perro o una persona. Si nos dijeran que se ha recuperado la capacidad para trabajar, que el medicamento tiene un sabor agradable y es fácil de tomar, y que la familia se muestra satisfecha con los resultados, reconoceríamos un grupo de respuestas humanizadas.

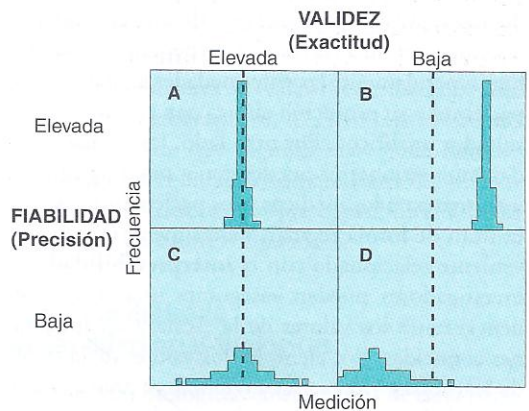
## Fiabilidad

La **fiabilidad** es el grado en que mediciones repetidas de un fenómeno estable, realizadas por personas e instrumentos diferentes, en momentos y lugá-

res distintos, expresan resultados similares. **Reproducibilidad** y **precisión** son otros términos que designan la misma propiedad.

La fiabilidad de las determinaciones analíticas se establece mediante mediciones repetidas (p. ej., del mismo suero o muestra tisular) llevadas a cabo, a veces, personas diferentes con instrumentos distintos. La fiabilidad de los síntomas puede comprobarse demostrando que diferentes observadores en condiciones diferentes los describen de forma similar.

En la figura 2-1, se muestran de forma simplificada las relaciones entre la fiabilidad y la validez. Las mediciones pueden ser válidas (exactas) de promedio, pero no fiables, puesto que se encuentran muy dispersas en torno al valor real, tal como muestra la gráfica C. Por otro lado, las mediciones pueden ser muy fiables, pero inexactas, si se alejan sistemáticamente del valor real, como sucede en la gráfica B. Un número reducido de mediciones con una fiabilidad deficiente también tendrán poca validez, ya que es muy posible que se alejen del valor real sólo por efecto del azar. Por lo tanto, la fiabilidad y la validez no son conceptos del todo independientes. La figura 2-1 constituye una aproximación; en general, una medición poco fiable no puede ser válida, y una medición válida debe ser fiable.



**FIGURA 2-1.** ■ Validez y fiabilidad. A. Validez elevada y fiabilidad elevada. B. Validez baja y fiabilidad elevada. C. Validez elevada y fiabilidad baja. D. Validez baja y fiabilidad baja. Las líneas de puntos son los valores reales.

## Amplitud

Puede que un instrumento no registre los valores muy elevados o muy bajos del fenómeno que se está midiendo, es decir, tiene una **amplitud** limitada, y la información que proporciona también lo es. Por ejemplo, la escala de Actividades Básicas de la Vida Diaria, que mide la capacidad de los pacientes para vestirse, comer, caminar, utilizar el retrete y mantener la higiene, no mide la capacidad de leer, escribir o tocar el piano (actividades que pueden ser muy importantes para determinados pacientes).

## Sensibilidad

Un instrumento demuestra su **sensibilidad** en la medida en que, al cambiar las condiciones, cambian los resultados. Por ejemplo, la escala de la NYHA (*New York Heart Association*), con clases de I a IV (sin síntomas, síntomas con esfuerzo ligero, con esfuerzo moderado y síntomas en reposo), no es sensible a los cambios sutiles que se producen en una insuficiencia cardíaca congestiva, precisamente aquellos que pueden ser importantes para los pacientes. Por otro lado, las mediciones de la fracción de eyección realizadas mediante ecocardiografía pueden detectar cambios tan sutiles que no los perciben ni los pacientes.

## Interpretabilidad

Los médicos aprenden a interpretar el significado de una PCO<sub>2</sub> de 50 mm Hg o de una glucemia de 460 mg/100 ml a través de la experiencia, tras calibrar repetidamente las enfermedades actuales de los pacientes y su evolución clínica con este tipo de resultados analíticos. Por otro lado, las escalas basadas en cuestionarios pueden tener poco significado intuitivo para los médicos y los pacientes que no las utilizan de forma regular. Para superar este inconveniente relacionado con la **interpretabilidad**, los investigadores pueden «asegurar» una correspondencia entre los valores de la escala y las situaciones conocidas. Por ejemplo, los valores de la escala de Karnofsky (*Karnofsky Performance Status Scale*), utilizada habitualmente en los estudios para medir la capacidad funcional de los pacientes con cáncer tratados con quimioterapia, oscilan entre 100 (normal) y 0 (muerte). Así, ¿cuál es la gravedad al obtener un valor de 60? Para que a los médicos les resul-

te más fácil interpretar los valores de la escala, se establece una correspondencia entre los números y las descripciones de las funciones cotidianas. Con un valor de 60 en la escala, los pacientes precisan ayuda ocasional, pero son capaces de solventar por sí mismos la mayor parte de sus necesidades.

## VARIACIÓN

Las mediciones clínicas del mismo fenómeno pueden adoptar todo un intervalo de valores en función de las circunstancias en las que dichas mediciones se realicen. A fin de evitar conclusiones erróneas a partir de los datos, los médicos deben ser conscientes de las razones que producen variaciones en una situación concreta, y saber cuáles pueden ejercer una gran influencia, poca influencia o ninguna influencia en aquello que se ha observado. Esto no sólo es válido para la interpretación de las investigaciones publicadas, sino también para la credibilidad de las mediciones clínicas del día a día.

La variación global es la suma de la variación relacionada con el propio acto de la medición, de las diferencias biológicas *en el interior* de las personas de un momento a otro y de las diferencias biológicas entre las personas (v. tabla 2-3).

## Variación derivada de la medición

Todas las observaciones están sujetas a variaciones, debido al funcionamiento de los instrumentos y a

TABLA 2-3. ■ Fuentes de variación

Fuente de variación	Definición
<i>Variación como consecuencia de la medida</i>	
Instrumento	Medios utilizados para realizar la medición
Observador	Persona que lleva a cabo la medición
<i>Variación biológica</i>	
Dentro de cada persona	Cambios que se producen en una persona con el tiempo y la situación
Entre personas	Diferencias biológicas entre una persona y otra

los observadores que intervienen en las mediciones. Las condiciones en que se realiza la medición pueden dar lugar a un resultado sesgado (falta de validez) o, sencillamente, a un error aleatorio (falta de fiabilidad). Esta fuente de variación puede reducirse si las mediciones se realizan con sumo cuidado y si se siguen protocolos normalizados. Sin embargo, cuando las mediciones dependen del juicio humano, y no de las máquinas, la variación que se genera puede ser especialmente amplia y difícil de controlar.

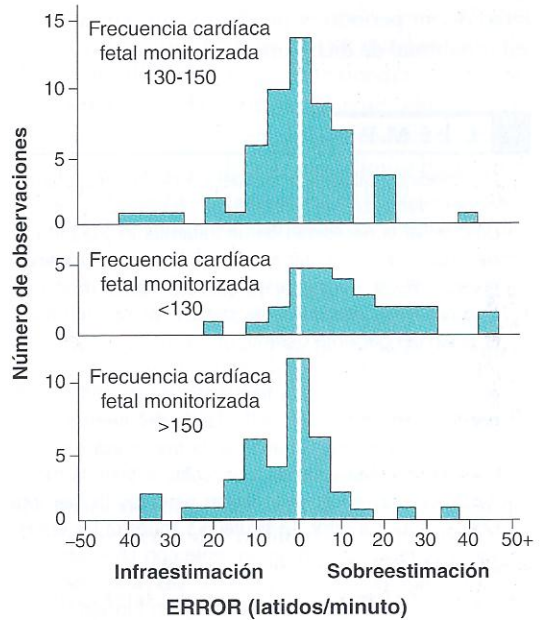
## EJEMPLO

La frecuencia cardíaca fetal suele determinarse por auscultación, un acto sujeto al error del observador. La monitorización electrónica proporciona la frecuencia real. Las frecuencias cardíacas fetales anormalmente elevadas o bajas son signos de sufrimiento fetal, y sugieren la necesidad de un parto prematuro mediante cesárea.

Day y cols. (2) compararon las frecuencias cardíacas fetales obtenidas por el equipo del hospital mediante auscultación con las frecuencias obtenidas mediante monitorización electrónica (fig. 2-2). Cuando la frecuencia cardíaca fetal real se encontraba dentro del intervalo normal, las frecuencias obtenidas por auscultación se distribuían uniformemente en torno al valor real (es decir, sólo se producía un error aleatorio). Pero cuando la frecuencia cardíaca fetal real era anormalmente elevada o baja, las frecuencias obtenidas por auscultación quedaron desviadas hacia el intervalo normal. Las frecuencias bajas tendieron a notificarse como más elevadas que las frecuencias reales, y las frecuencias elevadas, como más bajas que las frecuencias reales.

Este estudio ilustra los errores aleatorios y sistemáticos de las observaciones clínicas. En este caso, el sesgo hacia las frecuencias situadas dentro del intervalo normal podría deberse a que el equipo del hospital deseaba que el feto se encontrara en buen estado y se mostraba reacio a practicar una intervención mayor, basándose sólo en las observaciones de una frecuencia cardíaca anormalmente elevada o baja.

También pueden producirse variaciones en las mediciones porque se realizan en una sola muestra del fenómeno que se describe, lo que puede conllevar una representación errónea del conjunto. A menudo, la **fracción de muestreo** (fracción del conjunto incluida en la muestra) es muy pequeña. Por ejemplo, una biopsia hepática sólo representa apro-



**FIGURA 2-2.** ■ Variabilidad del observador. Errores en la notificación de la frecuencia cardíaca fetal en función de si la frecuencia real, determinada por un monitor electrónico, se sitúa dentro del intervalo normal, es baja o es elevada. (Reelaborado de Day E, Maddern L, Wood C. Auscultation of foetal heart rate: An assessment of its error and significance. Br Med J 1968;4:422-424).

ximadamente 1/100.000 del hígado. Puesto que se examina una parte tan minúscula del conjunto, es posible que se produzca una variación considerable entre una muestra y otra.

Si las mediciones se realizan con métodos diferentes, como laboratorios, técnicos o instrumentos distintos, puede que algunas sean poco fiables y proporcionen resultados sistemáticamente diferentes del valor correcto, lo que contribuiría a la dispersión de los valores obtenidos.

## Variación derivada de las diferencias biológicas

Los cambios biológicos en el interior de cada persona a lo largo del tiempo también pueden dar lugar a variaciones. La mayor parte de los fenómenos biológicos cambian de un momento a otro. Una medición en un punto del tiempo constituye una muestra de un conjunto de mediciones realizadas a lo

largo de un período, y puede que no represente el valor habitual de dichas mediciones.

## EJEMPLO

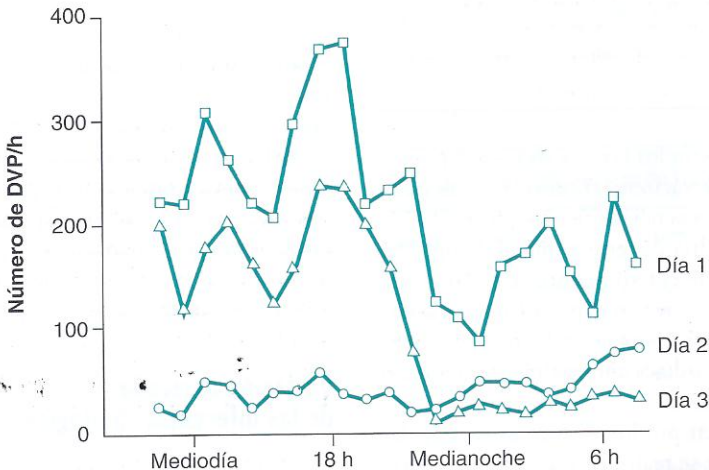
Los médicos calculan la frecuencia de las despolarizaciones ventriculares prematuras (DVP) para poder determinar la necesidad de un tratamiento y la eficacia de éste con más facilidad. Por razones prácticas, pueden llevar a cabo este procedimiento mediante observaciones relativamente breves, ya sea tomando el pulso del paciente durante un minuto o revisando el registro de un electrocardiograma durante algunos segundos. Sin embargo, la frecuencia de las DVP en un paciente concreto varía a lo largo del tiempo. Para obtener una muestra mayor de la frecuencia de DVP, a veces se utiliza un monitor portátil, si bien, la monitorización, incluso durante largos períodos de tiempo, puede inducir a error. La figura 2-3 muestra las observaciones realizadas en un paciente con DVP, similar a otros pacientes estudiados (3). Las DVP oscilaron entre menos de 20 y 380 por hora, en función del día y el momento del mismo, a lo largo de un período de tres días. Los autores concluyeron que «para distinguir una disminución de la frecuencia de las DVP atribuible sólo a la intervención terapéutica, y no a la variación biológica o espontánea, se requería una reducción superior al 83% en la frecuencia de DVP, si sólo se comparaban dos períodos de monitorización de 24

horas». Unos períodos de observación más cortos podrían ser más propensos al error debido a la variación biológica.

La variación también surge por las diferencias que existen entre las personas; en muchas situaciones, predominan las diferencias biológicas entre ellas. Por ejemplo, varios estudios han demostrado que la hipertensión arterial detectada en mediciones aisladas y fortuitas, a pesar de que éstas están sujetas a todas las demás formas de variación, está relacionada con una enfermedad cardiovascular posterior.

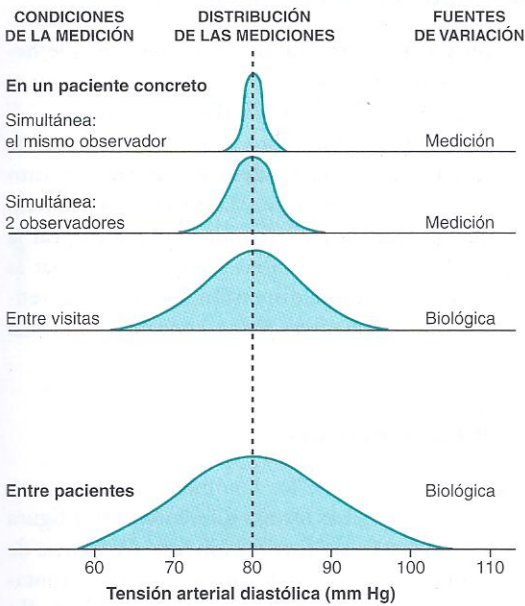
## Variación total

Las diferentes fuentes de variación son acumulativas. La figura 2-4 ilustra este hecho en la medición de la tensión arterial. Cuando se observa la distribución de una población, la variación de la medición en pacientes concretos se suma a la variación que se produce en dichos pacientes entre un momento y otro, que a su vez se suma a la variación entre pacientes diferentes. La variación de la medición contribuye relativamente poco a la variación total, aunque equivale a un intervalo de 12 mm Hg



**FIGURA 2-3.** ■ Variabilidad biológica. Número de DVP por hora en un paciente sin tratamiento previo durante tres días consecutivos. (Reelaborado de Morganroth J, Michelson EL, Horowitz LN, Josephson ME, Pearlman AS, Dunkman WB. Limitations of routine long-term electrocardiographic monitoring to assess ventricular ectopic frequency. *Circulation* 1978;58:408-414.)





**FIGURA 2-4.** ■ Fuentes de variación en la medición de la tensión arterial diastólica (fase V). La línea de puntos indica la tensión arterial real. (De: Fletcher RH and Fletcher SW (unpublisend); and Boe J, Humerfelt S, Wedervang F, Oecon C. The blood pressure in a population [especial Issue]. Acta Med Scand 1957;321:5-313).

entre diferentes observadores. Por otro lado, la tensión arterial de cada paciente varía en gran medida de un momento a otro del día, por lo que cualquier lectura de la misma puede no representar el valor habitual de ese paciente. Una gran parte de esta variación no es aleatoria: en general, la tensión arterial es más elevada cuando las personas están despiertas o excitadas, cuando acuden al médico o cuando toman fármacos de libre dispensación para el resfriado. Obviamente, lo que más nos interesa es saber cómo se compara la tensión arterial de un paciente con la del resto de los mismos, especialmente si el valor obtenido está relacionado con complicaciones de la hipertensión arterial y con la eficacia del tratamiento.

### Efectos de la variación

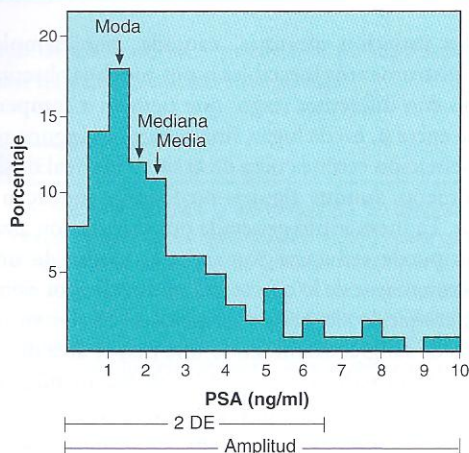
Otro modo de concebir la variación es desde la perspectiva de su efecto neto sobre la validez y la fiabilidad de una medición, y de lo que se pueda hacer al respecto.

La variación aleatoria, causada, por ejemplo, por instrumentos inestables o por muchos observadores con diferentes sesgos que tienden a compensarse entre sí, no da lugar, finalmente, a ninguna representación errónea neta de la situación real de un fenómeno, aunque algunas mediciones induzcan a error. La inexactitud generada por la variación aleatoria puede reducirse tomando la media de una muestra mayor de lo que se está midiendo, por ejemplo, contando un mayor número de células en un frotis de sangre, examinando una mayor área de sedimento urinario o estudiando un mayor número de pacientes. Asimismo, el grado de variación aleatoria puede calcularse mediante métodos estadísticos (v. cap. 10).

Por otro lado, los resultados sesgados son sistemáticamente diferentes del valor real, independientemente del número de veces que se repitan. Por ejemplo, la medición de la frecuencia cardíaca fetal (ejemplo descrito en la fig. 2-2) seguiría alejada del valor real si se repitiera muchas más veces. De forma similar, todos los valores elevados para las DVP descritas en la figura 2-3 se registraron el primer día, y la mayor parte de los valores bajos, en el tercero. Los días fueron estimaciones sesgadas unos respecto a otros, debido a la variación en la frecuencia de las DVP entre un día y otro.

## DISTRIBUCIONES

Los datos que se miden mediante escalas de intervalo se presentan, a menudo, como un diagrama, denominado **distribución de frecuencia**, en el que se muestra el número (o la proporción) de un grupo definido de personas que poseen los diferentes valores de aquello que se mide. La figura 2-5 muestra la distribución en varones sanos del antígeno prostático específico (PSA, *prostate-specific antigen*), cuya medición se emplea para detectar indicios de enfermedad prostática, especialmente cáncer. Aunque las concentraciones de PSA oscilan entre 0 y 10 ng/ml, la mayor parte de los resultados de las mediciones se sitúan en el intervalo entre 0 y 4 ng/ml. Mostrar los datos del intervalo como una distribución de frecuencia hace que la información quede presentada de forma relativamente detallada.



**FIGURE 2-5.** ■ Medidas de tendencia central y dispersión. Distribución de los niveles de PSA en varones supuestamente sanos. (De: Kane RA, Littrup PJ, Babaian R, Drago JR, Lee F, Chesley A, Murphy GP, Mettlin C. Prostate-specific antigen levels in 1695 men without evidence of prostate cancer. *Cancer* 1992; 69:1201-1207).

## Descripción de las distribuciones

Es conveniente resumir las distribuciones. De hecho, esta acción es obligada cuando se presenta y se compara un gran número de distribuciones.

Para ello, se utilizan dos propiedades básicas de las distribuciones: la **tendencia central**, o punto central de la distribución, y la **dispersión**, o grado en que se encuentran repartidos los valores. En la figura 2-5, se ilustran varias formas de expresar la tendencia central y la dispersión, así como sus ventajas e inconvenientes. Estos conceptos también aparecen resumidos en la tabla 2-4.

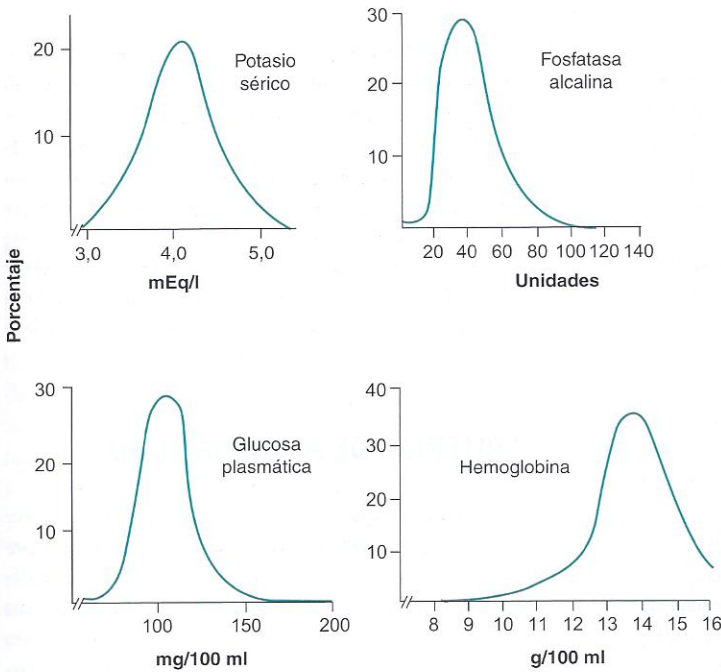
## Distribuciones reales

Las distribuciones de los fenómenos clínicos pueden adoptar muchas formas diferentes. En la figura 2-6, se muestran las distribuciones de frecuencia de cuatro tipos de análisis de sangre habituales (potasio, fosfatasa alcalina, glucosa y hemoglobina). En

**TABLA 2-4.** ■ Expresiones de tendencia central y dispersión

Expresión	Definición	Ventajas	Inconvenientes
<b>Tendencia central</b>			
Media	$\frac{\text{Suma de los valores de las observaciones}}{\text{Número de observaciones}}$	Idónea para la manipulación matemática	Se ve afectada por los valores extremos
Mediana	Punto de la distribución en el que el número de observaciones situado por encima es el mismo que el número de observaciones situado por debajo	Los valores extremos no la afectan fácilmente	Poco adecuada para la manipulación matemática
Moda	El valor obtenido con más frecuencia	Sencillez de significado	En ocasiones no existen valores más frecuentes o bien hay muchos
<b>Dispersión</b>			
Amplitud	Valores situados entre el valor más bajo y el más alto de una distribución	Incluye todos los valores	Se ve enormemente afectada por los valores extremos
Desviación estándar <sup>a</sup>	Valor absoluto de la diferencia media entre cada valor con respecto a la media <sup>a</sup>	Idónea para la manipulación matemática	En las distribuciones no gaussianas (v. a continuación) no describe una proporción conocida de las observaciones
Percentil, decil, cuartil, etc.	Proporción de todas las observaciones situadas entre valores específicos	Describe la «nfrecuencia» de un valor No realiza suposiciones sobre la forma de la distribución	Poco adecuada para la manipulación estadística

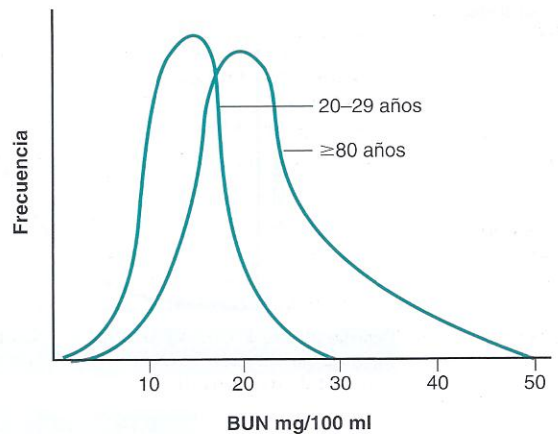
<sup>a</sup>  $\sqrt{\frac{\sum(X - \bar{X})^2}{N - 1}}$  donde  $X$  = cada observación;  $\bar{X}$  = media de todas las observaciones; y  $N$  = número de observaciones.



**FIGURA 2-6.** ■ Distribuciones clínicas reales. (De: Martin HF, Gudzinowicz BJ, Fanger H. *Normal values in clinical chemistry*. New York: Marcel Dekker, 1975.)

general, la mayor parte de los valores aparecen próximos al centro de la distribución y, con excepción de la parte central de las curvas, no se observan protuberancias ni irregularidades. Los extremos elevados y bajos de cada distribución se extienden formando colas, una de las cuales suele ser más alargada que la otra (es decir, las curvas están desviadas hacia el extremo más alargado). Algunas distribuciones están desviadas hacia los valores más elevados, mientras que otras lo están hacia los valores más bajos. En otras palabras, todas estas distribuciones son unimodales (poseen una sola protuberancia) y presentan, en cierta manera, forma de campana, aunque no necesariamente simétrica. Por lo demás, no se parecen entre sí.

En muchas pruebas analíticas, la distribución de los valores cambia en función de las características de los pacientes, como la edad, el sexo, la raza y la alimentación. La figura 2-7 muestra cómo la distribución de una de estas pruebas, el nitrógeno ureico en sangre (BUN, prueba para medir la función renal), cambia con la edad. Una concentración de BUN de 25 mg/100 ml sería un valor anormalmente elevado para una persona joven, pero no representaría un valor destacable en una persona anciana.



**FIGURA 2-7.** ■ La distribución de las variables clínicas cambia con la edad: BUN para personas entre 20 y 29 años de edad frente a BUN para personas con una edad de 80 años o superior. (De: Martin HF, Gudzinowicz BJ, Fanger H. *Normal values in clinical chemistry*. New York: Marcel Dekker, 1975.)

## Distribución normal

Otro tipo de distribución es la denominada **distribución normal** (o «gaussiana», bautizada así en honor de Gauss, el matemático que la describió por primera vez). La distribución normal, basada en la teoría estadística, describe la distribución de frecuencia de repetidas mediciones del mismo objeto físico realizadas por el mismo instrumento. La dispersión de los valores se debe únicamente a la variación aleatoria. En la figura 2-8, se muestra una curva normal, simétrica y con forma de campana. Esta curva tiene la propiedad matemática de que aproximadamente  $2/3$  de las observaciones se sitúan dentro de una desviación estándar de la media, y alrededor del 95%, dentro de dos desviaciones estándares.

Aunque a menudo las distribuciones clínicas se parecen a las distribuciones normales, el parecido es superficial. Tal como afirmó un estadístico (4): «El hecho experimental es que para la mayor parte de las variables fisiológicas, la distribución es uniforme, unimodal y desviada (asimétrica), y que la media  $\pm 2$  desviaciones estándares no se puede utilizar como punto de corte para el 95% deseado. No disponemos de teoremas matemáticos, estadísticos o de otro tipo, que nos permitan predecir la forma que adoptarán las distribuciones de las mediciones fisiológicas.»

Las formas de las distribuciones clínicas difieren entre sí porque a las distribuciones de las mediciones clínicas contribuyen, además de la variación aleatoria, las numerosas diferencias existentes entre las personas. Por lo tanto, si las distribuciones de las mediciones clínicas adoptan curvas normales es, en gran parte, por pura coincidencia. Aun así, a menudo se supone, por comodidad (puesto que las medias y las desviaciones estándares son relativamente fáciles de calcular y de manipular matemáticamente), que las mediciones clínicas se distribuyen de forma «normal».

## CRITERIOS DE ANORMALIDAD

Resultaría práctico que las distribuciones de frecuencia de las mediciones clínicas realizadas en pacientes sanos y en pacientes enfermos fueran tan diferentes que pudieran utilizarse para distinguir dos poblaciones diferentes. En realidad, esto es así para algunos genes anómalos. Las secuencias (que codifican anomalías genéticas) para la enfermedad autosómica dominante poliposis adenomatosa familiar pueden estar presentes o ausentes. Las personas que poseen el gen anómalo presentarán cientos de pólipos en el colon, mientras que las personas sin dicho gen rara vez tienen alguno. Sin embargo, ésta es la excepción que confirma la regla. Lo más

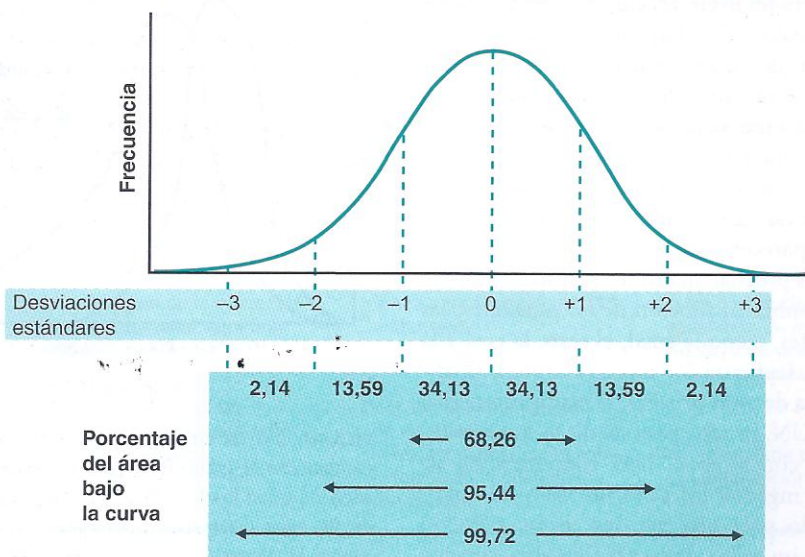


FIGURA 2-8. ■ Distribución normal (gaussiana).

habitual es que las diferentes anomalías genéticas que codifican la misma enfermedad produzcan toda una gama de expresiones. Incluso la expresión de una anomalía genética específica (como la sustitución de un solo par de bases) difiere sustancialmente de una persona a otra, probablemente debido a diferencias en el resto de la dotación genética y a la exposición a las causas externas de la enfermedad.

Por lo tanto, la mayor parte de las distribuciones de variables clínicas no se dividen tan fácilmente en «normales» y «anómalas». No son inherentemente dicotómicas, ni muestran brechas pronunciadas o dos picos que caractericen los resultados normales y anómalos. Esto es así porque habitualmente la enfermedad se adquiere por grados, de forma que se produce una suave transición entre los valores bajos y los valores elevados, con grados de disfunción cada vez mayores. Los resultados de las pruebas analíticas que reflejan insuficiencias orgánicas, como la creatinina sérica para la insuficiencia renal o la fracción de eyección para la insuficiencia cardíaca, se comportan de este modo.

Otra razón por la que pacientes sanos y enfermos no se aprecian como distribuciones separadas es que, incluso cuando las personas con una enfermedad y las que no la sufren presentan distribuciones de frecuencia sustancialmente diferentes, dichas distribuciones casi siempre se superponen. Cuando ambas distribuciones se mezclan, como sucede en las poblaciones que existen de modo natural, las anomalías no suelen percibirse de forma separada, debido a que ocupan una proporción muy pequeña del conjunto. La curva de las personas enfermas es «engullida» por la curva de mayor tamaño de las personas sanas.

## EJEMPLO

La fenilcetonuria (PKU, *phenylketonuria*) es una enfermedad congénita caracterizada por un retraso mental progresivo durante la infancia. Diversos alelos mutantes que codifican la fenilalanina hidroxilasa provocan la disfunción de la enzima y, con una alimentación normal, se produce una acumulación de fenilalanina que termina causando síntomas. En principio, los genes pueden ser anómalos o no (fig. 2-9 A). El diagnóstico, que se manifiesta durante el primer año de vida, lo confirman la persistencia de elevados niveles de fenilalanina (varias veces los valores habituales) y los bajos niveles de tirosina en la sangre.

Es una práctica habitual realizar pruebas de detección de PKU a los neonatos, mediante análisis de fenilalanina en sangre a los pocos días de nacer, a tiempo para recibir tratamiento antes de que los daños sean irreversibles. Sin embargo, aunque los genes anómalos estén presentes o ausentes (fig. 2-9 A), la prueba clasifica de forma errónea a algunos lactantes debido a que, a esa edad, las distribuciones de las concentraciones séricas de fenilalanina en los niños con y sin PKU se superponen, y a que los lactantes con PKU constituyen sólo una pequeña proporción de todos los que se someten a la prueba de detección, es decir, menos de 1/10.000 (fig. 2-9 B). Algunos recién nacidos con PKU se sitúan dentro del intervalo normal, ya sea porque aún no han ingerido suficientes proteínas que contengan fenilalanina, o bien porque presentan una combinación de alelos que se asocia a una manifestación leve de la enfermedad. Algunos niños que no están destinados a presentar PKU tienen unos niveles de fenilalanina relativamente altos, por ejemplo, porque sus madres sufren anomalías en el metabolismo de este aminoácido. El resultado de la prueba se ha establecido como positivo en el extremo inferior de la superposición, entre los niveles normales y anómalos, a fin de detectar a la mayoría de los lactantes con la enfermedad, a pesar de que sólo 1 de cada 5 lactantes con resultados anómalos en la prueba de cribado estará realmente afectado por la PKU.

Si no existe una clara línea divisoria entre lo normal y lo anómalo, y el médico puede elegir dónde se sitúa dicha línea, ¿qué normas básicas deben utilizarse para decidir? Existen tres criterios que han demostrado ser útiles: que se trate de un hecho infrecuente, que se asocie a un estado de enfermedad y que sea tratable. Para una medición concreta, los resultados de estos métodos no tienen por qué estar relacionados entre sí; lo que un criterio considera normal, puede que otro lo considere anómalo.

## Anormal = Infrecuente

El término normal a menudo se refiere a una situación que se produce con frecuencia o que es habitual. Lo que ocurre a menudo se considera normal y lo que se produce de forma infrecuente se considera anormal. Ésta es una definición estadística, basada en la frecuencia de una característica en una población definida. Habitualmente, la población de referencia se compone de personas sanas, aunque no tiene por qué ser así. Por ejemplo, podemos decir que es normal presentar dolor después de una

intervención quirúrgica, o prurito si se tiene un ecema.

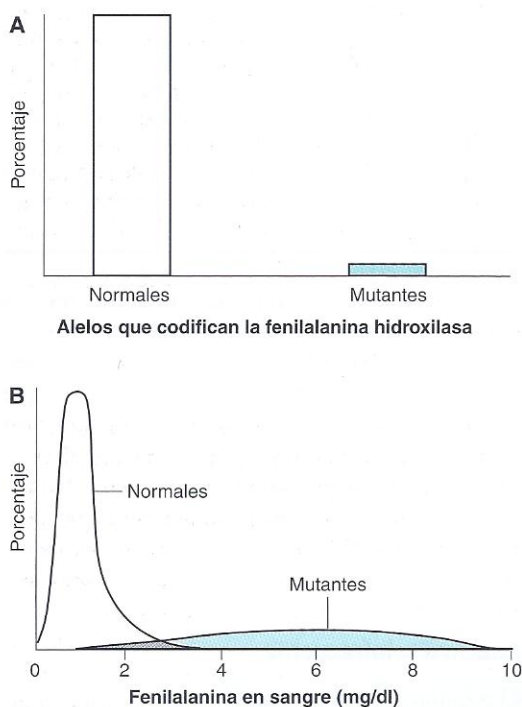
Resulta tentador tratar de lograr un mayor grado de especificidad definiendo lo que es anómalo mediante términos matemáticos. Una forma habitual de establecer un punto de corte entre lo normal y lo anómalo es acordar, de forma algo arbitraria, que todos los valores situados más allá de dos desviaciones estándares de la media son anómalos. Suponiendo que la distribución en cuestión se aproxima a una distribución normal, el 2,5% de las observaciones se situarían en cada cola de la distribución, y se considerarían anormalmente elevadas o anormalmente bajas.

Es obvio, como ya se ha señalado, que la mayor parte de las mediciones biológicas no producen dis-

tribuciones normales, por lo que es mejor describir los valores anómalos, independientemente de la proporción elegida, como una fracción (o percentil) de la distribución real. De esta forma, se puede realizar una afirmación directa sobre el grado de infrecuencia de un valor sin hacer suposiciones sobre la forma de la distribución de la que procede.

Habitualmente se utiliza una definición estadística de normalidad, que puede ser ambigua o conducir a error por diversos motivos:

1. Si todos los valores situados más allá de un límite estadístico arbitrario, supongamos el percentil 95, se consideraran anómalos, la frecuencia de todas las enfermedades sería la misma (si consideráramos que la distribución es normal, estaríamos hablando del 2,5% de los valores, teniendo en cuenta sólo los extremos elevado y bajo de la distribución). Y todo el mundo sabe que la frecuencia de las enfermedades es variable; la diabetes y la artrosis son mucho más frecuentes que la ovalocitosis y la tricoleucemia.
2. No existe una relación general entre el grado de infrecuencia estadística y la enfermedad clínica. La relación es específica de la enfermedad en cuestión y del entorno. Por ejemplo, la obesidad es bastante habitual en EE.UU., pero infrecuente en muchos de los países subdesarrollados. Para algunas mediciones, las desviaciones de los valores habituales sólo se asocian a un grado importante de enfermedad en los valores extremos, más allá del percentil 95, ó incluso el percentil 99. La insuficiencia de órganos como el hígado y los riñones sólo produce síntomas cuando se ha perdido la mayor parte de su función.
3. En ocasiones, los valores extremos pueden resultar, en realidad, beneficiosos. Por ejemplo, las personas con una tensión arterial anormalmente baja presentan, en promedio, un riesgo menor de sufrir enfermedades cardiovasculares que las personas con tensiones arteriales más elevadas y más habituales. Las personas con una densidad ósea anormalmente elevada poseen, de media, un riesgo menor de sufrir fracturas.

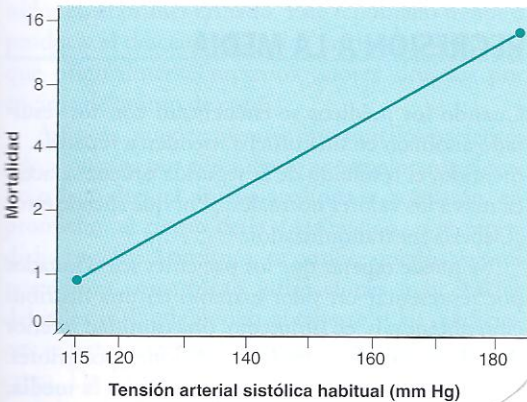


**FIGURA 2-9.** Detección de fenilcetonuria (PKU) en lactantes: distribuciones dicotómicas y superpuestas de lo normal y lo anómalo. A. Los alelos que codifican la fenilalanina hidroxilasa son normales o mutantes. B. Las distribuciones de los niveles de fenilalanina en sangre en los neonatos con y sin PKU se superponen, y la magnitud de ambas es considerablemente diferente. (La prevalencia de la PKU, que es inferior a 1/10.000, se ha exagerado, a fin de que pueda observarse su distribución en la figura.)

4. Muchas mediciones están relacionadas con el riesgo de enfermedad en un amplio intervalo de valores, sin que exista un umbral que separe el riesgo normal de un riesgo mayor. La tensión arterial es un ejemplo de ello.

### EJEMPLO

La figura 2-10 muestra que una tensión arterial sistólica más elevada de lo habitual está relacionada con mayores tasas de mortalidad por cardiopatía isquémica en un amplio intervalo de valores, que oscilan entre los 115 y los 185 mm Hg. Por cada aumento de 20 mm Hg en la tensión arterial sistólica, se dobla el índice de fallecimientos por cardiopatía isquémica, sin indicios de que exista un umbral por debajo de 115 mm Hg, el menor valor para el que existen datos suficientes (5). Aproximadamente, el 80% de la población general de adultos tiene una tensión arterial inferior a 140 mm Hg, punto en el que se recomienda el tratamiento farmacológico.



**FIGURA 2-10.** ■ La mortalidad por cardiopatía isquémica de las personas de entre 40 y 49 años de edad está relacionada con los valores de tensión arterial sistólica situados a lo largo del intervalo en el que se emplazan la mayoría de las personas. No existe un umbral entre lo normal y lo anómalo, ni para las personas de mediana edad ni para los ancianos. La «mortalidad» se presenta en forma de progresión geométrica respecto a la situación basal. (De: Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: A meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002;360:1903-13.)

### Anormal = Asociado a enfermedad

Un método más sólido para distinguir lo normal de lo anormal consiste en considerar anómalas todas aquellas observaciones que se apartan de forma clínicamente significativa de un buen estado de salud, es decir, que están asociadas al riesgo de sufrir o desarrollar una enfermedad, una discapacidad o la muerte.

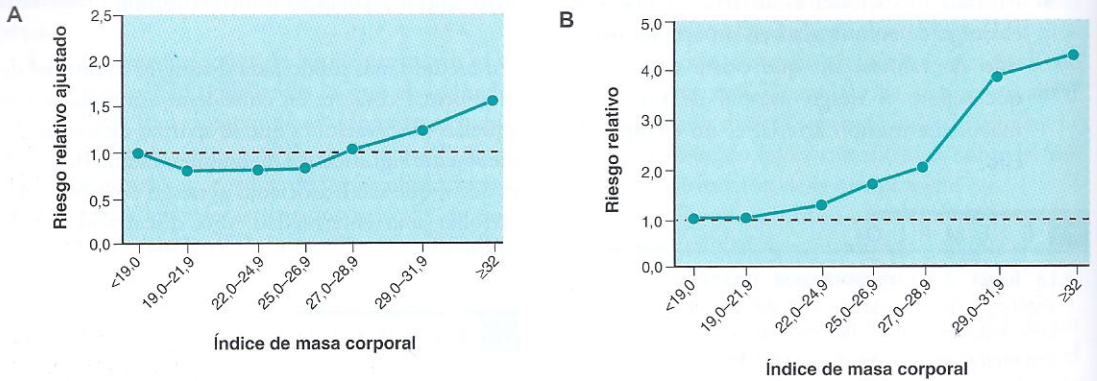
### EJEMPLO

¿A partir de qué punto un peso superior a la media con respecto a la altura se convierte en un problema de salud? La figura 2-11 A muestra la relación entre el índice de masa corporal (IMC, que se obtiene dividiendo el peso de una persona, en kilogramos, por su talla, en metros, elevada al cuadrado) y el riesgo de muerte por todas las causas en mujeres que nunca han fumado (6). El riesgo es casi el mismo a lo largo de un amplio intervalo valores de IMC, y sólo empieza a aumentar cuando se alcanza un IMC de 29. Sin embargo, el riesgo de enfermedad cardiovascular empieza a aumentar a partir de valores de IMC inferiores (fig. 2-11 B), lo que refleja que el punto de corte entre lo normal y lo anómalo también depende de la enfermedad en cuestión.

### Anormal = Tratable

En algunos procesos, especialmente en aquellos que por sí mismos no producen un cuadro clínico (asintomáticos), es mejor considerar que una medición es anómala sólo si el tratamiento del proceso que se ha conocido gracias a la medición conlleva un mejor resultado. La razón es que no todos los procesos asociados a un mayor riesgo pueden tratarse de forma eficaz: puede que la eliminación del proceso no elimine el riesgo, ya sea porque el propio proceso no es la causa de la enfermedad, sino que sólo está relacionado con la causa, o porque ya se ha producido un daño irreversible. Asimismo, etiquetar a una persona de «enfermo» puede provocar preocupación y un sentimiento de vulnerabilidad que no están justificados si el tratamiento no puede mejorar el pronóstico.

Lo que se considera tratable cambia con el tiempo. En el mejor de los casos, las decisiones terapéuticas se basan en pruebas procedentes de en-



**FIGURA 2-11.** ■ Asociación de anomalía y enfermedad. Relación entre el IMC y A, la mortalidad total y B, la incidencia de enfermedades cardiovasculares en las mujeres. El IMC se calcula dividiendo el peso en kilogramos por la talla en metros elevada al cuadrado. El «riesgo relativo» es la tasa de mortalidad relativa a las personas situadas en la categoría más baja del IMC (<19,0). (De: Manson JE, Willett WC, Stampfer MJ, Colditz GA, Hunter DJ, Hankinson SE, Hennekens CH, Speizer FE. Body weight and mortality among women. *N Engl J Med* 1995;333:677-685.)

sayos clínicos bien diseñados (cap. 8). A medida que se adquieren nuevos conocimientos derivados de los resultados de los ensayos clínicos, cambia el nivel a partir del cual el tratamiento se considera beneficioso.

## EJEMPLO

El ácido fólico, una vitamina que se encuentra principalmente en las verduras, fue descubierto a principios del siglo pasado. Se observó que las dietas con un bajo contenido de ácido fólico provocaban un síndrome de hipovitaminosis caracterizado por la aparición de anemia. Durante muchos años, el consumo suficiente de ácido fólico se definió en los estándares RDA (*recommended dietary allowance*) como la cantidad necesaria para evitar la anemia. Sin embargo, recientemente, han surgido argumentos a favor de unos nuevos niveles de consumo «normal» más elevados. El consumo de ácido fólico al comienzo del embarazo en unos niveles inferiores a los de muchas mujeres, pero muy por encima del nivel necesario para evitar la anemia, se ha vinculado con malformaciones del tubo neural en la descendencia de las gestantes. Los aportes de ácido fólico en las mujeres de alto riesgo han evitado aproximadamente tres cuartas partes de dichas malformaciones. Por lo tanto, ha surgido un nuevo nivel de consumo «normal» de ácido fólico debido a la presencia de información nueva relacionada con los niveles necesarios para evitar la enfermedad. El nivel óptimo es, al menos, dos veces (y algunos sugieren que hasta ocho veces) superior al antiguo criterio de aporte normal. Por una razón prác-

tica, el consumo de tales niveles sólo resulta factible, para la mayoría de las personas, mediante aportes vitamínicos, y no únicamente con la alimentación.

## REGRESIÓN A LA MEDIA

Cuando los médicos se encuentran con un resultado anómalo en una prueba, tienden a repetirla. A menudo, el resultado de la segunda prueba se acerca más a los valores normales. ¿Por qué sucede esto? ¿Debería ser tranquilizador?

Se puede esperar que los pacientes seleccionados por representar un valor extremo en una distribución obtendrán, en promedio, una cantidad inferior de valores extremos en las mediciones posteriores. Este fenómeno, denominado **regresión a la media**, se produce por puras razones estadísticas, y no porque los pacientes hayan mejorado necesariamente.

El origen de la regresión a la media es el siguiente (v. fig. 2-12): las personas son seleccionadas para su inclusión en un estudio, o para un posterior diagnóstico o tratamiento, porque la medición inicial de una de sus características se sitúa más allá de un punto de corte seleccionado arbitrariamente en la cola de una distribución de valores que pertenece a un conjunto de pacientes examinados. En mediciones posteriores, algunas de estas personas seguirán por encima del punto de corte, debido a que



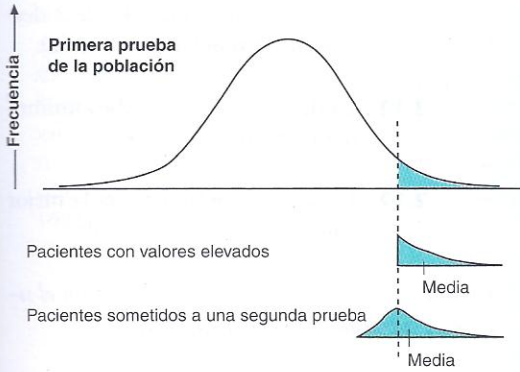


FIGURA 2-12. ■ Regresión a la media.

sus valores reales suelen ser más elevados que la media. Pero otras, cuyos valores se situaron por encima del punto de corte durante el cribado inicial, suelen presentar valores inferiores. Estas personas fueron seleccionadas únicamente porque dio la casualidad, originada por la variación aleatoria, de que presentaron un valor elevado en el momento de la primera medición. Al repetir la medición, estas personas presentan valores inferiores a los obtenidos en el primer cribado. Este fenómeno tiende a producir el descenso del valor medio del subgrupo que originalmente presentó valores situados por encima del punto de corte.

Así pues, puede esperarse que aquellos pacientes seleccionados por el resultado anormalmente elevado o bajo de una prueba analítica se acercarán, en promedio, al centro de la distribución, si se repite dicha prueba. Asimismo, es posible que los valores posteriores constituyan estimaciones más exactas del valor real, que podría obtenerse si las mediciones se repitieran muchas veces en un paciente concreto. De este modo, la práctica tradicional de repetir las pruebas analíticas con resultados anómalos y de considerar que el segundo es el correcto no es tan sólo cuestión de pensamientos bienintencionados, pues cuenta con el respaldo de una base teórica sólida y también de una base empírica. Por ejemplo, se ha demostrado que la mitad de las pruebas de  $T_4$  sérica que se situaron fuera de los límites normales durante el cribado, se hallaban dentro de los límites normales tras su repetición (7). Sin embargo, cuanto más extremo es el valor inicial, menos probable es que se encuentre dentro de los valores normales cuando se repita la prueba.

## PREGUNTAS PARA LA REVISIÓN

Para cada uno de los supuestos clínicos que se citan a continuación (2.1-2.5), seleccione el término más adecuado para definir el tipo de datos a que hacen referencia. Cada opción puede utilizarse una vez, más de una vez o puede no utilizarse nunca..

- 2.1. Reflejo tendinoso profundo, grado 0 (sin respuesta), 1+ (ligeramente disminuido), 2+ (normal), 3+ (más activo que la media) y 4+ (muy activo).
  - A. De intervalo: continuos
  - B. Dicotómicos
  - C. Nominales
  - D. Ordinales
  - E. De intervalo: discretos
- 2.2. Cáncer recurrente/no recurrente cinco años después del tratamiento inicial.
- 2.3. Sodio sérico de 139 mg/dl.
- 2.4. Tres crisis convulsivas al mes.
- 2.5. Causas de hemorragia digestiva alta: úlcera duodenal, gastritis, varices esofágicas o de otro tipo.
- 2.6. Cuando no es posible comprobar la medida de un fenómeno, como el prurito, mediante los sentidos físicos, ¿cuál de las siguientes afirmaciones es aplicable a su validez?
  - A. Es dudosa y hay que confiar en las medidas «sólidas», como las pruebas analíticas.
  - B. Su validez puede comprobarse demostrando que, cuando observadores diferentes repiten la medición en momentos diferentes, obtienen el mismo valor.
  - C. Su validez puede respaldarse si se demuestra que la medición está relacionada con mediciones de otros fenómenos, como la presencia de enfermedades que se sabe que provocan prurito.

- D. Su validez puede comprobarse demostrando que la medición da lugar a un amplio intervalo de valores.
- E. Su validez no puede establecerse.

**2.7.** Un médico o una enfermera miden la frecuencia cardíaca de una paciente tomándole el pulso durante 10 segundos cada vez que ésta acude a la consulta. Las frecuencias obtenidas pueden diferir de una visita a otra debido a todas las razones citadas a continuación, excepto:

- A. La paciente presenta diferentes frecuencias de pulso en momentos diferentes.
- B. La medición puede representar de forma errónea el pulso verdadero por efecto del azar y debido al breve período de observación.
- C. El médico y la enfermera utilizan técnicas diferentes (p. ej., tipo de presión sobre el pulso)
- D. La frecuencia del pulso varía entre los pacientes.
- E. Se empezó un tratamiento eficaz entre una visita y otra.

**2.8.** Habitualmente, el término «anormal» se corresponde con todas las definiciones siguientes, excepto:

- A. Nivel a partir del cual el tratamiento ha demostrado ser eficaz.
- B. Nivel a partir del cual aumenta la tasa de mortalidad.
- C. Valores estadísticamente infrecuentes.
- D. Valores que no corresponden a una distribución normal.
- E. Nivel a partir del cual existe un mayor riesgo de presentar síntomas.

*Indique si las siguientes afirmaciones son verdaderas o falsas.*

- **2.9.** La distribución normal describe la distribución de los fenómenos que se producen de forma natural.
- **2.10.** La distribución normal incluye un 2,5% de personas en cada cola de

la distribución (más allá de 2 desviaciones estándar de la media).

- **2.11.** La distribución normal es unimodal y simétrica.
- **2.12.** La distribución normal es la mejor base para definir lo anómalo.

*Indique si las afirmaciones relacionadas con el siguiente supuesto son verdaderas o falsas.*

En una primera medición, un paciente presenta un nivel de colesterol sérico superior a 200 mg/dl, valor que se sitúa en el percentil 65 para adultos en EE.UU. y que sobrepasa el umbral ampliamente aceptado a partir del cual se considera necesario el tratamiento farmacológico.

- **2.13.** Probablemente, este paciente presentará un nivel de colesterol sérico inferior en la próxima medición.
- **2.14.** Este paciente puede hallarse por encima del umbral por efecto del azar.
- **2.15.** El paciente debe ser tratado conforme a la observación de que su nivel de colesterol es superior al de sus semejantes.

Encontrará las respuestas en el apéndice A.

## BIBLIOGRAFÍA

1. Feinstein AR. The need for humanized science in evaluating medication. *Lancet* 1972;2:421-423.
2. Day E, Maddern L, Wood C. Auscultation of foetal heart rate: an assessment of its error and significance. *Br Med J* 1968;4:422-424.
3. Morganroth J, Michelson EL, Horowitz LN, Josephson ME, et al. Limitations of routine long-term electrocardiographic monitoring to assess ventricular ectopic frequency. *Circulation* 1978;58:408-414.

4. Elveback LR, Guillier CL, Keating FR. Health, normality, and the ghost of Gauss. *JAMA* 1970;211:69-75.
5. Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: A meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002;360:1903-13.
6. Manson JE, Willett WC, Stampfer MJ, Colditz GA, Hunter DJ, Hankinson SE, Hennekens CH, Speizer FE. Body weight and mortality among women. *N Engl J Med* 1995;333:677-685.
7. Epstein KA, Schneiderman LJ, Bush JW, Zettner A. The "abnormal" screening serum thyroxine (T4): Analysis of physician response, outcome, cost, and health and effectiveness. *J Chron Dis* 1981;34:175-190.