El cálculo de probabilidad y el método de las mayorías*

Émile Borel

De forma general, denominaremos *método de las mayorías* al procedimiento que consiste en considerar como prácticamente válida la opinión expresada por la mayoría, tras el escrutinio de las opiniones o pareceres de un número más o menos grande de personas. No nos detendremos, por el momento, a enumerar o discutir las diversas formas que puede adoptar el método, ni a intentar clasificar sus numerosas aplicaciones, realizadas o posibles. La cuestión que nos ocupará principalmente es la siguiente: ¿el conocimiento del cálculo de probabilidades puede ser de alguna utilidad para aquéllos que emplean o critican el método de las mayorías?

I

Josepth Bertrand ha reunido en su *Calcul des Probabilités* (XLIII-XLIX y 319-327) los argumentos que objetan, por así decirlo, toda cuestión previa a la introducción del cálculo de probabilidades en estas materias. Algunos pasajes darán a entender el punto de vista en el que se sitúa.

> La aplicación del cálculo a las decisiones judiciales es, según Stuart Mill, el escándalo de las matemáticas. La acusación es injusta. Podemos pesar cobre y ofrecerlo por oro, y el balance permanece de forma irreprochable. En sus trabajos sobre la teoría de los juicios de valor, Condorcet, Laplace y Poisson sólo han pesado cobre.
>
> ... «Condorcet ha tomado posesión del universo moral para someterlo al cálculo». Es el elogio que se le ha dado; pero uno se pregunta si es después de haberle leído. En su libro sobre la *probabilidad de los juicios de valor* propone, para empezar, dos problemas. Primero: ¿cuál es, para cada juicio de valor y para cada juez, la probabilidad de ser exacto? En segundo lugar, ¿cuál es la probabilidad de error a la cual la sociedad puede resignarse sin alarmarse?
>
> La primera cuestión le parece fácil.
>
> «Supongamos, dice Condorcet, que se haya elegido un número de hombres verdaderamente ilustrados y que se pronuncian sobre la verdad o la fal-

---

* Traducción de Beatriz Mañas. Ref: Borel, E., «Le calcul des probabilités et la méthode des majorités» *Année Psychologique,* t. 14, 1908, pp. 125-159.

EMPIRIA. Revista de Metodología de Ciencias Sociales. N.º 13, enero-junio, 2007, pp. 171-192.
ISSN: 1139-5737

sedad de una decisión. Si, entre las decisiones de este tribunal examinador, sólo se consideran aquéllas que han obtenido una cierta pluralidad, es fácil ver que, sin error sensible, se las puede contemplar como seguras».

Es simplemente un concilio infalible que él define y que pretende convocar. Sin dudar, vacila; no es que los hombres verdaderamente ilustrados sean escasos, no vayamos a creer eso, pero su tiempo es preciado. Para ahorrarlo, Condorcet propone un segundo método por el que, más tarde, Poisson no manifestó tanta ilusión. La probabilidad de error supuesta para un jurado puede, sin aumentar el número de jueces, disminuirse sin límite para el conjunto.

... Si, como pide seriamente Cournot, invitamos al escribano a anotar, tras cada juicio, la opinión de cada uno de los jueces cuando las cifras son numerosas, con el fin de aplicar la fórmula que muestra su valía, dado que la perspicacia de cada uno es controlada por la de sus dos colegas, el juez mejor considerado de Francia sería el que, sin discutir ni reflexionar, votaría siempre como su presidente: si creemos en la fórmula, tal juez no se equivoca jamás.

Ni Cournot ni Poisson han cometido el más pequeño error como geómetras; traducen rigurosamente sus hipótesis. Pero las hipótesis no tienen la más mínima relación con la situación de un acusado delante de los jueces.

Ellos han percibido las diferencias, y creen que señalándolas adquieren el derecho a no tenerlas en cuenta.

Poisson que, como Condorcet, ha consagrado a la teoría de los juicios de valor un volumen entero lleno de los más sabios cálculos, cree atenuar las objeciones que él no deja de percibir alterando, en sus enunciados, el significado de la palabra *culpable*. Él diría que el lenguaje se volvería más exacto al sustituir la palabra *condenable*, que necesitaba alguna explicación y que continuaremos empleando para conformarnos al uso habitual.

El inocente presionado por indicios engañosos o víctima de maquinaciones tan hábiles que ningún juez podría suponerlas, es un acusado *condenable*. Poisson, para *conformarse al uso*, lo clasifica entre los culpables. El error unánime de los jueces se convierte entonces en una prueba de sagacidad de la que da cuenta el álgebra evaluando su mérito con una infalible precisión. En el conjunto de cálculos estériles que, según Stuart Mill, van a quedar como el escándalo de las matemáticas, sólo Condorcet ha dado un sabio consejo, el de elegir, para componer las asambleas, hombres *verdaderamente ilustrados*.

He querido citar estos pasajes, donde la ironía y la paradoja ocultan frecuentemente un pensamiento muy acertado; hay mucho que retener de las apreciaciones de Bertrand y es útil recordarlas. Sería interesante investigar si Condorcet, Laplace, Poisson y Cournot fueron verdaderamente tan ingenuos como Bertrand parece creer; pero no entraré en esta discusión histórica, bastante inútil para nuestro objetivo; no se trata de saber si las afirmaciones de Condorcet son correctas, o si Bertrand tiene razón en criticarlas; más bien se trata de determinar bajo qué forma se puede considerar la aplicación del cálculo de probabilidades al método de las mayorías. En este estudio, los trabajos anteriores nos resultarán, naturalmente, útiles; pero citarlos y discutirlos a cada momento sólo conseguirá recargar sin provecho la exposición.

II

El primer punto es, obviamente, que la mayoría no puede poseer las luces que faltan en cada una de las individualidades que la componen: un jurado de ciegos no sabría juzgar los colores. La imposibilidad que se debe aquí a la naturaleza del jurado, puede, en otros casos, ser inherente a la cuestión planteada: *conociendo la altura del mástil mayor, encontrar la edad del capitán;* mil personas no resolverán mejor esta cuestión que una sola.

Quizás estas observaciones parezcan superfluas debido a su evidencia; sin embargo, podemos extraer una consecuencia que es fundamental en nuestro estudio: *el valor del juicio de la mayoría puede, en ciertos casos, ser mayor que el valor del juicio individual, pero NO PUEDE SER DE CUALIDAD DIFERENTE.*

Un ejemplo muy simple hará comprender bien el sentido de este enunciado: Paul juega a las cartas y es muy inexperto; se da cuenta de que tiene la dama, el as y el nueve de corazones, que son baza, el rey de tréboles y el rey de picas. Se pregunta si debe cortar el juego y consulta a unos amigos que saben más que él: ellos le animan, pero su adversario tiene el rey, la sota y el diez de la baza y dos pequeños diamantes; Paul pierde la partida. ¿Ha hecho bien en seguir los consejos que le han dado? O, más concretamente, ¿debemos creer que estos consejos eran malos? Seguramente no; estos consejos eran los mejores posibles teniendo en cuenta que los consejeros no conocían el juego del adversario de Paul: solamente un compinche hábil y poco honesto habría podido, entreviendo este juego, dar a Paul un consejo destinado a salir mejor parado de la circunstancia particular en la que se encontraba.

Podemos resumir esto diciendo que los consejos ofrecidos a Paul por sus amigos constituyen la *verdad relativa* en las condiciones en que se encuentra Paul, es decir, en la ignorancia en la que se encuentra acerca del juego de su adversario; pero puede ocurrir que esta *verdad relativa* sea contradictoria con la *verdad absoluta*, es decir, con la manera de jugar que sería la mejor para aquél que conociera los dos juegos. Por otra parte, está bien añadir que esta verdad, que llamamos *absoluta*, es solamente *menos relativa*; puede que la pérdida de esta partida, al influir en la disposición de las cartas para la partida siguiente sea, a fin de cuentas, ventajosa para Paul, de forma que el consejero que conociera, no solamente el juego de su adversario, sino la manera en que las cartas serán recogidas y barajadas para la jugada siguiente, podría dar un consejo diferente y mejor en la práctica. Se entiende así que la palabra *absoluta* sólo significa *menos relativa*; por tanto, no puede tener otro sentido en cualquier estudio que hable de la realidad.

Admitiendo esta forma de hablar, se pueden distinguir, desde un punto de vista práctico, tres categorías principales en las aplicaciones del cálculo de probabilidades al método de las mayorías.

1.º La *verdad relativa* que se alcanza tiene una significación interesante por sí misma, hasta el punto de que es legítimo tomarla como objetivo absoluto de la investigación.

2.º La verdad relativa que se alcanza sólo tiene una relación lejana y desconocida con la verdad absoluta, que es la que interesa.

3.º La cuestión se plantea de tal forma que se duda si existe una verdad relativa; por tanto, se debe plantear, al comienzo, la cuestión de saber si esta verdad relativa existe; sólo después se podrá proponer la determinación de su naturaleza.

Estas distinciones, repito, únicamente pueden tener un valor *práctico;* desde el punto de vista teórico, todos los casos intermedios son posibles; desde el punto de vista abstracto, todos los casos son semejantes cuando la traducción matemática coincide; pero la significación útil de estas mismas fórmulas es muy diferente según los casos: la misma balanza, diría Josepth Bertrand, pesa unas veces el cobre y otras veces el oro.

III

Ocupémonos primero de los casos donde la verdad relativa, no solamente es interesante por sí misma, sino que debe ser contemplada como poseedora de un valor absoluto: son los casos donde el hecho crea el derecho.

Uno de los mejores ejemplos que se pueden ofrecer es el de las cuestiones del lenguaje (Borel 1907); por ejemplo, plantearse si tal locución es usada actualmente o si es propia de un lugar determinado: el medio más seguro es interrogar a un gran número de habitantes de ese lugar; si pudiéramos consultarlos a todos individualmente e independientemente unos de otros, el conjunto de sus respuestas, supuestamente sinceras y con un recuento exacto, proporcionaría evidentemente, por su misma definición, la respuesta más satisfactoria posible a la cuestión planteada. Pero está claro que tal consulta es prácticamente imposible; el método de las mayorías consistirá, por tanto, en consultar solamente un pequeño número de personas, en tratar de prever el resultado que daría el recuento general. Por ejemplo, para saber si los parisinos de 1908 dicen *un* automóvil o *una* automóvil[**], se consultará a veinte personas, y se tomará la opinión de la mayoría. Esto es lo que hacen con frecuencia los periódicos cuando crean una «encuesta». ¿Qué valor tiene el método? Sin entrar en detalles, es necesario para que el cálculo pueda ser correctamente aplicado a una pregunta, que el grupo total en relación al cual esta cuestión es planteada, sea sensiblemente homogéneo; esta homogeneidad jamás es absoluta debido a que dos individuos cualesquiera no son idénticos; por tanto, es necesario además que el grupo parcial que se ha elegido como representante del grupo total tenga sensiblemente la misma heterogeneidad. Por ejemplo, sería totalmente incorrecto considerar los cincuenta viajeros que se encuentran en un tranvía determinado como representantes del conjunto de los parisinos; sería más correcto tomar como representantes del

[**] Traducido del original «*un* automobile ou *une* automobile». *N. T.*

conjunto de reclutas parisinos nacidos en 1887, aquéllos que nacieron, por ejemplo, el 14 de mayo de ese año; esta selección sería mejor que la de un barrio determinado, o incluso que una elección alfabética. Suponiendo que los nombres están inscritos alfabéticamente y seleccionando, como se hace frecuentemente para los exámenes o concursos, una porción de la lista designada por sorteo, nos arriesgaríamos a encontrarnos con porciones demasiado homogéneas, personas cuyos nombres comienzan por *La* o *Le*, o bien por W o Z que pueden tener conjuntamente algunos caracteres étnicos que los diferencian de los otros.

Si las condiciones de homogeneidad se cumplen, es decir, si el grupo restricto sobre el que se lleva a cabo el experimento es verdaderamente una imagen fiel del grupo total, ¿qué conclusiones podemos extraer del estudio de ese grupo parcial? Por ejemplo, sobre 36.500 reclutas de determinada región nacidos en 1887, hay 100 que nacieron el 14 de mayo; se constata que sobre esos 100, la mayoría absoluta, 51, tiene una altura mayor o igual a 1,65 m, teniendo los 49 restantes una talla inferior; por otra parte, se constata sobre la mayoría de los 100 tal ignorancia, tal opinión, o tal juicio particular; ¿qué podemos inducir de aquí para el conjunto total del cual han sido extraídos? No es aquí el momento de tratar matemáticamente estas cuestiones; basta con indicar que es legítimo plantearlas y señalar la forma en que el cálculo proporciona la respuesta. Esta forma es necesariamente un coeficiente de probabilidad. Por ejemplo, de los casos observados se podrá concluir lo siguiente: la probabilidad de que la mayoría de los casos no observados se ajuste a la mayoría de los casos observados es de 0,999, de lo que se sigue que la probabilidad opuesta será de 0,001. En otros términos, podemos apostar 999 contra 1 a que no nos equivocamos, aunque no tengamos la certeza de no equivocarnos. Esta forma particular de afirmación es común a todas las cuestiones donde interviene el cálculo de probabilidades; ciertas mentalidades se niegan a comprenderla y prefieren declarar que quieren ignorar un cálculo que conduce a resultados también inciertos; no se dan cuenta de que esta incertidumbre es común a todas nuestras afirmaciones, y no es menos peligrosa cuando está enmascarada bajo apariencias dogmáticas (Borel, 1906:424). En realidad, desde un punto de vista práctico, una probabilidad suficientemente próxima a la unidad debe, en la práctica, ser confundida con la certeza.

Indiquemos finalmente que el cálculo de probabilidades permite el estudio de cifras observadas, extraer conclusiones sobre la homogeneidad del grupo estudiado y, como consecuencia, sobre el valor que se puede atribuir a las observaciones hechas sobre ese grupo. Supongamos, por ejemplo, que los 10 primeros reclutas examinados en el consejo de revisión del Sena han medido, cinco de ellos 1,55 m. y los otros cinco 1,75 m.; la simple constatación del hecho de que ninguno de ellos tiene una talla *próxima a la media de 1,65 m. basta para inspirar las dudas más serias sobre la exactitud* de las conclusiones que se obtendrían suponiendo que este grupo parcial es la imagen exacta del grupo total del cual se ha extraído: por el contrario, debemos estar seguros de que han influido circunstancias particulares sobre la selección de los diez individuos que componen este grupo.

IV

Seré muy breve en relación al caso donde la *verdad relativa*, que el método de las mayorías permite alcanzar, no tiene relación real, o al menos relación conocida, con la verdad absoluta que sería útil conocer. Estos casos no son interesantes, debemos señalarlos solamente a título de indicación, a fin de evitar los errores a los que su estudio podría conducir. Podemos preguntarnos si todas las aplicaciones del cálculo a las decisiones judiciales entran en esta categoría, es decir, si la distinción hecha por Poisson entre *culpable* y *condenable* no es suficiente para justificar las severidades de Joseph Bertrand. Parece que hubiera planteado la cuestión sobre una forma demasiado abstracta. Para obtener juicios equitativos está claro que antes de nada hay que seleccionar jueces ilustrados y desprovistos de pasión: pero no es así como se plantea la cuestión en la práctica, y la verdad es que es demasiado simple proponer esta solución. La cuestión no es realizar un ideal imposible de alcanzar: es llegar al menor de los males sirviéndonos de instituciones imperfectas y de los hombres falibles de que disponemos, hasta que llegue el día en el que, como todos deseamos al igual que Joseph Bertrand y otros más, las instituciones sean perfectas y los hombres infalibles.

Cuando se deja, así, el terreno de la justicia ideal para situarnos en el de los hechos, está claro que pueden existir circunstancias donde tal acusado inocente deba ser fatalmente condenado por la casi unanimidad de los jueces o miembros del jurado llamados a pronunciarse sobre su destino. Naturalmente hay que lamentar que sea así, e intentar, mejorando las garantías que el procedimiento ofrece a los acusados y mejorando sobre todo la mentalidad de los jueces, que ese caso sea cada vez menos probable, entendiendo que el deseo de mejorar la realidad no debe impedir constatarla.

Por lo tanto, es perfectamente legítimo plantearse una cuestión como la siguiente: aumentando el número de miembros del jurado y modificando la mayoría requerida para la condena ¿aumentan o disminuyen las oportunidades de que un acusado sea condenado cuando su inocencia, no solamente es real, sino que además no está enmascarada por apariencias engañosas y maquinaciones hábiles?

Cuando la cuestión se plantea así, parece fuera de dudas que los cálculos de Laplace y de Poisson no proporcionen la respuesta correcta[1]. Lo único necesario para que esta respuesta sea correcta es no modificar el sentido haciendo abs-

[1] Es sin embargo un punto que Bertrand parece haber descuidado y cuya importancia puede, a veces, ser grande: ¿el conocimiento de las condiciones requeridas para la condena puede influir sobre la opinión de los jueces? Por ejemplo, entre los miembros de la convención que votaron la muerte de Luis XVI, ¿no podemos contemplar como muy probable que haya por lo menos uno que se haya pronunciado a favor de una condena menos severa si, siendo juez único o miembro de un jurado muy poco numeroso, su voto le pareció más decisivo? Esta cuestión es distinta a la del voto secreto de los jueces, que forma parte de las disposiciones que pueden ser modificadas por la ley y las costumbres.

tracción de las diferencias concretas que existen entre los diversos casos. Por ejemplo, de un cálculo simple resulta fácilmente que es preferible para un acusado ser juzgado por un tribunal de 7 jueces, donde 5 votos son necesarios para la condena, que por un tribunal de 12 jueces donde 7 votos bastan para condenar. Pero, evidentemente, sería abusivo concluir que es preferible para un soldado acusado ser llevado ante un consejo de guerra que ante un jurado civil. Porque, independientemente de cualquier otra consideración, el mero hecho de que el voto del jurado se efectúe con escrutinio secreto es un elemento que ha sido imposible tener en cuenta en los cálculos.

Entre los casos en que no hay *nada* que encontrar, y donde, como consecuencia, el método de las mayorías y todo método estadístico pierde sus derechos, es útil mencionar los juegos de azar, sin contar no obstante con que los jugadores hagan demasiado caso, pues se trata de gentes con prejuicios tenaces y más dispuestos a escuchar a los prometedores de martingalas que a los consejos del álgebra. Es una lástima ver cómo investigadores serios pierden, con investigaciones igualmente vanas, un tiempo que podría ser mejor empleado. Hago alusión aquí a un trabajo reciente del señor Charles Henry en el que, bajo una cubierta de distinción entre la verdad psicológica y la verdad matemática, reedita sofismas que podíamos considerar definitivamente superados hoy en día (Henry, 1908). La cuestión de saber si se puede ganar dinero seguro jugando a la ruleta, es una cuestión esencialmente práctica y concreta; este asunto pertenece esencialmente a la ciencia y cualquiera que sea la opinión que se profese acerca del interés de especulaciones metafísicas sobre el psicologismo, se debe admitir que no tienen nada que ver con la solución práctica de esta cuestión, solución que fue proporcionada hace mucho tiempo por Bernoulli y confirmada por todos los matemáticos: un jugador que juegue indefinidamente a un juego equitativo llega forzosamente a la ruina al cabo de un tiempo más o menos largo; este tiempo se vuelve muy corto si el juego no es equitativo, que es siempre el caso en la práctica.

V

Llegamos ahora al caso más interesante para los psicólogos, porque se presenta a menudo en sus investigaciones, a pesar de que la cuestión no sea siempre planteada por ellos bajo la forma que vamos a darle. Sea un conjunto de fenómenos observados por individuos diversos y después clasificados de tal manera que se extraiga una mayoría; el primer problema que se plantea es el siguiente:

¿La opinión de esta mayoría corresponde a alguna realidad?

Determinar cuál es esta realidad y si, en particular, se adecúa o no a la que se habría podido esperar instituyendo los experimentos, es un segundo problema del que ignoro totalmente la importancia, pero que es esencial distinguir del primero porque solamente cuando el primero se resuelva, el segundo podrá ser abordado de manera fructífera.

Precisemos esta distinción con un ejemplo; supongamos que hemos leído en voz alta una misma página en un gran número de escuelas primarias y se ha pedido a cada niño que indique por escrito la palabra que recuerdan con mayor frecuencia de esa página[2]. Si las tres cuartas partes de las respuestas coinciden en designar la misma palabra, se puede concluir con certeza que esta coincidencia no es fortuita, sino que tiene una razón. Esta razón puede ser el hecho de que la palabra designada es efectivamente aquélla que figura con mayor frecuencia en la página leída; puede también ser cualquier otra, pero es la existencia misma de una causa lo que constituye el resultado esencial proporcionado por el método de las mayorías en este ejemplo particular.

Igualmente, si hacemos comparar a numerosos experimentadores los pesos de dos objetos A y B cuyas formas y dimensiones son muy diferentes, y tres cuartas partes de las respuestas obtenidas coinciden en atribuir un peso más elevado al objeto A, se puede concluir que hay una razón para ello, pero que esta razón sólo puede ser que A sea efectivamente más pesado que B. Por el contrario, si sobre 1.000 experiencias, solamente 520 declaran como más pesado a A y 480 como más pesado a B (suponiendo que las respuestas dudosas se dejan de lado), la diferencia observada es la que se obtendría frecuentemente en una serie de 1.000 respuestas extraídas a cara o cruz; por tanto, lo único que se puede concluir es una fuerte suposición de que una nueva serie de 1.000 experimentos proporcionaría un resultado también incierto.

Antes de llegar a la discusión de los experimentos con resultados numéricos, quisiera señalar brevemente una experiencia muy interesante en la que el resultado se traducía en un dibujo. Se trata de observaciones de la superficie de un planeta hechas en un mismo instante, por observadores alejados los unos de los otros que transmitían cada uno un dibujo de su observación. Estos dibujos han sido reunidos y comparados por el señor Jean Mascart, del Observatorio de París, en el interesante estudio al que remito (Mascart, 1907). El experimento sugerido por el señor Nicolas Poutiata ha sido organizado por la Sociedad Astronómica de Francia, con la proposición del señor Camille Flammarion; el programa se estableció cuidadosamente y se realizó con la participación de 36 observadores, disponiendo de medios de acción muy diversos; duró 19 días (del 2 al 20 de enero de 1906), y aunque el mal tiempo impidió, naturalmente, la participación de ciertos observadores durante algunos días, el número de observaciones hechas un mismo día a la misma hora sobrepasa a menudo las 10 y se eleva hasta 17. No podemos entrar aquí en la discusión detallada a la que se consagra el señor Jean Mascart, ni en la descripción de los distintos procedimientos fotográficos empleados para *tomar la media* de dibujos diversos (señalemos no obstante la ingeniosa idea de proporcionar el tiempo de acción fotográfica, bien se deba al diámetro, o bien a la superficie del objetivo empleado en la observación particular). Contentémonos reteniendo la impresión que resulta de la

[2] Una experiencia análoga se ha efectuado recientemente por el señor Binet. Se leía a los niños 100 sustantivos y se les pedía escribir la lista de las palabras de las que se acordaban. No he contado con los elementos suficientes para tratarlo por medio del cálculo; tendré ocasión de volver sobre ello.

simple inspección de los dibujos yuxtapuestos: estos dibujos ejecutados el mismo día a la misma hora, son muy diferentes los unos de los otros y muy pocos detalles se reconocen sobre muchos de ellos. Esta simple constatación debe ser muy prudente en la interpretación de dibujos de esta naturaleza; es evidente que observadores cuya buena fe no se duda y que saben, además, que serán sometidos a un control, creen ver apariencias que no corresponden a ninguna realidad. ¿Se deben estas apariencias a las condiciones atmosféricas o a la imaginación del observador? Esto es lo que los nuevos experimentos podrían, quizás, determinar. En todo caso, la experiencia merece ser retomada, porque es tan interesante para el psicólogo como lo es para el progreso de la astronomía planetaria; en estos nuevos experimentos, se obtendrá gran provecho de muchos de los juiciosos comentarios de Jean Mascart.

VI

Entre las observaciones psicológicas que fueron ocasión de desarrollos matemáticos ligados al cálculo de probabilidades, se debe mencionar particularmente el método extrañamente denominado método *de los casos verdaderos y de los casos falsos Richtig und falsch; Right and wrong).* Este método ha dado lugar a numerosos cálculos en los cuales interviene la integral de Gauss y fórmulas relacionadas; no es aquí el lugar de hacer una exposición crítica, que exigiría un aparato matemático relativamente considerable; quizás tenga la ocasión de volver sobre ello en otro lugar; me contentaré con señalar aquí que el punto delicado es encontrar una expresión matemática del error elemental que satisfaga por igual las condiciones de continuidad y de discontinuidad inherentes a la naturaleza de la cuestión; las tentativas realizadas en este sentido no me parecen enteramente satisfactorias desde el punto de vista teórico y, por otra parte, desde el punto de vista práctico, conducen rápidamente a complicaciones analíticas que sería preferible evitar en relación a cuestiones que interesan a muchas personas cuya cultura matemática es forzosamente limitada.

Quisiera sobre todo indicar cómo el empleo sistemático del método de las mayorías permite plantear la cuestión bajo una forma bastante diferente de la forma usual: en lugar de interesarse, como parece que se ha hecho sobre todo hasta aquí, por el estudio de la sensación individual, que se pretendía desgajar del conjunto de experimentos, es la precisión de la «sensación colectiva» lo que me parece interesante separar de este conjunto. Consideremos un ejemplo concreto. Uno de los objetivos que los experimentadores de psicología se proponen es, por ejemplo, determinar cuál sería el aumento mínimo de peso que podría apreciar un observador medio; dado un peso de referencia de 1.000 gramos, se constata que es únicamente cuando el peso a juzgar es inferior a 950 gramos o superior a 1.050, cuando generalmente se percibe de forma notable la diferencia[3]; se dirá

[3] Dejo totalmente de lado la discusión de los errores sistemáticos que se introducen en los experimentos de este género; supongo, por ejemplo, que se elimina el error relativo al orden de las

entonces que 50 gramos representan la diferencia mínima perceptible en relación a 1.000 gramos; se ha trabajado mucho en la determinación de esta diferencia mínima después de la ley de Weber. Pero la cuestión se puede plantear de otra manera; en lugar de proponer el estudio de la ley de *un* experimento, se puede investigar qué precisión se podría obtener de un gran número de experiencias, y se observa que esta precisión sería mayor aunque, evidentemente, no sería indefinida. Supongamos, por ejemplo, que se hiciera comparar a un gran número de observadores dos pesos, uno de 1.000 gramos y otro de 1.010 gramos[4]; si, sobre 1.000 observadores hay 600 que proporcionan una solución exacta, será legítimo decir que este conjunto de 1.000 observadores tiene la percepción colectiva de la diferencia de 10 gramos, aunque si recomenzáramos muchas veces el experimento, puede ocurrir que cada uno de los 1.000 observadores se equivoque alguna vez, de forma que ninguno de ellos puede ser considerado como poseedor de la percepción individual de esta diferencia. He supuesto aquí que se exigía una respuesta afirmativa o negativa, excluyendo las respuestas dudosas; sería fácil plantear el problema de una manera análoga en el caso en que se admitieran estas respuestas dudosas. Supongamos, por ejemplo, que se dispone de una escala de peso de comparación, marcada en cifras variando de gramo en gramo desde 900 hasta 1.100 gramos; se da al observador un peso P de valor desconocido y se le pide señalar entre varios cuáles le parecen iguales, cuáles superiores y cuáles inferiores al peso a determinar. Pasando incluso aquí por encima de los detalles del experimento —y no es que desconozca su gran importancia, pero debo centrarme— supongo que se deduce de su resultado el valor más probable del peso P[5]. ¿Qué confianza se puede otorgar a tal determinación individual? ¿Y qué confianza se puede atribuir a la media de 1.000 determinaciones análogas hechas por observadores diferentes? Si se acepta de manera ciega la ley de Gauss o cualquier otra ley matemática precisa para expresar los errores individuales, llegaremos forzosamente a la conclusión de que la precisión crece como la raíz cuadrada del número de observaciones y puede, en consecuencia, agrandarse tanto como se quiera. Este método de estimación llegaría así a poder sobrepasar, si se emplea con cuidado, la precisión de las mejores medidas hechas a la media de instrumentos, lo que en apariencia es paradójico. Sería interesante determinar, gracias al experimento, la precisión colectiva que puede esperarse por medidas estimativas. Es éste un problema al que no me parece indigno que los observadores presten atención; se encontrarían elementos para

---

comparaciones haciendo cada comparación dos veces; dejo también de lado la definición precisa de la desviación *generalmente percibida*; si se admite que las desviaciones percibidas siguen la ley de Gauss, se puede suponer, tal y como sabemos, que la desviación probable o la desviación mediana están bastante próximas la una de la otra.

[4] Por supuesto, supongo siempre que los errores sistemáticos son evitados.

[5] Los métodos matemáticos para alcanzar este objetivo son muy simples si se supone el intervalo bastante débil para poder considerar el mínimo perceptible como constante; los métodos se complican si se tiene en cuenta su variación siguiendo la ley de Weber.

comenzar el estudio en publicaciones ya hechas y, sin duda, en los cuadernos de experimentos[6].

Para dar una idea del camino que se podría seguir, tomo prestadas las cifras siguientes del *Manual de psicología experimental* de Titchener[7] (Titchener, 1901:107). Un peso S de 1.071 gramos se comparaba con diversos pesos C —un total de 7—, cuyos valores crecían de 921 a 1.221 gramos en intervalos de 50 gramos. La tabla siguiente ofrece los resultados de 100 observaciones para cada peso C, siendo 700 observaciones en total; encontramos en las tres columnas los números respectivos de las observaciones donde se estimó a C inferior, igual o superior a S.

Valor de C	**C < S**	**C = S**	**C > S**
921	1	8	91
971	6	11	83
1.021	10	37	53
1.071	30	37	33
1.121	55	34	11
1.171	76	20	4
1.221	90	9	1

Si se examina esta tabla desde nuestro punto de vista, es decir, si tratamos de deducir del conjunto de las observaciones la relación real supuestamente desconocida entre C y S, se percibe inmediatamente que esta relación aparece sin ambigüedad; la lectura de las tres últimas columnas indica, sin que sea necesario observar la primera, si C es inferior, igual o superior a S. Por supuesto, la igualdad rigurosa no puede establecerse, sino simplemente suponerse según el resultado de los experimentos. Del hecho de que sobre 100 experimentos, se estimó que C era 30 veces inferior, 37 veces igual y 33 veces superior, podemos simplemente concluir esto: no hay ninguna razón[8], a la luz de los experimentos, para suponer C > S más que C < S; si no tenemos otros elementos de información, se

[6] Según la información proporcionada por el señor Ch.-Éd. Guillaume, director adjunto de la Oficina Internacional de Pesos y Medidas, se ha llegado, multiplicando las lecturas hechas por observadores experimentados, a alcanzar en las medias una precisión *mucho mayor* que la precisión de las observaciones individuales; por ejemplo, por medio de las lecturas hechas con lupa sobre una escala milimétrica, se ha alcanzado con certeza la *centésima de milímetro* con un número suficiente de observaciones. Me propongo hacer un estudio de estos resultados, en gran parte inéditos, respecto al punto de vista desarrollado aquí.

[7] He sumado los números correspondientes a dos series de experiencias (Time Order I, S first y Time Order II, S second).

[8] En realidad, la pequeña diferencia entre 30 y 33 ofrecería una débil razón para suponer que C > S; se trata de un problema de probabilidad de las causas que puede tratarse, siguiendo el método clásico, con la ayuda de una hipótesis suplementaria sobre la probabilidad *a priori*. Pero no hay necesidad de cálculos, por poco familiarizado que se esté sobre estas cuestiones, para estar seguros de que la probabilidad de que C > S sobrepasa muy poco 0,5 y, por tanto, está muy cerca de la probabilidad de que C < S.

debe entonces suponer que C = S. De igual forma, la mayoría distingue con seguridad un peso C de 1.021 gramos del peso S de 1.071 gramos, a pesar de que cada uno de los observadores pueda equivocarse individualmente; porque dejando de lado las 37 respuestas que afirman la igualdad o la duda, quedan 53 respuestas exactas contra 10 inexactas; tal reparto no puede ser efecto del azar; si en un juego de cara o cruz, de 63 pruebas, sale 53 veces cruz y solamente 10 veces cara, la certeza de que el juego esté trucado es casi absoluta.

Pero se puede ir más lejos preguntándonos qué ocurriría si C tuviera un valor intermedio entre 1.021 gramos y 1.071 gramos; vista la regularidad de la variación de los números inscritos en las tres columnas, se comete un error bastante débil procediendo por interpolación lineal; es decir, suponiendo que los números varían en progresión aritmética, ofreciendo los resultados siguientes[9]:

C	**C < S**	**C = S**	**C > S**
1.021	10	37	53
1.031	14	37	49
1.041	18	37	45
1.051	22	37	41
1.061	26	37	37
1.071	30	37	33

Obviamente, si el experimento se realizara realmente los números no serían exactamente los que hemos escrito debido a los errores fortuitos; pero el cálculo de probabilidades nos indica cuáles son los errores fortuitos que pueden ser contemplados como probables y cuáles son altamente improbables. Por ejemplo, sobre 63 estimaciones (que es nuestro caso aquí puesto que hay 37 neutros) obtenidas al azar, se debe observar como poco probable que haya más de 40 en un sentido y solamente 23 en el sentido opuesto[10].

El conjunto de las observaciones hechas en el caso en que C = 1.051 permite concluir con mucha verosimilitud que C < S, es decir, en el sentido de la verdad. Del conjunto de números de nuestra primera tabla, extraemos por tanto la conclusión de que el conjunto de los 100 observadores, considerando su juicio colectivo, llega a distinguir una diferencia de peso de 20 gramos. Pero, para cálculos bastante largos y muy correctos, Titchener encuentra para los valores de DL[11] deducidos de esa misma tabla, las cantidades siguien-

---

[9] Teniendo en cuenta las segundas diferencias, se obtienen números bastante próximos que conducen a las mismas conclusiones finales; he preferido presentar el cálculo bajo la forma más intuitiva.

[10] La probabilidad es $1-\Theta\left(\frac{9}{\sqrt{32}}\right) = 1-\Theta$ (1,6 = 0,024 (designando Θ la integral de Gauss).

[11] Es decir, del mínimo incremento medio perceptible para el observador individual.

tes[12]: 43, 36, 30, 53, 38, 34, cuya media de 39 es más o menos el doble del número 20 que encontramos (Titchener, 1901:109-113). Esta diferencia de resultados deducidas de los mismos experimentos muestra bien que los problemas son distintos: Titchener determina la sensibilidad individual mientras que nosotros determinamos la sensibilidad colectiva del conjunto de los observadores. Por otra parte, sería posible precisar más nuestra determinación, pero sería a costa del rigor, en vista de la pequeña cantidad de números empleados[13]. Es mejor atenerse a la clara constatación de que la diferencia de 20 gramos es perceptible casi con toda seguridad por el método de las mayorías aplicado al conjunto de 100 observaciones; así, este método proporciona, como se podría esperar, una precisión bastante más grande que la observación individual.

Me conformaré con este ejemplo, esperando que baste para mostrar el interés que podría tener la aplicación del método a experimentos combinados especialmente en vista a esta aplicación. Sin duda, ello conducirá a resultados importantes en distintos ámbitos.

VII

Para terminar, quisiera discutir algunas aplicaciones del método de las mayorías a cuestiones cualitativas, porque quizás sea fundamentalmente aquí donde el método proporciona resultados que no podrían ser obtenidos por ningún otro medio.

La primera idea de este tipo de aplicación me la sugirió una investigación que realicé sobre textos que habían servido para los experimentos grafológicos del señor Binet (Borel, 1906:244-366), que trataban de distinguir, entre 12 textos muy cortos reproducidos tipográficamente, cuáles de los 12 escritores eran hombres superiores y cuáles eran mediocres[14]. Las respuestas individuales fueron interesantes, pero la clasificación que se deducía del recuento de votos lo fue todavía más; los seis hombres superiores fueron así declarados por la mayoría absoluta; un solo hombre mediocre obtuvo también la mayoría absoluta y era, por así decirlo, el menos mediocre de los hombres «mediocres» (o «simplemente

[12] Las diferencias entre estos números son debidas, por una parte a la distinción entre los dos «time order», por otra a la distinción entre los incrementos positivos y negativos y, finalmente, a las diferencias entre los métodos seguidos; estas diferencias entre los resultados muestran cuán ilusorio es aplicar métodos de cálculo demasiado precisos a cantidades forzosamente aproximadas.

[13] Si se sigue un método paralelo al que se emplea generalmente en estas cuestiones cuando se usa la integral de Gauss, se buscaría el valor de C para que la probabilidad de la desviación observada sea inferior a $^1/_2$; se deducirían así cantidades dadas de forma que el valor de DL, para el conjunto de los 100 observadores considerados colectivamente, fuera aproximadamente de 8 gramos; pero el número de observaciones es demasiado débil para que el resultado pueda considerarse preciso.

[14] La clasificación se debe al señor Binet; ver su libro: *Les Révélations de l'écriture soumises à un contrôle scientifique*.

inteligentes») seleccionados por el señor Binet, que señala: «es un publicista científico al que no le falta mérito y que, si se reconoce, no querrá que se le haya preferido a Claude Bernard». La opinión agregada de la mayoría se ajustaba a la verdad para todos los hombres superiores y el único error que cometía clasificando como superior a uno de los hombres mediocres podía justificarse por la propia información que se tenía sobre él. No insistiré más en este experimento, pero quise recordarlo porque fue con ocasión del mismo que mantuve correspondencia con el señor Binet sobre el tema de este método de las mayorías (es él el primero en emplear esta expresión), y quiso invitarme a escribir este pequeño estudio para su *Année*.

* * *

Discutiré en detalle un interesante experimento realizado por la señora Rousson sobre la lectura de una fisonomía del niño (Rousson, 1906:76). Recordemos brevemente las condiciones dejando de lado los detalles. Se suministra separadamente a 20 observadores 40 fotografías de niños (cada una doble; haz y envés), de las cuales 23 eran retrasados y 17 normales; cada observador declaraba al niño retrasado o normal, o se negaba a responder (hubo solamente 18 negativas a responder sobre 800 observaciones[15]). Los porcentajes de respuestas correctas varían, siguiendo a los observadores, entre 67% y 92%; la media fue de 78%. De aquí se puede concluir que en las fotografías[16], o al menos en algunas de ellas, había índices de inteligencia. Pero esta conclusión general puede precisarse mucho aplicando el método de las mayorías. Quisiera mostrar en detalle, sobre este ejemplo, qué principios simples se deben utilizar en esta aplicación, procurando recurrir al mínimo de conocimientos matemáticos.

El número de votantes era 20; cada uno de ellos debía votar: *normal* (N) o *retrasado* (A)[***], de forma que el conjunto de votos relativos a una misma fotografía se representaba por una secuencia como la siguiente[17] (fotografía n.º 7):

N N A A A N A N A A A N A A A A A A A A

[15] Estas negativas son lo suficientemente poco numerosas para poder obviarse. La señora Rousson no las tiene en cuenta en el porcentaje de respuestas correctas, es decir, que si un observador se niega a responder 2 veces y se equivoca 3 veces sobre 40 niños, ella cuenta 35 respuestas correctas sobre 38. Este procedimiento aumenta ligeramente los porcentajes de la señora Rousson; los hemos reproducido tal y como ella los da.

[16] Digo *fotografías*, y no solamente *fisonomías*, porque uno de los observadores subrayó que habría sido mejor «descartar todas las diferencias exteriores que pueden influir en el juicio». Esta observación no tiene importancia para nuestras deducciones; debe intervenir en las conclusiones especiales relativas a la fisonomía que se podrían extraer.

[***] Las siglas N y A corresponden respectivamente a *normal* y *arriéré* en el texto original. N. T.

[17] Dejo de lado algunos votos *nulos* (un total de 18 sobre 800); para dar mayor certeza a mis conclusiones los contaré con la *minoría*; es el procedimiento más desfavorable y, con cualquier otro convenio, las conclusiones serían verdaderas *a fortiori*. En vista de su pequeño número, el asunto es, por otra parte, de poca importancia.

¿Cuántas secuencias de 20 letras N o A como esta pueden existir? La primera letra puede ser N o A, lo que lleva a dos posibilidades; para cada una de estas dos posibilidades, la segunda letra puede ser N o A, de lo que surgen 2×2 posibilidades (NN, NA, AN, AA); para cada una de estas 2×2 posibilidades, la tercera letra puede ser N o A, lo que lleva a $2 \times 2 \times 2$ posibilidades (NNN, NNA, NAN, NAA, ANN, ANA, AAN, AAA), etc. Utilizando el mismo razonamiento hasta la vigésima letra, encontramos un número de posibilidades igual al producto de 20 factores equivalente a 2, es decir a 2^{20} o a 1.048.576. Este es la cantidad de las distintas posibilidades del escrutinio. Si cada votante, en vez de deliberar su voto, se lo juega a cara o cruz, cada una de estas posibilidades sería igualmente probable y, como consecuencia, habría exactamente 1 oportunidad sobre 1.048.576 de obtener cualquiera de ellas (por ejemplo, la que hemos señalado anteriormente).

Ahora hay que determinar cuántas de las diversas posibilidades del escrutinio dan a N un cierto número de votos; está muy claro, por ejemplo, que N puede tener un voto contra 19 de A, de 20 maneras diferentes, porque el único voto N puede ser el del primer votante, o el del segundo... o el del vigésimo. La solución completa de esta cuestión se proporciona por el análisis combinatorio, y el cálculo numérico se hace simplemente por el algoritmo llamado «triángulo aritmético de Pascal»; me limitaré a indicar el resultado.

Se puede obtener	De un número de formas igual a	Suma de números de la columna precedente
20 N y 0 A	1	21
19 N y 1 A	20	211
18 N y 2 A	190	1.351
17 N y 3 A	1.140	6.196
16 N y 4 A	4.845	21.700
15 N y 5 A	15.504	60.460
14 N y 6 A	38.760	137.980
13 N y 7 A	77.520	263.950
12 N y 8 A	125.970	
11 N y 9 A	167.960	
10 N y 10 A	184.756	
9 N y 11 A	167.960	
8 N y 12 A	125.970	

No tiene sentido escribir el final de la tabla que reproduce, evidentemente, el comienzo en orden inverso. En la tercera columna tenemos inscrita la suma de las cantidades de la segunda (21.700, por ejemplo, es la suma de 1 + 20 + 190 + 1.140 + 4.845 + 15.504); se verá enseguida cuál es su utilidad.

La tabla anterior es el elemento esencial de la discusión del método de las mayorías; cuando el número de votantes es pequeño, se obtiene sin dificultad con la ayuda del triángulo aritmético de Pascal[18]; cuando es grande, los cálculos serían inextricables y se utilizan fórmulas aproximadas establecidas por el cálculo integral que conducen a considerar la célebre integral de Gauss; este es el significado de esta integral, que no tiene nada de misterioso pero que, para el ejemplo actual, me parece preferible evitar su empleo.

El número total de posibilidades es 1.048.576; si los votos se jugaran a cara o cruz, lo que haría que todas las posibilidades fueran igualmente probables, la probabilidad para que se produzca una de las 38.760 combinaciones que da 14 N y 6 A sería igual al cociente de 38.760 entre 1.048.576, es decir, en cifras redondas, al 4%. Se pueden así deducir de la tabla precedente los resultados probables que darían los votos emitidos al azar; los tenemos aquí, redondeando las cifras:

Se obtendría	20N y 0A	....	1	vez sobre	1.000.000
—	19N y 1A	....	1	—	50.000
—	18N y 2A	....	1	—	2.500
—	17N y 3A	....	1	—	1.000
—	16N y 4A	....	1	—	200
—	15N y 5A	....	1,5	—	100
—	14N y 6A	....	4	—	
—	13N y 7A	....	8	—	
—	12N y 8A	....	12	—	
—	11N y 9A	....	17	—	
—	10N y 10A	....	18	—	

Resulta que, sobre 40 escrutinios se debería esperar, en líneas generales, los siguientes resultados:

7	veces cada una de las combinaciones		11 N y 9 A, 10 N y 10 A, 9 N y 11 A
5	—	—	12 N y 8 A, 8 N y 12 A
3	—	—	13 N y 7 A, 7 N y 13 A
3	veces en el total de ambas combinaciones		14 N y 6 A, 6 N y 14 A

[18] Este triángulo tiene la siguiente disposición:

1								
1	1							
1	2	1						
1	3	3	1					
1	4	6	4	1				
1	5	10	10	5	1			
1	6	15	20	15	6	1		
1	7	21	35	35	21	7	1	
1	8	28	56	70	56	28	8	1

El número 56, por ejemplo —el cuarto de la última línea— se obtiene sumando el número 35 situado encima de éste más el número 21 que precede al 35; los números de nuestra tabla formarían la 21.ª línea del triángulo, comenzando por 1, 20, seguidos de 190, 1140, etc.

La ocurrencia de las otras combinaciones no es imposible, pero es bastante poco probable para 15 N o 16 N, muy poco probable para 17 N y 18 N, y totalmente improbable para 19 N o 20 N. En particular, esta última combinación (la unanimidad) sólo debería producirse 1 vez de cada 1.000.000 porque, en realidad, se produce 5 veces y cada vez en la dirección de la verdad[19]. Atribuir este acontecimiento al azar es tan improbable como admitir el hecho siguiente: en una ciudad como Londres que comprende alrededor de 1.000.000 hombres adultos, 40 individuos investidos del poder supremo son forzados a dimitir para otorgar el poder a 5 ciudadanos escogidos por sorteo entre el millón, y el sorteo designa precisamente a 5 de los dimitidos. El pueblo estará seguro de que hay fraude; y tendría razón. De igual forma, debemos estar seguros de que no ha sido sólo el azar el que ha producido el resultado de la votación; hay por tanto una causa, y es bastante natural pensar que sea la causa efectivamente investigada, es decir, la inspección de las fotografías (podría ser también un acuerdo entre los votantes, o una sugerencia, pero parece que deben descartarse estas hipótesis según la información disponible sobre las condiciones del experimento).

Pero entremos en el detalle de los resultados; los reagruparemos en exactos, dudosos e inexactos.

I. Resultados exactos[20]

5 a la unanimidad de 20 votos
5 con 19 votos contra 1 (o contra 1 abstención)
6 con 18 votos contra 2 (o abstenciones)
2 con 17 votos contra 3 (—)
4 con 16 votos contra 4 (—)
3 con 15 votos contra 5 (—)
2 con 14 votos contra 6 (—)

27 resultados exactos en total

II. Resultados dudosos

2 a la mayoría de 13 votos
2 — 12 —
1 — 11 —
4 a la igualdad de los votos (10 contra 10)
1 a la minoría de 9 votos (contra 10 y una abstención)
1 5 votos[21] (contra 12 y 3 abstenciones)

11 dudosos

[19] Debido a que ocurre en la dirección de la verdad decimos 1 vez de cada 1.000.000 y no 2 veces de cada 1.000.000, que sería la probabilidad de que se produjera la unanimidad en un sentido no indicado de antemano (siendo 20 N, 20 A).

[20] Estos resultados exactos o dudosos se reparten de forma más o menos igual.

[21] Este resultado también podría haberse colocado entre los inexactos; es dudoso si se añaden las abstenciones a la minoría.

III. Resultados inexactos

1 donde la solución exacta tiene solamente			5 votos contra		15
1	—	—	3	—	17
2 inexactos					

Los resultados clasificados como dudosos son aquéllos para los cuales *no es improbable que la exactitud del resultado sea debida al azar*. Su número es 11; si vamos a la tercera columna de la tabla de la página 143 se ve que sobre 1.024.576 pruebas regidas por el azar, hay 60.460 para las cuales la mayoría alcanza o sobrepasa 14, y 137.980 para los cuales esta mayoría alcanza o sobrepasa 13; por tanto, sobre 11 pruebas regidas por el azar, el número probable de aquéllas para las cuales la mayoría en un sentido dado de antemano sería igual o superior a 14, es inferior a 1, mientras que el número de aquéllas para las que la mayoría alcanza 13 es superior a 1. Debido a ello hemos clasificado los 2 escrutinios donde la mayoría alcanza 14 entre los resultados exactos, y los 2 escrutinios donde la mayoría alcanza 13 entre los dudosos. Hay, evidentemente, una cierta arbitrariedad en esta demarcación absoluta; sería pueril disimular esta arbitrariedad, pero sería igualmente excesivo sacar argumentos contra las conclusiones en lo que tienen de general; simplemente, no hay que olvidar que estas conclusiones implican algún tipo de juego, es decir, que en lugar de 27 resultados exactos sobre 40, se podría contar de 25 a 29, quizás incluso de 24 a 30; pero esta ligera incertidumbre en la precisión numérica no mancilla el principio mismo de las deducciones: en la propia naturaleza de todo cálculo basado en experiencias está el hecho de proporcionar únicamente resultados aproximados. Hechas estas reservas, consideraremos como exactos los números de la página 186.

Vemos que, para 27 fotografías sobre 40, es decir, para alrededor de *los dos tercios*, hay que admitir que el resultado del escrutinio ha sido determinado por el aspecto de las fotografías; es totalmente probable que estas 27 fotografías sometidas a otros observadores de las misma naturaleza (profesores y profesoras) hayan conducido a los mismos resultados; sería interesante probar el experimento.

Para 13 fotografías, es decir para *un tercio*, el conjunto de las respuestas no da una exactitud superior a lo que habría proporcionado el azar. Al parecer, debemos concluir que para estas fotografías los caracteres sobre los que se basan en general las observaciones a la hora de emitir un juicio, estaban ausentes o eran parcialmente contradictorias. En resumen, parecería que, para un tercio de las fotografías, el experimento deba interpretarse como conducente a la conclusión de que los caracteres de estas fotografías no permiten determinar a qué categoría pertenece el niño. Tal es la consecuencia lógica de este experimento; me apresuro a añadir que esta consecuencia podría ser parcialmente modificada por un experimento más extenso sobre las mismas fotografías; estoy convencido de que no se invalidaría totalmente.

Finalmente, vamos a hablar brevemente de dos escrutinios cuyo resultado es inexacto, a los cuales quizá se deba añadir uno de los que se clasificaron como dudosos. Se trata de tres niños retrasados que fueron declarados normales por 17 votos contra 3, 15 contra 5 y 12 contra 5. Hay presunciones bastante serias de que al menos dos de las fotografías (cuando no las tres) poseen las características de los niños normales; es decir, de que el error de la mayoría corresponda a una verdad relativa, aunque sea contradiciendo la verdad absoluta (es decir, la apreciación de las personas que conocen muy bien a los niños). Podríamos preguntarnos si estas fotografías condujeron a conclusiones inexactas porque son más difíciles de juzgar, y si jueces más hábiles llegarían o no a la verdad acerca de ellas. No lo parece, debido a que aquéllos observadores que han dado una apreciación exacta sobre esas fotografías no parecen distinguirse por ninguna habilidad particular; los porcentajes respectivos de respuestas exactas para aquellos observadores que han respondido bien en las tres votaciones al resultado inexacto, son:

82%, 75%, 67%, 78%, 67%
82%, 67%, 92%
77%, 76%, 67%, 82%, 75%

Estos porcentajes oscilan entre el mínimo general y el máximo general de los 20 observadores (67% y 92%); su media, 76%, es ligeramente inferior a la media general, 78% de los 20 observadores. No son, por tanto, los observadores más hábiles lo que han respondido mejor en este caso.

Como discusión general de esta aplicación del método de las mayorías a la experiencia de la señora Rousson, llegamos a la conclusión de que las fotografías encierran los elementos de una solución exacta en aproximadamente el 65% de los casos[22], deben dejar en la duda en alrededor de un 30% de los casos, y deben conducir a una solución inexacta (retrasado tomado por normal) en alrededor del 5% de los casos.

* * *

He aplicado una crítica análoga a un experimento muy interesante del señor Binet sobre fotografías de manos[23]. Se trataba de determinar, sobre una fotografía de la mano de un niño (o más bien dos fotografías: palma y reverso), 1.º cuál era el sexo del niño; 2.º si el niño era inteligente o torpe. La respuesta a esta segunda cuestión, planteada así por el señor Binet, debía proporcionarse sin ambigüedad. El experimento se realizaba con 20 fotografías sometidas a 20 observadores; yo mismo planteé a 4 observadores la cuestión del señor Binet relativa

[22] Esto no contradice el hecho de que el porcentaje general de respuestas exactas haya sido del 78%, porque en el caso dudoso hay alrededor de una oportunidad sobre dos de que la respuesta sea exacta; si todos los casos fueran dudosos, el porcentaje general de las respuestas exactas sería del 50%.

[23] Este experimento es inédito: agradezco al señor Binet que me haya transmitido su dossier.

a la inteligencia, y he añadido, en lo que sigue, estas cuatro respuestas a las 20 obtenidas por el señor Binet[24]; aquí estás los resultados:

1.º *Cuestión relativa al sexo.*

13 respuestas exactas:

2 con 19 votos sobre 20 votantes, 1 con 18 votos, 6 con 17 votos,
1 con 16 votos, 1 con 15 votos, 1 con 14 votos

5 respuestas dudosas:

1 con 13 votos, 2 con 12, 1 con 11, 1 con 7

2 respuestas inexactas:

5 votos por la verdad contra 15

Las conclusiones son más o menos parecidas a las que extrajimos del experimento de la señora Rousson salvo que, sin embargo, las mayorías en las respuestas exactas son menos considerables; las fotografías de las manos revelan el sexo un poco peor que las fotografías de las fisonomías revelaban la inteligencia. Hay que subrayar que las dos respuestas inexactas se relacionen con dos niños tomados por niñas: estos dos chicos son inteligentes: uno de ellos ha sido juzgado así por 18 votos sobre 24 y el otro ha obtenido en esta segunda encuesta 12 votos sobre 24.

2.º *Cuestión relativa a la inteligencia.*

6 respuestas exactas:

1 con 19 votos sobre 24, 2 con 18, 1 con 17 y 1 con 16;

Estas respuestas exactas se refieren a dos niños juzgados inteligentes y 4 juzgados torpes.

12 respuestas dudosas:

Solución exacta habiendo obtenido, sobre 24 votantes, 3 veces 14 votos,
4 veces 12 votos, 2 veces 11 votos, 2 veces 10 votos, 1 vez 9 votos

2 respuestas inexactas:

Solución exacta no habiendo obtenido más que 7 votos sobre 24.

Estas dos respuestas inexactas se refieren respectivamente a las dos categorías.

[24] Esta adición disminuye en una unidad el número *de dudosos;* el sentido general del resultado queda igual. En el dossier que me transmitió el señor Binet, he utilizado, para cada observador, las respuestas de la primera prueba, sin tener en cuenta las modificaciones en los juicios de la segunda prueba, de lo que hablaré brevemente a continuación.

Estos resultados son, como cabría naturalmente esperar, muy inferiores en exactitud a los que proporciona el estudio de las fisonomías; la mayoría nunca es neta y, en las dos terceras partes de los casos, las desviaciones son del orden de lo que se podría esperar del azar. Parece, por tanto, que hay alrededor de un 60% de los casos de los que poco se puede deducir de las fotografías de las manos. Lo que corrobora esta impresión es el estudio de un segundo experimento en el cual un cierto número de los mismos observadores variaron sus juicios sobre las mismas fotografías (en esta segunda prueba conocen el sexo de los niños); estas variaciones de juicio son un total de 62, de donde 46 eran sobre los 12 casos que clasificamos como dudosos, 15 sobre los 6 casos donde la respuesta fue exacta, y solamente 1 sobre los dos casos donde la respuesta fue inexacta. Ello proporciona alrededor de 4 variaciones para cada uno de los casos dudosos, 2,5 solamente para las respuestas exactas y 0,5 para las respuestas inexactas. Hay, por tanto, casos donde el juicio es muy incierto, y otros casos donde es más seguro; y donde el examen descubre «algo objetivo». Este «algo objetivo» viene a coincidir con la inteligencia en 6 de los 8 casos; estas cifras son demasiado débiles para que el cálculo de probabilidades permita decir algo sobre el tema; sería ilusorio hacer porcentajes; reemplazando 6 sobre 8 por el 75% se cometería un grave error, porque podría perfectamente ocurrir que el azar puro (es decir, una probabilidad del 50%) ofrezca 6 éxitos sobre 8 experimentos, mientras que 75 éxitos sobre 100 experimentos desvelan, con una certidumbre prácticamente absoluta, una causa.

Por tanto, es a través de un estudio concreto como podríamos tratar de determinar la naturaleza de ese «algo objetivo» que desvela el experimento. El examen de las fotografías sugirió al señor Binet la siguiente hipótesis: cuando los dedos son cortos y feos, se juzga que el niño es torpe; cuando son largos y afilados, se le juzga inteligente. Para verificar esta hipótesis, el señor Binet pidió a dos nuevos experimentadores que hicieran una clasificación basando su juicio únicamente en el criterio precedente. Este experimento figura en el dossier que él me ha remitido y, para no dejarme influir por él, no lo he analizado hasta después de haber redactado todo lo que precede: aporta una confirmación notable, de modo que la hipótesis del señor Binet me parece enteramente confirmada por mi estudio aritmético. En efecto, en los 6 casos en los que la respuesta de la mayoría ha sido exacta, y en los 2 casos donde ha sido inexacta, *todas* las respuestas de estos dos últimos observadores concuerdan exactamente con las de la mayoría[25], y este ejemplo muestra bien cómo este método debería ser aplicado en aquellos casos donde, como éste, no se sabe exactamente *a priori* a qué se refiere la investigación. El método de las mayorías nos enseña que hay «algo objetivo»; la observación directa de los hechos conduce enseguida a formular hipótesis, que deben ser verificadas por

[25] Una sola diferencia insignificante a señalar; para uno de los niños juzgado torpe por la mayoría (respuesta exacta), uno de los dos observadores ha señalado «medio», mientras que el otro señaló «torpe».

nuevos experimentos. El cálculo juega solamente aquí, como siempre ocurre en las ciencias experimentales, un rol de intermediario entre dos experiencias, pero este rol de guía de la experiencia, siendo modesto, es muy a menudo de los más útiles.

ÉMILE BOREL,
Profesor adjunto de la Sorbona

BIBLIOGRAFÍA

BERTRAND, J. (1889): *Calcul des probabilités*. Gauthier-Villars. Paris.
BINET, A. (1905): *Les Révelations de l'écriture soumises à un contrôle scientifique*. E. Flammarion. Paris.
BOREL, E. (1906): «Sur la valeur pratique du calcul des probabilités» *Revue du Mois*. 10 avril; t. I.
— (1906): *Revue du Mois*. Aôut-septembre; t. II.
— (1907): «Un paradoxe économique; le calcul des probabilités des verités statistiques» *Revue du Mois,* 10 décembre; t. IV.
HENRY, C. (1908): «La loi des petits nombres» *Recherches sur le sens de l'écart probable dans les chances simples à la roulette, au trente, quarante, etc., suivies d'une instruction pratique pour le joueur.* Laboratoire d'Energétique d'Ernest Solvay. Paris.
MASCART, J. (1907): «Observations simultanées de la surface de Jupiter» *Bulletin de la Société astronomique de France*.
ROUSSON (1906): *Bulletin de la Société libre pour l'étude psychologique de l'enfant*. Juin, 6.º année.
TITCHENER, E. (1901): *Experimental Psychology. Students' Manual-Quantitative*. The Macmillan Company. New York.