LAS MATEMÁTICAS DE LO INCIERTO

Pedro Gil Álvarez
Catedrático de Estadística e Investigación Operativa

Las Matemáticas son el lenguaje con el que Dios ha escrito el Universo
(GALILEO)

Al recibir el grato encargo de preparar esta lección inaugural, recordé que en tal menester fui precedido hace ya muchos años (en 1913 para ser exactos), en esta misma Universidad, por el padre o padrino científico de los matemáticos españoles e hispanoamericanos del siglo XX: D. Julio Rey Pastor. Acudí presuroso a nuestra querida biblioteca y allí conseguí el discurso escrito por el maestro.

Mi primera tentación tras la lectura precedente fue la de proceder a una discreta retirada de la misión, argumentando algún problema personal, para no tener que verme en la situación de afrontar comparaciones, aunque fuera yo mismo el autor de las críticas.

Más tarde pensé en leer el propio discurso escrito por D. Julio; a fin de cuentas la parte central del mismo se refiere a la historia de la matemática española del siglo XVI y los comentarios previos y posteriores sobre nuestra querida Universidad mantienen, por desgracia, suficiente actualidad.

Finalmente opté por mirar al futuro, y no al pasado, y exponerme al juicio de los oyentes, a los que supongo algo aterrados ante la perspectiva de escuchar una soporífera lección de la que ha sido llamada «reina de las ciencias», aunque dudo mucho que tan laudatoria calificación fuera aceptada por muchos alumnos e incluso algunos profesores de nuestra Institución.

Por su desarrollo histórico las matemáticas son, efectivamente, el edificio mejor acabado de la ciencia; pero es también bastante general la opinión de que moverse en el interior de ese edificio resulta demasiado complejo para el común de los mortales, y la de que sólo unos pocos elegidos lo logran. El origen de tales ideas habría que buscarlo tanto en el carácter mágico que siempre han tenido los números y las figuras geométricas, como en el elevado índice de suspensos (no necesariamente imputables siempre al alumnado) que la materia en cuesción proporciona a muchas personas en su etapa educativa.

Así pues, fui conducido a la búsqueda de un tema que exponer hoy ante Vds. «Pero (de nuevo son palabras de Rey Pastor) desarrollar un tema propio de esta disciplina, con todas sus aparatosas y para los no iniciados casi espeluznantes notaciones simbólicas, y abusar de vuestra desventajosa posición, cuando la ley os obliga a oírme, para desarrollar uno de mis insignificantes trabajos matemáticos, hubiera sido caso inaudito de crueldad, que no podría perdonarme toda vuestra indulgencia».

No obstante, mis limitados conocimientos me impiden hablar de otro tema distinto del que constituye el eje de mi vida científica; carezco de la versatilidad necesaria para encarar otras posibilidades, como puedan hacerlo muchos de mis colegas a los que sanamente envidio.

En consecuencia decido hablar de varios asuntos matemáticos, si bien procuraré que el tratamiento de los mismos sea tan ligero de fórmulas como cargado de ideas.

Es mi deseo procurar, en lo posible, huir del lenguaje críptico, a que tan acostumbrados nos tienen los «especialistas» en cualquier materia (obsérvense, por ejemplo, las explicaciones de algunos economistas, médicos, etc. en los medios de comunicación), probando así que los objetivos de las Matemáticas —en general— y la Estadística y otras ciencias asimiladas —en particular— no constituyen sólo la filosofía de una secta, sino que pueden ser entendidos en mayor o menor grado por el que, en consonancia con esta lección, podríamos llamar «ciudadano medio».

El tema genérico elegido es el estudio matemático de la incertidumbre. La incertidumbre es inherente a la naturaleza: hoy en día nadie pone en cuestión que los movimientos de las partículas, la distribución de los genes y cromosomas, y el comportamiento de los propios individuos necesitan de teorías basadas en el estudio de la incertidumbre más que en leyes deterministas.

¿Qué criterios seguir para tomar decisiones en situaciones de incertidumbre? ¿Cómo generalizar los datos particulares hasta lograr descubrir nuevos fenómenos o crear nuevas teorías? ¿Qué es, como señala el Prof. Rao, el proceso implicado en estas tareas: un arte, una tecnología o una ciencia?

No hubo intentos de responder a estas preguntas hasta comienzos de este siglo, tratando de cuantificar la incertidumbre. En los últimos cincuenta años, si bien no puede decirse que el éxito conseguido sea completo, sí es cierto que los resultados obtenidos han producido una revolución en todas las esferas del saber, han cambiado nuestros hábitos de pensamiento y han permitido descubrimientos notables que nuestros prejuicios acerca del determinismo habían impedido anteriormente.

Ahora, al alcance ya de la mano el nuevo siglo, queremos dar una visión sencilla, que no simple, de estos temas que, 'a la chita callando', se han colado en nuestra vida diaria.

El fin perseguido está claro; cuando concluya la exposición, Vds. tienen la palabra para decir si se ha conseguido.

He bautizado la lección como «Las Matemáticas de lo incierto» y la he desglosado en tres pequeños apartados y un epílogo. Comencemos.

1. LAS MATEMÁTICAS DEL AZAR: EL CÁLCULO DE PROBABILIDADES Y LA ESTADÍSTICA

El azar es, quizá, el seudónimo de Dios cuando no desea firmar.
(ANATOLE FRANCE)

Lo que comenzó en el siglo XVII siendo un 'divertimento' para ociosos amigos del juego —recuérdese al Caballero de Meré que propició la célebre correspondencia entre Pascal y Fermat— o una curiosidad para los políticos con aficiones aritméticas, como las tablas de Graunt o Halley, se ha convertido en la actualidad en el arma más poderosa de las ciencias humanas y de la naturaleza.

Para algunos desaprensivos ha sido incluso un arma de doble filo que ha permitido, con bastante frecuencia, el fraude de una clase de 'pseudodemostraciones' que nos han convencido simultáneamente de la veracidad de cualquier afirmación y de su contraria, hasta ponernos, por ejemplo, el corazón en un puño con la seria amenaza de contraer gravísimas dolencias por el consumo de humildes zanahorias (por otra parte maravillosas para poseer una envidiable vista de lince).

1.1. El uso de la Estadística...

Pero, dejando por el momento las curiosidades, veamos cómo funciona la ciencia estadística.

No voy a intentar dar una definición rigurosa; existen muchas de carácter jocoso, como aquella tan conocida de «la ciencia que nos enseña que si Juan ha comido un pollo y Pedro no ha comido nada, han comido medio pollo cada uno», o esa otra de «la ciencia que nos señala como situación ideal para la temperatura corporal tener la cabeza en el horno y los pies en la nevera», ambas coincidentes en la crítica hacia el uso desmedido del valor medio. Y existen otras muchas más serias entre las que sería muy difícil elegir.

Citaré, pese a todo, la definición dada por un gran metamático de nuestro siglo, el Prof. Fréchet, que señala a la Estadística como «una ciencia que se ocupa de resolver problemas que, con rigor matemático, carecen de solución». En mi opinión en esta frase se encierran simultáneamente la grandeza y las limitaciones de esta ciencia: grandeza como soporte de otras ciencias a las que otras ramas de las matemáticas no pueden ayudar; limitaciones inherentes a todos los resultados estadísticos, que se encuentran en la propia naturaleza de la ciencia estadística, considerada como estudio de las situaciones en que hay intervención del azar.

Pero, si establecer una definición no es tarea sencilla, sí resulta fácil, señalar las tres partes esenciales de que consta, a las que pasaremos una breve revista:

La estadística descriptiva se ocupa de la presentación de la información disponible de forma ordenada; existía ya cuando aún la palabra 'estadística' hacía referencia solamente a los trabajos de «estado» relativos a censos, nacimientos o defunciones.

El cálculo de probabilidades, que es el modelo matemático (y, como tal, utópico) para representar esos datos; constituye la parte más refinada en cuanto al rigor, y permite que la estadística sea considerada como una parte de las matemáticas desde hace unos sesenta años (¡mucho menos tiempo que la, hasta hace poco, denominada «matemática moderna»!). En esas fechas, en España, la Estadística sólo formaba parte de los planes de estudio de las Facultades de Leyes, quizá las únicas que ahora no la incluyen en los mismos.

La inferencia se nutre de las dos partes anteriores y es la que, en rigor, merece el nombre de ciencia estadística propiamente dicha. Su objetivo es averiguar, a veces casi 'adivinar', algo (sea poco o mucho) sobre esas leyes del azar que influyen en un fenómeno y que, o bien se conocen sólo en parte, o bien son totalmente desconocidas.

Todas sus conclusiones están sujetas a limitaciones de carácter probabilístico; nada de lo que se afirma o se niega es totalmente seguro, sólo se sabe que ocurre con una cierta probabilidad. Dicho de otro modo: se pueden dar ciertas garantías de que las cosas van a ocurrir tal como indica la predicción, pero siempre existe el riesgo (a veces muy fuerte, si no se toman las precauciones debidas) de que ocurra lo contrario o, al menos, de que no suceda exactamente lo predicho.

Tal como se ha indicado, el objetivo primordial que persigue la Estadística es el de hacer averiguaciones, habitualmente sobre una característica de interés en la 'población' que es objeto de estudio.

Si existe la posibilidad de observar tal característica en todos los 'individuos' (nótese que los nombres empleados indican los orígenes de la Estadística como ciencia del estado), la averiguación carecería de sentido, se obtendrían resultados indiscutibles; tal es el caso de los censos o de un recuento de votos emitidos.

Si, por el contrario, como ocurre con mucha frecuencia, la observación de toda la población no es posible por su coste económico o temporal, por tratarse de un proceso destructivo, etc., el investigador debe conformarse con el estudio de una 'muestra', que es una parte de la población y, a través de ella, inferir los valores poblacionales. De nada serviría, por ejemplo, estrellar todos los vehículos producidos para averiguar su resistencia a los impactos frontales y laterales; no quedarían automóviles para la venta.

En principio parece razonable pensar que cuanto mayor sea el tamaño de la muestra (número de individuos que la componen), más acertadas serán las averiguaciones que llevemos a cabo; y así puede demostrarse aunque, a veces, un incremento desmesurado del tamaño de muestra no lleve aparejado un incremento sustancial de las garantías conseguidas, y sí del gasto correspondiente, como sucede en alguna de nuestras encuestas oficiales más conocidas y prestigiadas. Debe, pues, buscarse el tamaño de muestra óptimo en algún sentido.

Es imprescindible que el investigador mantenga el rigor matemático de los resultados teóricos que obtiene y, si en algún paso se efectúan aproximaciones, éstas deberán indicarse para ser añadidas a la lista de limitaciones de la conclusión.

Por último, la ética profesional debe caracterizar la exposición de conclusiones, señalando claramente, como ya se ha indicado, que ninguna de ellas constituye en sí misma una verdad absoluta sino —todo lo más— que tiene una alta probabilidad de ser cierta.

Sería infantil pretender dar un ejemplo de cada una de las ciencias que se han beneficiado del desarrollo de los métodos estadísticos; renuncio expresamente a ello en beneficio del tiempo disponible para criticar sus usos incorrectos.

1.2. ...y el abuso de la Estadística

Y hasta aquí las indicaciones a seguir para un uso correcto de la Estadística; desgraciadamente los «abusos» cometidos en su nombre tienen una gran transcendencia sobre los receptores de la información, nosotros mismos. ¿Quién no ha escuchado o leído la frase «hay mentiras, grandes mentiras y estadísticas»? ¿Por qué existen libros con títulos como «mentir con estadísticas»? ¿Por qué los mismos datos sirven para que alguien nos convenza de algo y, acto seguido, otra persona justifique lo contrario?

Con frecuencia la respuesta a estas preguntas se puede encontrar en la falta de cultura estadística de los que reciben la información; soy un ardiente defensor de que tal cultura sea obligatoria en nuestro sistema educativo porque, en palabras de mi maestro, el Prof. Sixto Ríos, «en la actualidad las ideas estadísticas más sencillas deben ser parte indispensable del equipo mental del hombre educado», y con ello las posibilidades de engaño serían menores.

Otras veces, no pocas, es la impericia y la falta de conocimientos de los responsables de la elaboración de estudios estadísticos; no en vano es la Estadística una de las especialidades con mayor índice de intrusismo.

Finalmente, a nadie se le escapa que en este terreno resbaladizo de las aproximaciones y la falta de garantías absolutas, es fácil la manipulación (incluso legítima, sin fraude) de los resultados. Un precursor de estas manipulaciones fue el propio Mendel, de quien se ha podido probar que, muy posiblemente, falseó los resultados obtenidos en sus experiencias con guisantes, para dar mayor fuerza a sus famosas leyes.

Veamos otros ejemplos:

Si a cualquiera de los aquí presentes le dicen que «después de recorrer 100 Km. a 75 Km./h. y otros 100 Km. a 125 Km./h. la velocidad media ha sido de 100 Km./h.» estaría, si no lo piensa un poco, dispuesto a admitirlo como cierto; un cálculo sencillo muestra, sin embargo, que tal velocidad media no llega a los 94 Km./h. ¿Qué ha ocurrido? Sencillamente que nos hemos dejado engañar con el uso de un promedio.

El propio Bernard Shaw decía, irónicamente, que «el uso del paraguas aumenta el perímetro torácico, prolonga la vida y confiere inmunidad para las enfermedades, pues se puede probar con estadísticas que las personas que usan paraguas son más gruesas y saludables y viven más tiempo que las otras». No hacen falta muchos conocimientos estadísticos para comprender que la causa de esta diferencia no es el paraguas (artículo de lujo en la época) sino la calidad de vida de quienes lo tenían.

Afirmaciones similares a la burlesca de Shaw son moneda corriente hoy día; una simple ojeada a la prensa nos revela todas las «demostraciones estadísticas» de lo nocivo que resulta para nuestra salud el consumo de tal o cual alimento (o de lo maravilloso que resulta otro de análogo composición), de lo perjudicial que resulta el vino para el corazón (o de lo excelente que es beber en las comidas para elevar el tono vital), de lo malísimo que es el humo de nuestro pitillo para el que pasa a nuestro lado (o de la influencia beneficiosa de unos cigarrillos para combatir la demencia senil); y así podríamos seguir casi indefinidamente.

¿Qué secreto se esconde en estas contradicciones? ¿Son los famosos duendes de la imprenta (hoy también en las ondas) los que juegan estas malas pasadas a nuestro entendimiento? La respuesta es, desde luego, no. La falacia en la «demostración» se produce, como se han señalado, en dos puntos esenciales: hay ocultación de las limitaciones de los resultados por parte del informador (el original, no el periodista) y, aunque no la hubiera, necesitaríamos una cultura estadística más amplia para dar a la información recibida su auténtico valor.

1.3. Los sondeos y algo más

¿Y los sondeos de opinión? Si hay algún fracaso sonoro de la estadística que se refleje en nuestra vida cotidiana, ese fracaso está en los dichosos sondeos. Alguno pensará que, hábilmente, estoy escurriendo el bulto, intentando pasar de puntillas sobre un tema que es objeto de polémica cada vez que se convocan unas elecciones (algo que, por otra parte, es demasiado frecuente desde que unos aprendimos a hacerlo y los mayores pudieron recordar cómo se hacía). Pues no; no es esa mi intención y por ello procuraré arrojar luz sobre algunos aspectos de los sondeos que no siempre son bien entendidos si bien, aunque siga intentando eludirlas, precisaré de alguna fórmula.

¿Por qué los resultados no son coincidentes en todos los sondeos? ¿Con qué criterios juegan los diseñadores del cuestionario para decidir el tamaño de la muestra? ¿Cómo debe hacerse la lectura objetiva de los resultados presentados?

Trataremos de ver solamente los elementos esenciales suponiendo, en todo caso, que los procedimientos de selección de la muestra han sido correctos.

El objetivo perseguido inicialmente es localizar el porcentaje de la población que se inclina por una de las opciones que se le presentan; llamaremos P a ese porcentaje. Este valor P es el parámetro poblacional que se desconoce (de hecho no se conocerá hasta efectuar una encuesta exhaustiva, esto es hasta el día siguiente al de las elecciones en el caso de un sondeo sobre el voto a los diferentes partidos) y que debe estimarse a través de la muestra; se elige para tal fin el valor del porcentaje correspondiente a las respuestas de los individuos encuestados; denotaremos este valor P^*.

Se puede (con los recursos matemáticos disponibles en el Cálculo de Probabilidades), dar un intervalo en el que se encuentra el porcentaje P desconocido un tanto por ciento de veces tan alto como se quiera. Ahora ya todo parece bajo control, excepto por un pequeño detalle: ¿qué error estamos dispuestos a admitir? O, dicho de otro modo: ¿qué amplitud deseamos que tenga el intervalo?

Pues aún aparece otro valor controlable, el tamaño de la muestra —que debe determinarse antes de proceder a realizar la encuesta— y que depende de ese error que se considera admisible.

En general, tal error es la mitad de la longitud del intervalo construido (un porcentaje pequeño), la distancia de P^* (que es el punto central) a cada extremo; si lo representamos por E se cumple en las condiciones más comunes que el tamaño de muestra necesario es

$$n = \frac{10.000}{E^2}$$

Puede comprobarse que para los valores E=2%, E=1% y E=0,5% aparecen los números «mágicos» de los tamaños de muestra en las encuestas y «macroencuestas»: 2.500, 10.000 y 40.000 respectivamente. Así, por ejemplo, un valor estimado P* =30% permite decir que el valor poblacional P está en los intervalos (28, 32), (29, 31) o (29.5, 30.5) (o, con el lenguaje habitual de las empresas encuestadoras, (30± 2%, 30± 1%, 30± 0,5%) con la garantía del 95,5%, lo que significa (¡atención!) que si fuéramos capaces de construir muchos de estos intervalos con muchas muestras diferentes del mismo tamaño, el auténtico valor de P se encontraría en el 95,5% de ellos.

Y, como en todo procedimiento estadístico, las limitaciones de lo conseguido: sea cual sea el margen de error que estemos dispuestos a admitir, un 4,5% de los intervalos que podríamos construir no contienen el valor verdadero de P. Este 4,5%, ¿es poco o es mucho? La respuesta es particular de cada solicitante de un estudio, depende de los riesgos que esté dispuesto a correr cada uno; personalmente pienso que no me arriesgaría a encargar un estudio en el que una de cada veinte veces (más o menos) no se acierta (¿y si me toca una de esas veces?). Nótese que en caso de no acertar no sólo es incorrecto el valor de P estimado por P* sino cualquier valor de intervalo.

Desde luego, de lo anterior no se deduce en absoluto que los resultados de los sondeos electorales sean extrapolables a la composición del parlamento, en parte debido a los márgenes de error (de las 2.500 encuestas necesarias para un error del 2%, a nuestra querida Asturias, considerada como provincia media, le corresponderán unas 50, valor que, suponiendo que todos los procesos de selección de la muestra y demás son correctos, se corresponde con un error de casi un 15%), y en parte a nuestra ley electoral; en este sentido sólo tiene validez un recuento similar al que se efectúa con las primeras papeletas en mesas seleccionadas de cada circunscripción electoral (lo que constituye el primer muestreo de la auténtica población, la de los votos emitidos).

Además, en cada sondeo, existe un porcentaje no despreciable de contestaciones «no sabe-no contesta» cuya clasificación no es sencilla por depender de muchos factores individuales (no sólo de la fidelidad de voto). Y, aunque existen técnicas de estimación indirecta, no parece que sean utilizadas por nuestras empresas de encuestas.

Lo que acabamos de ver es aplicable a los estudios de audiencia de las cadenas de radio y televisión, gracias a los cuales tenemos los programas que tenemos; las limitaciones son idénticas (o aún peores, pues hasta hace muy poco sólo se controlaban, por ejemplo, 2.000 hogares, y ni siquiera estaban elegidos al azar pues habían prestado su consentimiento para tener conectado el aparatito correspondiente). Menos mal, y esto es aplicable al caso de las encuestas, que los procedimientos de muestreo ayudan a mejorar los resultados (aunque a veces sea inconscientemente) por parte de las empresas dedicadas a este tipo de estudios. Sin ánimo de ofensa, me gustaría señalar que los responsables de las mismas no son expertos en Estadística, sino profesionales de otros campos como la Sociología, Ciencia Política, Economía, etc.

Dejemos las encuestas y veamos unos últimos equívocos que se producen ante la estadística: la generación constante de resultados aleatorios conduce a la aparición de cualquier resultado posible; así, la sorpresa inicial ante una combinación como 1-2-3-4-5-6 aparecida en un sorteo de la lotería primitiva, no debe inducirnos a creer que tal secuencia tiene menor probabilidad de aparición que cualquier otra.

Es una situación semejante a la del mono que, mecanografiando continuamente, podría llegar a escribir las obras completas de Shakespeare en un tiempo muy largo aunque finito (se calcula que la probabilidad de reproducir Hamlet con sus 27.000 letras y espacios es aproximadamente la unidad dividida entre una cantidad que tiene 41.600 ceros, lo que da idea de la magnitud del tiempo que cabe esperar para obtener el resultado apetecido).

En otro orden de equívocos se sitúa la conocida falacia del jugador que considera que su probabilidad de acierto aumenta con el número de fracasos precedentes; sobre tal error señala Polya la anécdota del médico que da ánimos a su paciente del siguiente modo: «Su enfermedad es muy grave; sólo uno de cada diez enfermos sobrevive. Pero no se preocupe; ha tenido suerte de acudir a mí: recientemente he tenido nueve pacientes con esta enfermedad y todos han muerto».

Como contrapartida de lo anterior, un ejemplo más: durante mucho tiempo se ha creído en la conocida como teoría de la acumulación según la cual un mismo suceso tiende a repetirse en períodos de tiempo breves; así lo ponen de manifiesto nuestros refranes como «las desgracias nunca vienen solas» o el más castizo «cuando la semana está de piojos, no vale mudar la camisa».

Y así, ad infinitum. Tal vez por continuar creyendo que los fenómenos son siempre determinísticos, creencia mantenida en vida por un científico tan notable como el propio Einstein. En cualquier caso, el genial físico aceptó la teoría de Bose sobre el comportamiento aleatorio de las moléculas, convencido quizá por la llamada ley de los grandes números que 'introduce el orden dentro del desorden'.

2. Las matemáticas de la comunicación: teoría de la información

> Aparecieron, como divididas, lenguas de fuego, que se posaron sobre cada uno de ellos, quedando todos llenos del Espíritu Santo; y comenzaron a hablar en lenguas extrañas, a medida que el Espíritu les otorgaba expresarse.
> (Hechos de los Apóstoles, 2, 3-4)

La información es poder. Se dice que será el gran poder del próximo siglo y, a juzgar por la situación actual, tal afirmación será un hecho.

Pero, ¿se puede medir la cantidad de información? ¿Puede darse un valor a la información contenida en una noticia, a la información dada por una foto, a la que circula por una línea telefónica o en el interior de un ordenador? ¿Es verdad que, como nos enseña el refranero, una imagen vale más que mil palabras?

La respuesta fue dada simultáneamente por dos de los grandes científicos de nuestro siglo: Norbert Wiener y Claude Shannon. No es la primera vez que en la historia de las matemáticas sucede algo parecido (recuerdo entre mis lecturas de infancia la de la 'carrera' entre Newton y Leibnitz por llegar el primero a 'patentar' el cálculo diferencial), pero tal respuesta tiene pequeños matices que diferencian la posición de ambos investigadores:

Wiener piensa en la información que recibe una vez que se produce la experiencia; mide esa información a posteriori; Shannon, por el contrario, observa la situación a priori y mide la información que espera obtener de ella; por eso la entropía de este último es el valor medio de las informaciones del primero.

Hay otros muchos conceptos de información desarrollados antes que los anteriores, y otros que se desarrollan a partir de la década de los 50: todos tienen propiedades comunes. ¿Casualidad?. ¿O es algo más?. Como diría Philip Jourdain puede tratarse de un mismo Concepto Matemático (con mayúsculas) que se expresa a través de desarrollos matemáticos (con minúsculas) diferentes.

Cuando Hartley en 1928 señala el logaritmo del número de resultados como medida de la incertidumbre ante una experiencia desconoce que está suponiendo que tales resultados son equiprobables y, para evitar conflictos con las situaciones en que tal equiprobabilidad no es cierta, argumenta que las probabilidades son algo que tiene mucho que ver con las motivaciones psicológicas (¡quizá pensaba ya en una probabilidad subjetiva!) y que, por tanto, el problema debe ser estudiado por los psicólogos y no por los matemáticos o ingenieros.

Se equivoca al adoptar esta estrategia y eludir el problema.

Pero acierta en el camino a seguir: la información debe medirse como una disminución de la incertidumbre y, en consecuencia, debe comenzarse por estudiar ésta.

Y con la idea de Hartley se desarrolla la Teoría de la Información por Shannon, quien define la entropía para un sistema probabilístico cualquiera como medida de la incertidumbre asociada (ligando así la incertidumbre con la probabilidad), inspirada en la cantidad del mismo nombre definida en termodinámica por Boltzmann. Es la célebre fórmula:

$$H = -\sum p_i \log p_i$$

Define también la información entre dos sistemas como la disminución media de la incertidumbre que se produce en uno de ellos por el conocimiento del resultado del otro. De aquí a los canales de información, su capacidad, las fuentes de información, el estudio de los ruidos, etc., en un desarrollo vertiginoso en el que participan numerosos científicos.

Todo ello partiendo de un esquema sencillo que hoy se conoce en todas las ciencias de la comunicación: la secuencia emisor, codificador, canal, descodificador y receptor.

Todo ello sobre la idea del efecto sorpresa: cuanto menos probable es un suceso más información proporciona cuando tiene lugar (el famoso 'hombre muerde a perro' del periodismo). Y, además, con el fin de transmitir mensajes cuyo contenido se considera sólo en su valor sintáctico, como significantes sin significado.

Y todo ello funcionando a las mil maravillas y logrando que la compañía telefónica Bell (la empresa en la que trabajaba Shannon) consiga su objetivo: ofrecer a sus clientes una mayor fiabilidad en la transmisión de las conversaciones, procurando que los ruidos provoquen interferencias mínimas.

¿Todo resuelto? No, aquello no había hecho más que empezar; al menos los dos siguientes problemas estaban sin resolver.

2.1. *Codificación y Criptología*

¿Cómo reproducir en el receptor los mensajes originales cuando, a su paso por el canal, se produce una distorsión de los mismos? ¿Cómo compartir los recursos de almacenamiento y transmisión respetando la privacidad de la información?

Es cierto que siempre se puede establecer un sistema de repetición de mensajes entre emisor y receptor tantas veces como sea necesario hasta tener garantías de ausencia de error; o bien se puede codificar el mensaje de modo tan especial que ni uno mismo sepa lo que escribió (como en ciertos exámenes que todos los profesores hemos corregido en alguna ocasión). Pero no es menos cierto que ni nuestros bolsillos soportarían las facturas del teléfono ni nuestras mentes el esfuerzo de la descodificación.

Fue el propio Shannon el primero en demostrar la existencia de un sistema de codificación-descodificación que asegura que la probabilidad de error sea tan pequeña como se quiera, pero se trata de una demostración de esas a que los matemáticos somos tan aficionados y que, como señalaba Bertrand Rusell, forma parte del «campo en el que no sabemos bien de qué estamos hablando ni si lo que decimos es verdad».

No se ha podido construir un sistema real que responda al teorema de Shannon pero, en el empeño, se ha desarrollado una poderosa teoría de la codificación, de naturaleza fundamentalmente algebraica, que ha proporcionado soluciones satisfactorias al problema inicial: todos los sistemas digitales de registro, conservación y transmisión de la información que empleamos en la actualidad (ordenadores, discos compactos, teléfonos móviles de última generación, etc.) son una buena muestra de la potencia de la herramienta conseguida. ¡Y todo ello empleando solamente los dígitos 0 y 1!

En cuanto a la segunda pregunta, de máxima actualidad, la respuesta está siendo positiva (en presente activo, en términos de gramática inglesa): cada día es más importante la protección de los datos para impedir que alguien pueda utilizarlos de forma fraudulenta o, simplemente, para otros fines distintos de los pensados, cada día es más importante que muchos datos estén protegidos de los piratas informáticos que pueden provocar desde nuestra pequeña pero respetable ruina (con las claves de trabajo de nuestras tarjetas de crédito) hasta una tensión mundial de imprevisibles consecuencias (introduciéndose en los sistemas de defensa de las grandes potencias).

Surge así, como prolongación natural de la teoría de la codificación, una nueva ciencia denominada criptología, con una rama de cifrado (la criptografía) y otra de descifrado (el criptoanálisis), cuyos orígenes son tan antiguos como la humanidad: recuérdense la escritura sobre la tira enrollada al bastón o 'escítalo' de los lacedemonios o el cifrado de César por desplazamiento de magnitud constante sobre las letras del alfabeto.

2.2. El huevo y la gallina: ¿Información o Probabilidad?

Pasemos a otro punto: ¿sin probabilidad no hay información?. ¿Cuál de los dos conceptos es el primario?. ¿Cómo tener en cuenta el valor semántico de la información transmitida? A estas preguntas da respuesta la teoría axiomática de la información, creada por los Profs. Kampé de Fériet y Forte hace unos 30 años.

En las situaciones prácticas en que la teoría desarrollada por Wiener y Shannon ha sido aplicada con éxito, las probabilidades empleadas se originan a partir de frecuencias correspondientes a la repetición de una experiencia un número suficientemente grande de veces.

Por supuesto es válido emplear las expresiones de incertidumbre e información dadas para probabilidades a priori, pero tales valores de probabilidad no pueden ser determinados si no es en función de toda la información de que se disponga sobre el modo en que se produce un resultado; así, por ejemplo, si se conociera que la posición del centro de gravedad de un dado no corresponde a su centro geométrico, no debería asignarse la misma probabilidad a las seis caras.

Con frecuencia, sin embargo, las situaciones en las que, sin duda, existe una información, no permiten la consideración de una probabilidad; en un ejemplo clásico de este tipo aparece la figura del Cardenal Roncalli, cuya elección como Papa Juan XXIII mediante un sólo cónclave impediría cualquier consideración frecuentista. Además, en una consulta previa no debería darse el mismo valor a la opinión de un taxista, por ejemplo, que a la de un alto cargo del Vaticano y a pesar de todo, evidentemente, antes de la elección existía «incertidumbre» y, después de conocer el resultado, una ''información''».

Otro ejemplo: al tratar de aproximar el número π, las sucesivas acotaciones que puedan obtenerse ($3.1 < \pi < 3.2$, $3.14 < \pi < 3.15$, etc.) aumentan la información recibida, que podría medirse, salvo algún factor de escala, por medio del tiempo necesario para lograr la precisión deseada. Esta medida, perfectamente válida, también se sale del marco de la teoría de Shannon y carece de sentido hablar de la probabilidad de que el número π se encuentre entre ciertos valores.

Una última situación, en la que queda clara la «subjetividad» del valor de la información, es la correspondiente a la celebración de un sorteo de lotería con un millón de números en el que el premio recae sobre, pongamos por caso, el número 123456. De acuerdo con el modelo probabilístico la información recibida es el logaritmo del número de resultados posibles. Sin embargo, al leer el resultado, tres lectores distintos pueden recibir una información muy diferente: el lector A, que no juega en ese sorteo, recibe una información nula; el lector B que tiene un décimo no premiado recibe una información que podría calcularse, por ejemplo, por medio del gasto efectuado al comprarlo; finalmente un tercer lector C, que posee un décimo premiado, recibe una información muy grande. La probabilidad de aparición del 123456 es, evidentemente, la misma en todos los casos.

Como ya se ha señalado, la entropía y, en consecuencia, la cantidad de información asociada, no tienen en cuenta el valor semántico de los resultados, sino sólo su valor sintáctico, prescindiendo de la «importancia» del mensaje a transmitir, lo que ha conducido a ciertos autores a considerar medidas en las que la probabilidad no lo es todo, aunque sea tenida en consideración. Los mensajes a transmitir quizá tengan el mismo valor para los operarios o las máquinas que codifican y descodifican, pero no para el emisor y receptor humanos a los que, en última instancia, afectan dichos mensajes.

De cualquier modo existe una comunicación entre informador y receptor, plasmada en una proposición explícita o implícita, entendida la proposición en el sentido de la lógica clásica. Será a dicha proposición a la que asignaremos una medida de la cantidad de información que transmite.

Con los conjuntos de objetos y propiedades puede construirse la familia $\mathcal{P}$ de las proposiciones elementales formada por las iniciales, sus respectivas negaciones y todas las proposiciones compuestas deducidas de las elementales por medio de las operaciones clásicas de la lógica distributiva, construyendo un retículo de proposiciones $\mathcal{F}$, que junto con $\mathcal{P}$ proporciona la estructura adecuada para definir una medida de información.

En su formulación conjuntista, la teoría axiomática de la información parte de un espacio medible como el que permite construir la probabilidad; pero, hay ciertos aspectos de interés, que obligan a contemplar de algún modo la estructura de partida.

El último elemento, necesario para construir la teoría axiomática de la información, es el concepto de independencia, cuyo análisis vamos a realizar en el espacio de las proposiciones:

Dos proposiciones, P y q, son lógicamente independientes cuando cada una no implica ni excluye la otra. Pero quizá esta condición no baste; por ejemplo: la proposición «X es fumador» no implica ni excluye la proposición «X tiene cáncer de pulmón» y viceversa; sin embargo, ¿quién no se resiste a considerar ambas proposiciones como independientes? Pues al menos la primera permite hacer previsiones sobre la segunda y por tanto proporciona información sobre ella. Así, pues, en lo relacionado con la información debemos distinguir dos niveles de independencia:

a) El nivel lógico o sintáctico, que queda reflejado en la condición de independencia dada, basada en las operaciones lógicas habituales.

b) Un segundo nivel, semántico, más exigente que el anterior, que juzga la independencia de las proposiciones a través de su significado y que rebasa el ámbito de la estructura de información en que nos movemos.

La terna descrita se denomina estructura de información medible. Su estudio abre nuevos horizontes para el estudio de modelos, más generales que los probabilísticos, en los que o bien no existe la probabilidad o no se encuentra sola. Quizá, como se ha señalado por algunos estudiosos del tema, sea ésta la ciencia de conexión que pueda vertebrar las ya muy dispersas actividades científicas del tiempo que nos ha tocado vivir.

3. LAS MATEMÁTICAS DE LA IMPRECISIÓN: CONJUNTOS BORROSOS

Tanto la precisión como la certidumbre son ideales falsos. Son imposibles de alcanzar y... debería renunciarse a ambas. No se debería nunca intentar ser más preciso de lo que requiere la situación del problema.
(KARL POPPER)

Hace treinta años aparecía publicado el trabajo Fuzzy Sets con el que el Prof. Zadeh marcaba un hito en el desarrollo de las matemáticas aplicadas al estudio de los problemas del hombre.

Hace ya treinta años y, como señalan Dubois y Prade, «la llamada del Prof. Zadeh a la búsqueda de técnicas no convencionales ha sido mal entendida, e interpretada como un ejemplo de pensamiento permisivo que tratara de escapar al rigor de las matemáticas»; de hecho son muchos los científicos que no quieren creer en las ventajas del modelo para adaptarse al estudio de los fenómenos que nos rodean. ¿Por qué?

En primer lugar porque es de una gran osadía enfrentarse a la todopoderosa «matemática moderna». La base de la teoría conjuntista está en garantizar que todo elemento o bien pertenece o bien no pertenece a cualquier conjunto dado. O a la lógica clásica que garantiza que una proposición es verdadera o es falsa. Siempre el «o sí o sopas» de nuestro refranero.

Pongamos algún ejemplo más: en matemáticas una función es o bien continua o discontinua; no puede ser continua en cierto grado. De forma análoga, una matriz es simétrica o no; no puede ser algo simétrica, más o menos simétrica, o simétrica en cierto grado. Del mismo modo, un trabajo publicado en una revista matemática se espera que contenga definiciones, axiomas y teoremas establecidos con precisión. Generalmente, un trabajo no se consideraría aceptable para publicación, si sus conclusiones se establecieran como afirmaciones que no fueran inequívocamente ciertas.

En claro contraste con el mundo idealizado de las matemáticas puras, nuestra percepción del mundo real está invadida por conceptos que no tienen fronteras nítidamente definidas, como por ejemplo, alto, gordo, muchos, la mayoría, lentamente, viejo, familiar, relevante, mucho mayor que, amable, etc.

Y éste es precisamente el quid de la cuestión en palabras del propio Prof. Zadeh:

«Los elementos clave en el pensamiento humano no son números, sino etiquetas de conjuntos borrosos[1], esto es, de clases de objetos en los que la transición de la pertenencia a la no-pertenencia es gradual en lugar de ser brusca (...) Está claro que 'la clase de todos los números reales mucho mayores que uno', o 'la clase de las mujeres guapas' (sorprende este ejemplo en un hombre tan galante como el Prof. Zadeh para quien todas las mujeres son guapas), o 'la clase de los hombres altos' no constituyen conjuntos en el sentido matemático usual del término (...) aunque juegan un papel

(1) La palabra inglesa 'fuzzy' debería quizá haberse traducido por 'difuminado' y no 'difuso', ya que los verbos difuminar y difundir tienen significados muy distintos. Los problemas lingüísticos han surgido cuando se han tratado de buscar las palabras derivadas.

importante en el pensamiento humano, particularmente en los dominios del reconocimiento de formas, la comunicación de información y la abstracción».

El problema estaba planteado. ¿Cuál fue la solución?

Es frecuente que los grandes avances científicos sean de cuasiimposible comprensión para un lego en la materia (e incluso para muchos no suficientemente expertos). Sorprendentemente este no es el caso; lo expresaremos de nuevo con las propias palabras del Prof. Zadeh:

«Un conjunto difuso (Fuzzy Set) en un referencial está caracterizado por una función de pertenencia que asocia a cada elemento del referencial un número entre cero y uno, su grado de pertenencia. Así, cuanto más próximo a uno sea el valor de la función en un punto mayor será el grado de pertenencia del elemento al conjunto. Cuando el conjunto en cuestión es un conjunto ordinario su función de pertenencia puede tomar sólo dos valores, uno o cero, según que el elemento considerado pertenezca o no al conjunto».

¿Tan fácil? Pues sí; se resolvió el problema: hay hombres claramente altos (grado de pertenencia uno o próximo a uno), otros no tanto (grado de pertenencia intermedio entre cero y uno) y otros claramente no altos (grado de pertenencia cero o próximo a cero).

Además, la ventaja de la formulación en términos conjuntistas sobre la formulación lógica, está en el inmenso campo que la matemática ya establecida ofrece a la nueva teoría. Como ejemplo, en el marco de la teoría de conjuntos borrosos es natural plantearse la existencia de números borrosos (tan útiles para dar valoraciones imprecisas, como por ejemplo las calificaciones de los exámenes, ¿cuál es la frontera entre el aprobado y el notable?), de funciones borrosas (excelentes para dictar algunas órdenes, como 'la calefacción debe ser fuerte si el día es frío', ¿o es mejor decir que si el termómetro exterior señala 2 grados sobre cero, el de la caldera debe ponerse a 73 grados, si señala 3 debe ponerse a 71, etc., indicando además que la escala es la de los grados centígrados?), etc.

Algunas críticas feroces a los conjuntos borrosos han sido hechas por investigadores de gran prestigio. Fíjense en los comentarios que el Prof. Kalman, pionero de la Teoría de Sistemas, hizo en 1972:

«Sin duda, el entusiasmo del Profesor Zadeh por la borrosidad se ha visto favorecido por el clima político que impera en los Estados Unidos, de una permisividad sin precedentes. La 'borrosización' es un tipo de permisividad científica; tiende a dar lugar a slogans socialmente atractivos, que no van acompañados por la disciplina del trabajo científico sólido y la observación paciente.

Dejenme decir, de forma bastante categórica que, en mi opinión, no existe algo como un concepto científico borroso».

Por otra parte, además de las críticas de los matemáticos que pudiéramos llamar 'puros' está el ataque de los estadísticos, particularmente los bayesianos, que han intentado considerar la nueva herramienta de trabajo en situaciones de incertidumbre como un caso particular de sus propios métodos. Tal parece que se hubieran armado hasta los dientes para defender su plaza fuerte, la de las decisiones en ambiente incierto, de la que se consideran legítimos y únicos moradores, actuando la mayor parte de las veces con un gran desconocimiento del modelo.

No son pocos los que han pretendido interpretar los grados de pertenencia, convenientemente normalizados, como probabilidades, ignorando que en muchas de las situaciones de ambiente borroso carece de sentido hablar de probabilidad: no hay probabilidad de que Fulano sea alto, o de que hoy haga un buen día; se describen situaciones que dependen de la vaguedad de los conceptos empleados (alto, bueno) y no del azar.

En cuanto al 'enfrentamiento' con los métodos estadísticos y como detalle anecdótico, les contaré que, hace no muchos años, fuí invitado a formar parte de un tribunal de tesis doctoral en el que se presentaba un modelo de diagnosis médica apoyado en cálculos probabilísticos. En el transcurso de la defensa se rechazaban los modelos borrosos con argumentos no demasiado convincentes. ¡Cuál no sería mi sorpresa cuando comprobé que las variables de entrada tomaban «valores» que respondían a criterios tan borrosos como «mucha» tos, «poca» fiebre, «alto índice de colesterol», etc.!

La confrontación carece de sentido y, como muy bien indicaba el título de la conferencia impartida el pasado curso por Zadeh en esta Universidad, «la teoría de la probabilidad y la lógica difusa son complementarias y no adversarias». Además, como ya se ha señalado con respecto a otros aspectos matemáticos, el concepto de probabilidad y las técnicas estadísticas habituales pueden también difuminarse.

Hoy día, como si se tratara de un episodio de la mitología, las criaturas casi han devorado a su creador: no es difícil encontrar quienes preguntan por «los conjuntos de Fuzzy», con el mismo lenguaje que utilizan para el «teorema de Pitágoras», los «espacios de Hilbert» o «la distribución de Gauss».

3.1. La tecnología fuzzy y...

Hoy por hoy las aplicaciones de la lógica borrosa son demasiado visibles como para ignorarlas; ha nacido una nueva tecnología desarrollada fundamentalmente en Japón. Los controladores borrosos son simples y robustos; pero, quizás más importante, el control borroso permite la ejecución de tareas tales como aparcar un coche, que no se prestan a resolución mediante métodos convencionales: los controladores de enfoque de las videocámaras de Sanyo o Panasonic, el controlador de frenos ABS de Nissan, el del acondicionador de aire de Mitsubishi, los sistemas de reconocimiento de caracteres manuscritos de Sony y del habla de Hitachi, el sistema de control del metro de la ciudad de Sendai, controladores de microondas, secadoras, ascensores, y un largo etcétera que proporciona prestaciones mejores y más económicas a los usuarios, consiguiendo en definitiva colaborar a la consecución de una mejor calidad de vida y de una mayor libertad del género humano. Lo 'borroso' está bien visto por los usuarios, es una garantía de buen funcionamiento; lo borroso, en términos de estudios de mercado, 'vende'.

No obstante, aún son demasiados los científicos que siguen sin convencerse de que la lógica borrosa tiene algo importante que ofrecer. El ya citado respeto por lo que es cuantitativo y preciso, y el desdén por lo cualitativo e impreciso, está demasiado profundamente arraigado en la sociedad como para desprenderse de él sin oponer resistencia.

Analicemos brevemente algunos detalles recientes sobre la borrosidad: Zadeh emplea el término granulación para referirse al proceso de formar clases borrosas de objetos, que están agrupados por similaridad. Cuando el número de clases diferentes que deben manejarse es demasiado elevado, tales clases deben agruparse para formar gránulos. Es lo que hacemos para, por ejemplo, determinar los colores: la gama de verdes es tan amplia que 'el color verde' pasa a ser una agrupación, desde luego nada nítida, de las distintas longitudes de onda que se corresponden con diferentes tonalidades de verde.

La necesidad de la granulación obedece, pues, a la capacidad limitada de las personas para almacenar detalles. Desde este punto de vista, la borrosidad y la granulación son consecuencias de la complejidad, y desempeñan un papel clave en la tolerancia de la imprecisión para logar eficiencia, robustez y bajo coste en el producto final.

Una implicación importante de esta observación, como señalaba Zadeh en su investidura como Doctor 'honoris causa' por nuestra Universidad, es que con el rápido crecimiento en la complejidad de las tareas de proceso de información que se pide que realicen los ordenadores, estamos llegando a un punto en el que los ordenadores tendrán que diseñarse de manera que procesen información borrosa.

De hecho, es la capacidad para manipular conceptos borrosos lo que distingue la inteligencia humana de la inteligencia de la máquina en los ordenadores de la generación actual. Sin esa capacidad, no podemos construir máquinas que pueden recoger historias no estereotípicas, traducir bien de un lenguaje natural a otro, o realizar muchas otras laboras que los humanos pueden hacer con facilidad debido a su capacidad para granular y manipular los conceptos borrosos restantes.

Todo ello está conduciendo a una situación que escandaliza a los científicos más ortodoxos: está creándose un modo de obtener conclusiones que puede denominarse 'computación con palabras', que sustituirá en ocasiones a la computación habitual realizada con números; no en vano las variables que estudia la lógica borrosa se denominan variables lingüísticas, y no en vano este sistema es el empleado por la humanidad desde tiempo inmemorial para calcular y razonar cuando la información disponible no es suficientemente precisa como para justificar el empleo de números.

Con ello se tiende a aprovechar la tolerancia de la imprecisión para alcanzar eficiencia y mejor relación con la realidad. Y, aún más, a proporcionar bases para el desarrollo de lenguajes de programación que pudieran aproximarse a los lenguajes naturales en apariencia y en capacidad de expresión.

Dice Neil Postman en 'Tecnópolis' que más matemáticas, más ciencia y más ordenadores no van a resolver el problema del hambre, la soledad, etc. y que 'la sociedad informatizada se encamina peligrosamente hacia una sociedad global y autoritaria'.

No comparto este pesimismo: creo que quizá en unos años la computación con palabras llegue a ser una metodología por derecho propio, cuyo modelo final es la mente humana. Quizá así se consiga que las máquinas se hagan un poco más 'humanas'.

4. EPÍLOGO

Si no vencí reyes moros
engendré quien los venciera.
(ROMANCERO)

He comenzado esta lección haciendo referencia al Prof. Rey Pastor y a su discurso de apertura del año académico 1913-14 y, a pesar de los años transcurridos, no ha sido posible que las empresas de todo tipo recibieran los benéficos estudios de profesionales de la Estadística, la Investigación Operativa y la Matemática Aplicada en general que, con el rigor característico de las ciencias llamadas exactas, dieran orientaciones a la dirección para la toma adecuada de decisiones en presencia de incertidumbre, puliendo la información para que ésta se presente adecuadamente a la vista de los gestores.

Otros países y, en la propia España, otras regiones más dinámicas ya lo han entendido así, y hoy son auténticos ejércitos de matemáticos los que llenan algunos departamentos de los grandes bancos mundiales o de las poderosas empresas multinacionales.

¿Qué va a ocurrir en Asturias? Recogiendo de nuevo palabras de D. Julio Rey podríamos decir que «en matemáticas no es Asturias (D. Julio decía España) un pueblo moderno; pero tampoco es un pueblo decadente ni un pueblo inepto. Es sencillamente un pueblo atrasado, que no se ha incorporado todavía a la civilización moderna; pero que conserva en su seno energías y entusiasmos suficientes para salvar la distancia producida por años de aislamiento y desorientación».

Hoy, felizmente, ya llevamos dos años dando a la sociedad asturiana, desde la recreada Facultad de Ciencias, promociones de matemáticas, con una preparación de la que nos sentimos orgullosos (en la que se considera fundamental el aspecto práctico de las teorías sin descuidar éstas) y capaces de rendir a la sociedad frutos que ésta necesita para su desarrollo.

Ellos están dispuestos. Es ahora la sociedad la que debe darles la oportunidad de rendir: en la enseñanza, en la empresa, en las administraciones... La Universidad ha jugado su baza: la preparación. Ha sido una buena e inteligente jugada. La sociedad asturiana toda ha de jugar la suya, el empleo, aún con más inteligencia, para que no sea incierto el futuro de las matemáticas en Asturias.