La opinión de los expertos

JOSE ANTONIO LÓPEZ DÍAZ. AEMET, MADRID

Introducción

El lector atento a las noticias y debates en torno a la cuestión del cambio climático causado por el hombre habrá sin duda encontrado varias veces ejemplos de apelación a la opinión de los expertos. Puede que haya leído que un determinado porcentaje de los expertos consideran que el cambio climático (antropogénico) es un problema grave para el futuro de la humanidad. O tal vez que otro porcentaje de los expertos (o quizás de los climatólogos) tienen por seguro que el hombre ha modificado el clima.

Sin duda tales apelaciones a la opinión de las personas que son consideradas expertas en una actividad o campo del conocimiento sirven para orientar a los menos conocedores sobre el tema en cuestión. Después de todo está claro que, dada la complejidad y número creciente de cuestiones y dilemas a los que debemos enfrentarnos, no queda más remedio, en la gran mayoría de los casos, que guiarnos por lo que los especialistas opinen. Esto, desde luego, parece perfectamente razonable. Pero lo que me planteo aquí es profundizar en aquellos elementos que son relevantes para valorar el grado de apoyo real que las cifras de porcentajes de expertos en apoyo (o rechazo) de determinada tesis confieren a esta (o su contraria).

Empecemos por plantear la cuestión de la manera más concreta y sencilla. Digamos entonces que tenemos cierta tesis Φ de carácter incierto (si fuera de carácter evidente ella o su negación no haría falta recurrir a la opinión experta) que vamos a suponer que tiene una probabilidad Φ de ser cierta. En el primer ejemplo anterior Φ sería "el cambio climático (antropogénico) es un problema grave para el futuro de la humanidad". Como es lógico debemos suponer la probabilidad Φ desconocida *a priori* y en consecuencia sujeta a debate. Suponemos además que existe un conjunto E de expertos cada uno de los cuales debe responder con un sí o un no a la pregunta sobre la verdad de Φ. Este esquema es el que funciona para derivar enunciados de la forma "Un porcentaje x de los expertos considera que Φ es cierta".

¿Cómo debe ser la opinión de los expertos E para que el porcentaje x refleje verazmente la probabilidad teórica φ? Una primera posibilidad, la más sencilla, es que la probabilidad de que cada experto conteste sí a la pregunta "¿Considera que es cierta Φ?", sea φ, y que además cada experto opine independientemente de los demás. En este caso la teoría de probabilidades garantiza que, a medida que el número de expertos que integran E crece hacia infinito, el porcentaje x tiende a φ. Pero sin duda este modelo es demasiado sencillo. En la vida real no es realista suponer que todos los expertos tienen exactamente la misma valoración de la tesis Φ, salvo para casos de Φ muy sencillas. Tampoco es realista suponer que los expertos opinan de forma independiente de la opinión de los demás. Más bien cada experto conocerá mejor algunos aspectos de la cuestión Φ y peor otros, de forma que su valoración final tendrá que ver con la incidencia de los temas que domina mejor en la veracidad de Φ.

Un modelo sencillo: expertos independientes

Empecemos permitiendo que las valoraciones de Φ por los expertos varíen, pero simplifiquemos suponiendo que los expertos opinan de forma autónoma. Concretando, supongamos que E lo integran N expertos, cada uno de los cuales asigna una probabilidad e(i), i= 1,…, N, a la veracidad de Φ. Por tanto podemos modelar la respuesta de cada experto i con la variable aleatoria I(i), que vale 1 si dice que Φ es cierta y vale 0 sin dice que Φ es falsa, exigiendo que pr (I(i) = 1)= e(i). Al computar el porcentaje final de expertos que contestan sí tenemos que hallar el promedio de las I(i), notémoslo por I = 1/N Σ I(i). Si queremos que los expertos acierten en conjunto con la probabilidad teórica que hemos postulado para Φ, debemos exigir que la esperanza matemática de I sea igual a φ. De aquí deducimos, tras un cálculo sencillo de probabilidades, la condición

$$1/N \, \Sigma \, e(i) = \varphi \qquad \text{(A)}$$

Preguntémonos ahora cómo varía la dispersión de los resultados de la encuesta a los expertos. Precisando, lo que preguntamos es por la varianza de la cantidad I. Tras un cálculo probabilístico un tanto prolijo llegamos al siguiente resultado:

$$\mathrm{var}(\,I\,) = 1/N\,(\varphi\,(1-\varphi) - \mathrm{disp}(e)\,) \qquad (1)$$

en la que disp(e) designa la dispersión de los valores e(i) entendida como la varianza de una variable aleatoria e' que tome los valores e(i), i=1, 2, …., N con la probabilidad 1/N cada uno (de acuerdo a la condición (A) de antes tenemos E(e') = φ). El resultado (1) muestra, en primer lugar, que con el aumento del número de expertos N la varianza disminuye a ritmo 1/N aproximadamente, y en el límite tiende a cero. Pero para N fijo vemos dos componentes en var (I). La primera φ (1 – φ) no depende más que de la probabilidad teórica de Φ, y alcanza su máximo cuando φ = 0.5, que es el caso de más incertidumbre *a priori* respecto a la veracidad de Φ. El otro sumando disp(e) mide el grado en que las opiniones de los expertos, en términos de sus probabilidades e(i), varían entre sí. En el caso, que antes mencionamos, de que todos los expertos valoren igual Φ, e(i) es constante y por tanto disp(e) = 0.

Es llamativo y, a primera vista, contra la intuición, el signo negativo de disp(e) en (1). Esto significa que cuanto mayor es la uniformidad de las predisposiciones (medidas por las probabilidades e(i)) tanto mayor es la incertidumbre en el resultado final de la encuesta de expertos. La mayor uniformidad se obtiene cuando todos los expertos tienen la misma e(i) (igual a φ por (A)) Pero se entiende mejor el caso contrario: un valor alto de disp(e) significa que las e(i) son próximas a 0 o a 1, lo que significa que los expertos tienen opiniones claras al respecto, y por tanto producen resultados menos inciertos. Cabe señalar que, en problemas complejos y para los que φ no sea extrema, es difícil que los expertos puedan tener opiniones demasiado claras, e incluso podría quizá afirmarse que si las tienen pueden hacer dudosa su valoración del problema.

Para finalizar el análisis del resultado (1) veamos el valor máximo que puede alcanzar disp(e). Es bastante intuitivo que esta varianza de las e(i) se maximiza cuando las e(i) son lo más extremas posible, esto es, cuando valen todas 1 o 0. Teniendo en cuenta la restricción $1/N\ \Sigma\ e(i) = \varphi$ y suponiendo para simplificar que $N\varphi$ sea entero, entonces $N\varphi$ de las e(i) valen 1 y $N(1-\varphi)$ de las e(i) valen 0. En este caso un cálculo muestra que $disp(e) = \varphi(1-\varphi)$. Cuando $N\varphi$ no sea entero el resultado es aproximado, pero para N moderadamente grande es muy aproximado. Por tanto, designando por χ a la incertidumbre intrínseca de Φ que no depende de las e(i), $\chi = \varphi(1-\varphi)$, concluimos que var(I) varía entre χ/N (cuando las probabilidades de todos los expertos son la misma) y cero (cuando sus probabilidades son lo más extremas posibles). En este último caso es también evidente a priori que var(I) debe ser cero pues en la encuesta cada experto responde seguro bien sea sí o bien sea no (al ser e(i) igual a 0 o 1 para todos los expertos).

Expertos y no tan expertos

Pasemos ahora a relajar la condición de independencia entre las opiniones de los expertos, que es sin duda una simplificación excesiva en casos de cuestiones tan complejas como el papel del hombre en el cambio climático presente y previsto. En cambio, para simplificar algo el análisis, vamos a suponer que no hay variación en las opiniones de los expertos autónomos.

Es razonable esperar que la mayoría de los expertos se guíe en gran medida, en su valoración de la tesis Φ, por la opinión de otros expertos a los que juzguen más competentes, y que para muchos aspectos que no sean de su especialidad supediten totalmente su juicio a las opiniones de estos otros expertos. Vemos pues que se abren muchas posibilidades. Pero para analizar, siquiera sea cualitativamente, alguno de los efectos más obvios de estas complejas interacciones, podemos proponer un modelo relativamente sencillo.

En este modelo distinguimos dos clases de experto, según su grado de independencia de juicio sobre la cuestión Φ. En concreto suponemos que de los N expertos en total, hay una proporción ε de ellos que actúan como expertos autónomos stricto sensu, es decir, actúan de forma totalmente independiente con probabilidades e(i), como supusimos en el modelo anterior. Pero la proporción de expertos $1-\varepsilon$ restante suponemos que actúan totalmente supeditados a la opinión de los otros expertos. Si además actúan racionalmente se limitarán a contestar sí a la pregunta sobre la validez de la tesis Φ cuando la mayoría de los expertos autónomos digan sí, y contestarán no en caso contrario. Naturalmente, en la práctica, uno espera una gradación más o menos continua en los grados de autonomía de los expertos, pero este sencillo modelo bifásico puede ilustrar cualitativamente los nuevos rasgos que aparecen.

Suponemos, como en el modelo de monofásico anterior, que el promedio de las probabilidades de los expertos *stricto sensu* es igual a la probabilidad teórica desconocida φ (como en (A)). Pero con este modelo bifásico aparece una nueva característica importante: el que (A) se cumpla para los expertos autónomos no garantiza que la esperanza matemática del porcentaje de síes en la encuesta de los N expertos sea igual a φ, como sería sin duda deseable. Dicho de otro modo, aparece en general un sesgo sistemático en el resultado de la encuesta. De hecho se tiene, suponiendo que todos los expertos autónomos tienen la misma $e(i) = \varphi$, que:

$$E(I) = \varepsilon\varphi + (1-\varepsilon)\ pr\ \{Bi[N\varepsilon, \varphi]/\ N\varepsilon \geq 0.5\} \qquad (2)$$

donde Bi[n, p] denota la variable aleatoria binomial con parámetros n y p. Esta variable aleatoria describe el resultado del siguiente experimento: tenemos n variables aleatorias que toman valor 1 con probabilidad p y valor 0 con probabilidad 1-p y hallamos la suma de unos. En esta y sucesivas expresiones para E(I) con este modelo supondremos para simplificar que el número de expertos autónomos $N\varepsilon$ es impar: esto simplifica las expresiones ya que impide el empate entre expertos. Hay un caso particular en que el modelo bifásico es insesgado (aparte de los dos triviales $\varphi = 0$ y $\varphi = 1$): cuando $\varphi = 0.5$ se tiene que la probabilidad del miembro derecho de (2) vale justo 0.5 por simetría, y entonces $E(I) = 0.5$.

Para el caso de $N\varepsilon$ grande podemos usar la aproximación normal a la variable binomial, con lo que, denotando por Z a la variable normal estándar, obtenemos:

$$E(I) = \varepsilon\varphi + (1-\varepsilon)\ pr\ \{Z \geq [N\varepsilon/(\varphi(1-\varphi))]^{1/2}\ (0.5-\varphi\} \qquad (3)$$

Veamos qué sucede cuando $N\varepsilon$, número de expertos autónomos, tiende a infinito. En este caso la probabilidad del segundo miembro de (3) tiende a 1 si $\varphi > 0.5$ y a 0 si $\varphi < 0.5$. De aquí deducimos que el porcentaje esperado de la encuesta E (I) tiende a:

$$E(I \mid N\varepsilon = \infty) = \varepsilon\varphi + (1-\varepsilon)\ \Omega\ (\varphi > 0.5) \qquad (4)$$

donde $\Omega\ (\varphi > 0.5)$ vale 1 si $\varphi > 0.5$ y 0 en caso contrario.

Las dos aproximaciones (3) y (4) a (2) se representan en la figura 1. En esta figura se ha supuesto $N = 200$ y $\varepsilon = 19/200$, lo que supone aproximadamente un 10% de expertos autónomos. Se ve que para este caso la aproximación límite (4) funciona muy bien a partir de φ = 0.8, pero la aproximación normal es muy buena para todo el rango de valores de φ. Como tanto para φ = 0.5 como para φ = 1 el modelo bifásico es insesgado para E(I) esperamos las mayores desviaciones o sesgos E(I) – φ para valores de φ intermedios, lo que es claramente apreciable en la figura. En concreto para este caso el mayor sesgo sucede para φ = 0.7 y vale 0.24.

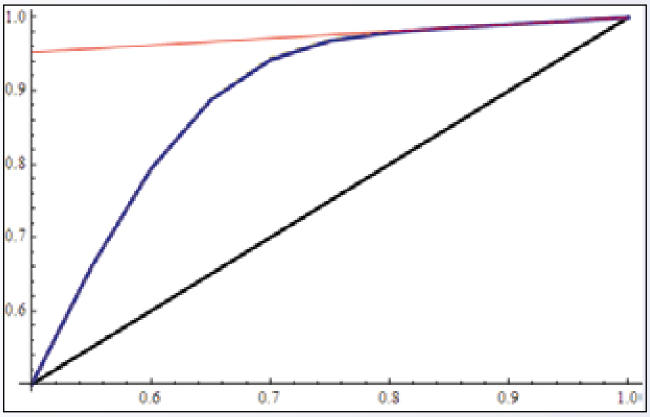

Fig.1 : Para N = 200 y ε = 19/200, ploteo del valor teórico φ (línea negra gruesa), la expresión exacta (2) para E(I) (línea gruesa azul), la aproximación normal (3) (línea de trazos casi indistinguible de la anterior) y la aproximación límite (4) (línea roja). En abscisas valores de φ.

En la figura 2 se representa para N fijo = 200, los sesgos E(I) – φ en función de φ para tres valores de ε (0.05, 0.1 y 0.3). La apariencia de las curvas es similar, con forma de campana más o menos apuntada. Una característica importante del sesgo

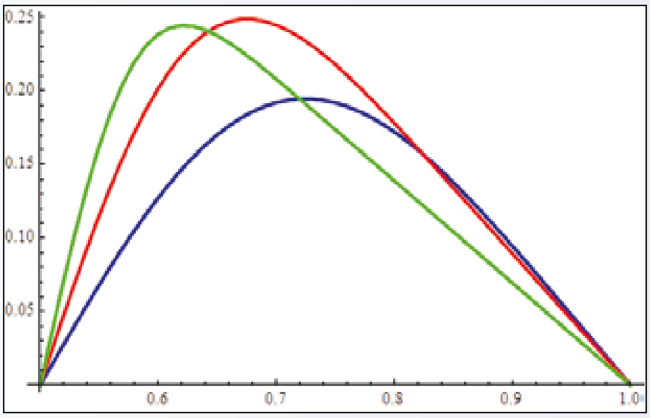

Fig.2 : Para N= 200 , sesgos E(I) – φ en función de φ para tres valores de ε: 0.05 (línea azul), 0.1 (línea roja) y 0.3 (línea verde)

de este modelo es que es siempre positivo (para φ > 0.5 como estamos suponiendo, es siempre negativo para el caso simétrico φ < 0.5). Esto quiere decir que este modelo bifásico siempre tiende a exagerar, bien sea la plausibilidad a priori, bien sea su contrario, de la tesis Φ.

En la figura 3 se ha representado para distintos valores de N = 50, 100, 200 y 500, y distintos valores del parámetro ε, el valor máximo del sesgo. Se ha añadido la recta límite correspondiente al límite N ε = infinito que, según (4) da un valor máximo para E(I) – φ = 1/2 (1 - ε). Vemos que al crecer N crece esta diferencia máxima, llegando para N = 500 a rondar 0.3 para to-

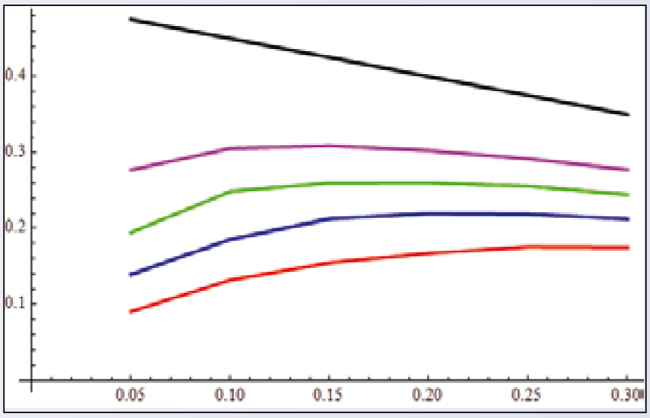

Fig. 3: Sesgos máximos con el modelo de expertos bifásico para distintos valores de ε en abscisas. Las líneas corresponden a distintos valores de N = 50, 100, 200 y 500, que aparecen en la figura de abajo a arriba. La línea recta superior corresponde al límite N ε = infinito

dos los valores de la proporción de expertos autónomos ε considerados. La variación con ε es modesta para N fijo.

En la figura 4 se han representado los valores de φ para los que se alcanzan los mayores sesgos de la figura 3. Vemos que hay una tendencia a que a valores menores de N correspondan valores de φ que maximizan el desvío mayores. Para el caso límite N ε = infinito el máximo de desviación corresponde a φ = 0.5. Excluyendo este caso límite, para los valores de N y ε considerados φ se mueve en el rango [0.6, 0.8].

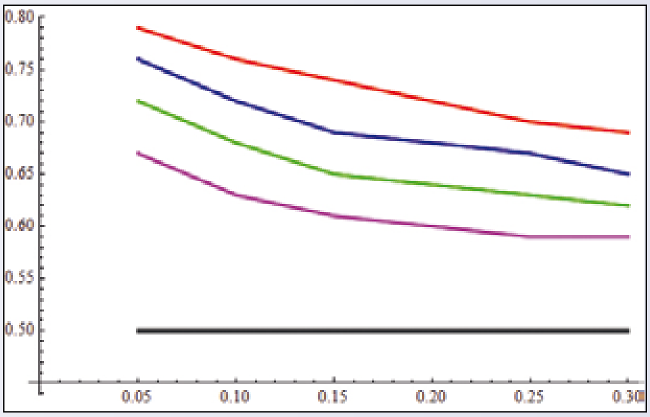

Fig. 4: Valores de φ para los que se producen las desviaciones máximas de la figura 2. El color de las líneas corresponde a los mismos valores de N que en la fig. 2, al igual que el eje de abscisas.

Hasta ahora no hemos dicho nada acerca de la incertidumbre (varianza) en los resultados de la encuesta de expertos en el modelo bifásico. Usando el teorema sobre la descomposición de la varianza y condicionando a minoría/mayoría de expertos autónomos se puede derivar una expresión analítica para la varianza, pero es complicada. Para hacernos una idea cualitativa del efecto de la introducción de expertos "heterónomos" en la encuesta, podemos ver la relación entre la varianza del modelo bifásico con un número determinado de expertos autónomos y la varianza de una encuesta realizada exclusivamente a esos expertos autónomos. Esto se ha hecho en la fig. 5, en la que se han dibujado líneas, cada una para un valor de φ, de la razón de las desviaciones estándar del modelo bifásico con N=200 y del modelo monofásico para varios valores del número de expertos autónomos (igual en ambos modelos) en abscisas. La línea superior azul corresponde a φ=0.6, y en este caso, de bastante incertidumbre en la veracidad de la tesis Φ *a priori*, lo que vemos es que la desviación típica del modelo bifásico supera con mucho a la del monofásico. En cambio la línea inferior, de color verde, corresponde al caso de mucha certidumbre a priori sobre Φ, en concreto φ = 0.9. En este caso la varianza de la encuesta bifásica se reduce drásticamente con relación a la encuesta limitado a los expertos autónomos. Por tanto en este caso la variabilidad que cabe esperar a priori entre las opiniones de expertos autónomos se filtra grandemente por efecto del "seguimiento de la mayoría" de los expertos "heterónomos".

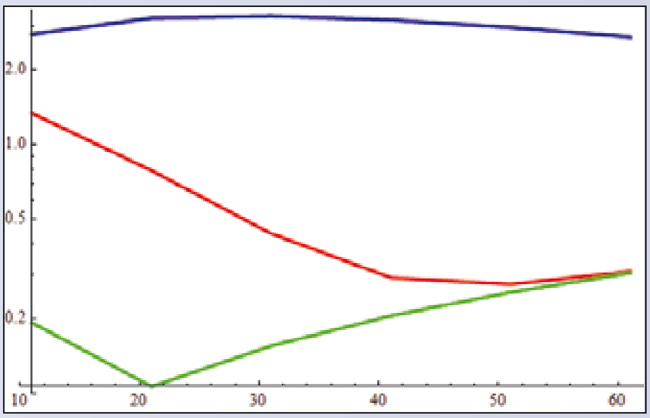

Fig.5: Representación, con escala logarítmica, de la razón de desviaciones estándar entre el modelo bifásico con número de expertos autónomos en abscisas y el monofásico con el mismo número de expertos. Para el bifásico se ha tomado N = 200. Las líneas, de arriba abajo, corresponden a φ = 0.6, 0.75 y 0.9 resp.

Conclusiones

Del anterior análisis con este modelo de dos tipos de experto, deducimos que siempre se produce un acercamiento del porcentaje esperado de síes en la encuesta hacia 100 para φ>0.5 (hacia 0 si φ < 0.5 por simetría). La magnitud de este desplazamiento del porcentaje respecto de valor teórico φ puede alcanzar valores del orden de 0.2 o más para valores teóricos de φ en la parte central del intervalo [0.5, 1] (o simétrico respecto de 0.5). Además para tesis con incertidumbre a priori reducida los porcentajes de la encuesta muestran mucho menos variabilidad que en el caso de una encuesta limitada a los expertos autónomos. Dicho de forma más fácilmente comprensible, la conclusión es que *el resultado de la encuesta de expertos en este modelo bifásico exagera la certidumbre respecto de la veracidad/falsedad de Φ.*

Estas conclusiones se ilustran con ejemplos numéricos en la tabla 1. Aquí se supone un número fijo de expertos autónomos = 21, pero se deja variar tanto la probabilidad teórica φ de veracidad de la tesis propuesta a los expertos como el número total de expertos en el modelo bifásico. Las entradas de la tabla dan el valor esperado del porcentaje de síes en la encuesta más/menos la desviación típica. Aquí podemos ver, yendo a lo largo de las filas, cómo el valor esperado del porcentaje de respuestas positivas se va inflando hacia el 100%, por ejemplo para φ = 0.7 vemos cómo partimos de un valor esperado de 0.7, igual a la probabilidad teórica φ cuando no hay expertos "heterónomos", hasta llegar a un valor esperado de 0.94 cuando en total hay 200 expertos. Además vemos una ilustración del efecto en la incertidumbre del resultado de la encuesta reflejado en la fig. 5: para valores de φ que connotan más incertidumbre, el 0.6 por ejemplo, la incertidumbre en el resultado de la encuesta aumenta con el número total de expertos. Pero en claro contraste, cuando hay más certidumbre a priori, como para φ = 0.9, vemos que la incertidumbre en el resultado de la encuesta disminuye drásticamente al aumentar en número de expertos. En la celdilla inferior derecha de la tabla se muestra que con este φ alto y 200 expertos en total es casi seguro que el resultado de la encuesta estará muy próximo a 99%.

φ \ N	21	40	100	200
0.6	0.60 ± 0.11	0.71 ± 0.22	0.78 ± 0.32	0.80 ± 0.35
0.7	0.70 ± 0.10	0.83 ± 0.11	0.92 ± 0.14	0.94 ± 0.15
0.8	0.80 ± 0.09	0.89 ± 0.05	0.96 ± 0.03	0.98 ± 0.03
0.9	0.90 ± 0.07	0.95 ± 0.03	0.98 ± 0.01	0.99 ± 0.01

Tabla 1: Valor esperado de la proporción de expertos que contestan positivamente a una tesis de probabilidad teórica φ en filas, con el modelo bifásico para un número fijo de expertos autónomos = 21, y número total de expertos en columnas.

El mensaje que cabe deducir del anterior análisis es que al utilizar, como se hace a menudo, el porcentaje de expertos que apoyan determinada tesis sobre un tema de gran complejidad como el cambio climático antropogénico, es de esperar que se transmita una impresión exagerada de certidumbre. Esto se debe a que muchos (o todos) los expertos opinarán de forma en mayor o menor grado condicionada a la opinión de otros expertos.