

Frequentistes *versus* bayesians: uns apunts des del sentit comú

Jordi Vallverdú Segura

1. Un món causal?

Quan ens preguntem per l'origen de la racionalitat humana trobem una idea que sobresurt per damunt de les altres: la causalitat. Els primers homínids desenvoluparen habilitats manuals i tècniques que permeteren la seva supervivència immediata, però, en les seves ments, la noció de causa s'anava conformant a mesura que el mateix llenguatge progressava en precisió i abastament de la realitat. Calia establir lligams correctes entre els esdeveniments de la vida quotidiana per a la supervivència: la migració dels animals per a la caça, les estacions per a la recollida d'aliments. Podien observar una relació entre el cel i la terra, que semblava constituir l'ordre de les coses. Els astres, amb els seus processos aparentment immutables i eterns, foren un mirall d'eternitat regular per a uns éssers dèbils, però actius i curiosos alhora, com ho eren els humans. Aviat s'establiren nexes entre els esdeveniments del cel i els de la terra. Hi havia un lligam físic entre el cel i la terra, més enllà d'una hierogàmia mística, que els estudiosos i savis havien d'explicar. El pensament racional estava emergint a poc a poc a través de desenes de generacions mitjançant l'observació de les estrelles. No pas gratuïtament, Epicur afirmà: «Si no ens preocupés tot allò referent a les estrelles i a la mort, no tindríem cap necessitat de la filosofia»¹. Amb Aristòtil, la idea de causalitat (en els seus quatre vessants) quedà formalment constituïda i formaria part de la posterior concepció ontològica del món. Val a dir que l'Estagirita no tancà completament la possibilitat de l'existència de fenòmens que avui considerariem probabilistes i que aleshores ell agrupà amb els conceptes de «sort» i «casualitat» (*Física* II, 4-6). Ara bé, aviat aparegueren dubtes sobre la idea de l'atzar i el seu possible paper en la natura, com mantenien els atomistes Demòcrit i Lucreci. Segons Hald (1988), la idea de probabilitat, encara que no matematitzada, ja existeix en l'antiguitat clàssica i es va transme-

1. DIÒGENES LAERCI, *Vides dels filòsofs*, Llibre X, Epicur.

tent durant segles. Podem trobar rastres d'idees probabilistes en textos talmúdics i rabínics, i la influència aristotèlica sobre els filòsofs romans², els quals, al seu torn, transmeteren llurs idees a la filosofia escolàstica, la qual s'enfrontà al repte de conciliar la rica filosofia aristotèlica amb les concepcions deterministes de l'Església. Tot i que durant el període medieval no hi ha constància sobre l'existència de cap càlcul de probabilitats, sabem que l'any 960 el bisbe Wibold de Cambrai enumerà correctament 56 diverses possibilitats de jugar amb tres daus. Posteriorment, en un poema del segle XIII, *De vetula*, trobem les 216 possibilitats de resultats de les tirades amb tres daus.

Haurem d'esperar fins l'any 1662 per a trobar els primers exemples de probabilitats numèriques, sorgits d'Antoine Arnauld i de Pierre Nicole (1662, *La logique ou l'art de penser*, usualment denominada com la «Lògica de Port Royal»). Els aspectes matemàtics implícits en els jocs d'atzar interessaren a autors com Cardano, Pascal, Fermat o Huygens (Weatherford, 1982), tot i que desestimaren considerar la realitat de la incertesa i s'immergiren en un paradigma newtonià estrictament mecanicista. De fet, Huygens (1657, *De ratiotiniis in aleae ludo*) és considerat com el primer autor que reflexionà sobre els aspectes racionalitzables dels jocs d'atzar, o sia sobre la idea de probabilitat. Posteriorment, De Moivre redactà l'influent *De mensura sortis* (1711) i donà pas als estudis sobre probabilitat estadística preocupats per la naturalesa real de les coses. En 1789 Laplace escriví l'*Assaig filosòfic sobre les probabilitats*. Deixant de banda el vessant purament matemàtic per centrar-se en els aspectes epistemològics, Hume atacà durament el concepte de causa com a producte dels hàbits adquirits, no pas de la mateixa realitat. La revolució de la física a la fi del segle XIX i a començaments del segle XX provocà l'esfondrament de la idea de causa i exigí un replantejament del concepte d'«observació», alhora que Gödel dinamitava la matemàtica axiomàtica defensada per Hilbert. Així, doncs, Bertrand Russell afirmà l'any 1913: «*The law of causality [...] is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm [...] The principle "same cause, same effect", which philosophers imagine to be vital to science, is therefore utterly otiose*» (Suppes, 1970, p. 5). Malgrat l'opinió de Russell, la reflexió al voltant de la causalitat i la seva inferència ha continuat fins als nostres dies, encara que en alguns casos, com el de Nancy Cartwright, en forma d'una teoria probabilística de la causació³.

Ara bé, cal entendre la discussió al voltant de la idea de causa des de la dicotomia determinisme-indeterminisme i la seva conseqüent realisme-relativisme. En realitat, s'enfronten dues visions d'entendre la constitució mateixa del món i del paper de la racio-

2. Més enllà de l'àmbit filosòfic, trobem les taules del jurista romà D. Ulpià (s. III de la nostra era), les quals contenien pautes per a la conversió legal de les pensions. La taula d'Ulpià, continguda en el *Corpus Iuris Civilis* Digesta XXXV, 2, 68, fou analitzada detingudament per Nicholas Bernoulli, el qual hi veié per primer cop la presència de càlcul estadístic.

3. L'estudi de referència és N. CARTWRIGHT, N. (1979) «Causal Laws and Effective Strategies», *Noûs* 13, pp. 419-437.

nalitat. Un exemple el trobem en la petita obra *A World of Propensities* de Karl R. Popper, amb la que defensa l'existència d'indeterminacions objectives tot i haver renunciat a la causalitat. Entrem en un món de propensions estadístiques⁴.

2. Uns apunts sobre les dues escoles estadístiques

Podem dir que dins la disciplina de l'estadística s'han format dos grans corrents que difereixen en la seva concepció ontològica de la realitat⁵: d'una banda, el model de l'*estadística clàssica* o *freqüentista*, de l'altra, el model *bayesià*⁶, els quals defensen opinions diferents pel que fa a la relació entre els judicis de probabilitat i l'evidència empírica de les freqüències trobades en les mostres.

La clau per a diferenciar ambdues escoles⁷ consisteix a preguntar-nos quina és la seva posició davant allò que sigui la probabilitat: per als freqüentistes, la probabilitat és una freqüència relativa, una proporció, mentre que, per als bayesians, la probabilitat és una mesura de la creença personal. Posem un exemple clàssic: el que consisteix a llançar una moneda a l'aire i mirar si surt cara o creu. Per a un freqüentista, la probabilitat (P) que surti cara s'obté a partir del llançament múltiple d'aquesta moneda, cosa que dona com a resultat final una probabilitat (P) de sortir cara o creu, que és de 0.5; el bayesià podria jutjar el valor de P sense llançar una sola moneda, a partir dels coneixements previs de qualsevol tipus, en aquest cas, de física, i allò que l'experiència li recomana. Els freqüentistes, a més, mai no considerarien l'estimació de la probabilitat d'esdeveniments singulars o únics, ja que sempre treballen a partir de diversos resultats, de què obtenen una mitjana o una freqüència. El bayesians compten amb la fórmula que el prevere Thomas Bayes creà⁸:

$$P(T|E.B) = \frac{P(T|B)P(E|B.T)}{P(T|B)P(E|B.T) + P(\sim T|B)P(E|B.\sim T)}$$

T equival a la teoria o la hipòtesi d'analitzar B, la informació disponible de base, i E, a qualsevol nova evidència que tot just hem adquirit. Així, el costat esquerre de l'equació (és a dir, $P(T|E.B)$) representa la probabilitat (P) de la nostra hipòtesi (T) a partir de la informació (B) disponible de base i la nova evidència (E). Per dir-ho d'una altra manera, quan llegim $P(T|E.B)$ podem també entendre-ho com a $(E \wedge B \rightarrow T) \rightarrow P_T$. Aquesta equació es coneix com a «probabilitat a posteriori» («*posterior probability*»).

4. H. PRICE: «Agency and Probabilistic Causality», *British Journal of the Philosophy of Science*, UK, 1991, n. 42, pp. 157-176, en un intent d'abandonar la complicada i difícilment defensible idea de causa, present, encara que de forma apagada i exemplificada, en les relacions causals mantingudes per Cartwright, proposa canviar-la per la de relacions d'associació.

5. En va no s'ha considerat pas l'estadística un art, tot i ser alhora una ciència, ja que requereix la destresa i l'experiència de l'estadista per a la seva correcta execució. W. J. REICHMANN: *Use and Abuse of Statistics*, UK: Penguin Books, 1961, p. 16; C. RADHAKRISHNA RAO: *Estadística y verdad* Barcelona: PPU, 1994., p. 68. Aquesta divisió és molt general, però bàsica quants als meus propòsits. Existeixen altres teories sobre la probabilitat, com la de l'*a priori*, la de freqüència relativa o la subjectivista, R. WEATHERFORD: *Philosophical Foundations of Probability Theory*, USA: Routledge & Kegan Paul, 1982.

6. Thomas BAYES el presentà en 1763 dins l'obra *An Essay Towards Solving a problem in the Doctrine of Chances*.

7. Exemples extrets de R. L. LILFORD, D. BRAUNHOLTZ: «For Debate: The statistical basis of public policy: a paradigm shift is overdue», dins *British Medical Journal*, 1996, n. 313, pp. 603-607.

8. Per al desenvolupament i les explicacions del teorema de Bayes, segueixo el text de SALMON (1990).

La secció de la dreta conté quatre expressions de probabilitat: les dues primeres, $P(T|B)$ i $P(\sim T|B)$, són denominades «probabilitats *a priori*» («*prior probability*»); representen la probabilitat que la nostra hipòtesi sigui veritable o falsa a partir únicament de la informació de base, sense tenir en compte la nova evidència. Les dues probabilitats que queden, $P(E|T.B)$ i $P(E|\sim B.T)$, són conegudes com a «funcions de versemblança» («*likelihoods*»). La primera ens explica la probabilitat que es donés nova evidència si la nostra hipòtesi fos veritable i la segona recull el mateix cas, però essent la nostra hipòtesi falsa. A diferència de les dues probabilitats prèvies, les probabilitats poden ser establertes independentment, ja que el valor de l'una no determina el valor de l'altra.

Quan coneixem les probabilitats *a priori*, aquestes són la conseqüència del teorema del càlcul de probabilitats, però si les probabilitats *a priori* són desconegudes i hom assumeix amb Bayes que són *equiprobables*⁹, estem obligats a determinar la probabilitat inversa de les causes. Amb l'equació, Bayes pretén quantificar la incertesa respecte de diverses hipòtesis en una distribució *a posteriori*, de manera que un «*rational agent should identify as admissible all and only those feasible options wich maximize conditional expected utility*»¹⁰. Així, la probabilitat, entesa des d'una perspectiva bayesiana, mesuraria el grau subjectiu de creença.

La discussió filosòfica s'ocupa únicament d'una discussió de tipus epistèmic, a un nivell metateòric proper a reflexions ontològico-metafísiques, en què el problema de la inferència científica esdevé de vital importància. És en aquest darrer àmbit on podem analitzar les contribucions de Reichenbach, Suppes, Salmon o Cartwright.

3. La metareflexió filosòfica sobre ambdues escoles estadístiques

En l'àmbit filosòfic, podríem parlar, doncs, de dues escoles¹¹ confrontades, a partir de dues actituds clares, com recull Mayo (1997a):

Actituds	Enfocament	Posició epistèmica	Autors
Bayesians	Relació d'evidència (<i>Evidential-relation view</i>)	Subjectivistes	Bayes, Carnap, Salmon
Clàssics / Freqüentistes	Error estadístic (<i>Error statistical view</i>)	Objectivistes	Reichenbach, Giere, Mayo

L'interès renovat pels models bayesians s'inicià l'any 1954 amb la publicació *Foundations of Statistics*, de L. J. Savage. Durant la

9. L'*equiprobabilitat* és coneguda com a «Postulat de Bayes». Jean Bernoulli establí, a partir del teorema, el «principi de raó no suficient»: «cal jutjar les alternatives sempre com a *equiprobables* si no tenim cap raó de pes per a escollir-ne una en comptes de l'altra» R. WEATHERFORD: *Philosophical Foundations of Probability Theory*, USA: Routledge & Kegan Paul Ltd, 1982, p. 24.

10. I. LEVI: *Hard Choices. Decision making under unresolved conflict* USA: Cambridge University Press, 1986, p. 61, citant R. C. JEFFREY.

11. També tenim autors que, sense pertànyer directament a aquestes escoles, han analitzat aspectes de l'estadística. Un cas curiós és el de l'alemany Karl Marbe, que en els seus estudis de començaments del segle xx incorregué a interpretar els resultats estadístics en la «fal·làcia del jugador». Després d'estudiar 200.000 registres de naixements en quatre ciutats de Baviera, conclougué que la probabilitat que una parella tingués un fill mascle augmentava si uns dies abans havia nascut una gran quantitat de femelles (exemple citat per I. HACKING: *El surgimiento de la probabilidad*, BCN: Gedisa, 1995).

dècada següent, la psicologia intentà esbrinar si els éssers humans actuaven com a agents bayesians, cosa que semblaren desmentir. La discussió respecte de la racionalitat dins la ciència i el paper de l'estadística ha continuat essent un punt clau en el programa de recerca de la filosofia de la ciència, atès que conté la clau de volta per a entendre el funcionament i la dinàmica de la ciència. Thomas Kuhn (1977) considerà que els científics eren agents bayesians quan havien de triar entre teories confrontades, malgrat que altres autors s'oposessin radicalment a tal plantejament¹². Salmon (1990) proposa una línia de connexió directa entre Kuhn i Bayes, on l'autor introdueix el que denomina l'«expectació de l'evidència», que va més enllà del tancat enfocament de Carnap. L'expectació de l'evidència explicaria la divergència entre el tipus d'hipòtesi obtingut per part de diversos científics: «*Individual scientists, using the same algorithm, may arrive at different evaluations of the same hypothesis because they plug in different values for the probabilities*»¹³.

Per a Salmon, aquest plantejament no implica caure en un subjectivisme imprevisible sinó en un model dels processos racionals científics on es reconeix un paper subjectiu que, de fet, sembla donar-se en el fer de les ciències, sense que això signifiqui un rebuig dels protocols acceptats per la comunitat científica. Salmon consideraria que la racionalitat del procés recolza en el que jo he denominat anteriorment (Vallverdú, 2002) «arguments formacionals», és a dir, aquells que recorren a la *formació* (acadèmica, tècnica) i l'*experiència* dels científics implicats. De fet, hem pogut veure al llarg de la tesi com en moments clau aquestes són les característiques que defineixen un expert i que constitueixen alhora els elements polèmics d'una controvèrsia en la investigació científica, ja que s'apel·la al valor qualitatiu acadèmic d'un individu, quelcom difícil de saber en un context quantitatiu (premis, beques, publicacions, formació...).

Deborah Mayo desenvolupà en contra del bayesianisme una extensió de la metodologia Neyman-Pearson en una obra controvertida: *Error an the Growth of Experimental Knowledge* (1996). Howson (1997) té un article amb un títol explícit: «*Error Probabilities in Error*», en què critica la lleugeresa amb què aparentment Mayo ha exclòs de la ciència el mètode bayesià, quan fins i tot hi ha constància del seu ús conscient per part de diverses disciplines¹⁴. Mayo defensaria una *estadística de l'error* a partir de la submissió de les hipòtesis H a tests severes. Segons Mayo (1997b): «*This passing test T (with result e) is a severe test of H just to the extent that there is a very low probability that test procedure T would yield such a passing result, if hypothesis H is false.*»

Al capdavant, Mayo intenta controlar -introduint-la dins el model formal- la presència de l'error estadístic¹⁵ per tal d'evitar un

12. Segons T. KUHN: *The Essential Tension*, Chicago: Chicago University Press, 1977, p. 328. R. N. GIERE: *Explaining science: a cognitive approach*, USA: The University of Chicago Press, 1988, per contra, ataca a aquells que consideren que els científics són bayesians, quan en realitat la major part -els bons, segons Giere- són freqüentistes. Em remeto als articles de J. M. BLAND, G. D. ALTMAN: «Bayesian and frequentists», dins *British Medical Journal*, 1998, n.317, pp. 1151-1160 i R. L. LILFORD, D. BRAUNHOLTZ: «For Debate: The statistical basis of public policy: a paradigm shift is overdue», dins *British Medical Journal*, 1996, n. 313, pp. 603-607, en què aquests autors mostren l'ús en augment de la metodologia bayesiana dins la comunitat mèdica.

13. SALMON, 1990, p. 181.

14. Em remeto a l'editorial amb data de 7 de setembre de 1996 del *British Medical Journal*: «Bayesian Statistical Methods. A Natural Way to Assess Clinical Evidence», pp. 569-570.

15. En tot cas, Popper havia enumerat dotze principis respecte de l'actitud davant l'error que es troben recollits en l'obra *Sociedad abierta, universo abierto. Conversación con Franz Kreuzer*, Madrid: Tecnos, pp. 139-158.

bayesianisme, que considera extremadament subjectiu i contrari a l'esperit universalista de la ciència. Giere, tot criticant també el bayesianisme, opta per una evolució del model de la teoria de la decisió, denominada «*satisficing model*», en què l'activitat científica ha de preferir escollir una teoria veritable a una teoria falsa. La forma de resoldre els problemes o les controvèrsies passa per l'ús de nova informació, obtinguda, per exemple, a partir d'experiments crucials entre altres opcions.

Les crítiques al bayesianisme com a no-part integrant de la ciència contemporània demostren la total desconeixença de l'estat de la qüestió en què es troba aquesta ciència. En l'anàlisi de l'activitat dels investigadors, hom ha d'admetre la presència de divergències dins la recerca científica que no poden ser merament atribuïbles als resultats experimentals sinó a la posició prèvia de l'investigador, que parteix de la seva visió particular i les seves eines per a captar les característiques del món, que alhora modelitza i redueix en un espai limitat de recerca.

4. El context d'ús d'ambdós models

A partir de la idea segons la qual l'objectiu de la ciència consisteix a «resoldre problemes» (Lakatos, 1977), veurem alguns exemples dels usos pràctics d'ambdós models, bo i evitant tot debat metateòric sobre la naturalesa estadística del món.

I. Bayesians

La polèmica al voltant de la toxicitat carcinogènica de la sacarina, que abraça el període comprès entre els anys 1897 i 2000, implicà l'ús d'ambdós models. En 1985, Chankong emprà un mètode bayesià, el CPBS, creat a mitjans dels vuitanta per tal d'analitzar estudis en els quals part dels resultats obtinguts són positius, mentre altres es mostren negatius, tot evitant al màxim elements de judici subjectius. Segons una prestigiosa institució nord-americana que emprà aquest mètode, «*In this procedure, Bayes' decision theory is used to calculate the probability that a test chemical is or is not a carcinogen, taking advantage of the known characteristics of the various assays*» (NRC, 1985, p. 73). Cal admetre, però, que aquest mètode ha estat emprat per a disposar d'un mètode ràpid i barat a l'hora de treballar en tests de curta durada i no els habituals bioassaigs animals.

II. Freqüentistes

Aquí topem amb els conflictes del valor de probabilitat o *valor-p*, defensat per àmbits de la ciència reguladora, la qual necessita punts clars per a la inferència i els judicis legals. Aquest valor con-

sisteix en la probabilitat que, bo i suposant que la hipòtesi nul·la¹⁶ sigui certa, l'estadístic de contrast adopti un valor igual o més extrem que l'observat en la mostra; o sia, es tracta de mesurar la discrepància de les dades amb les hipòtesis. Giere (1991) es pregunta com hauria de ser d'alta una probabilitat perquè es poguessin establir relacions d'evidència, i hi troba una resposta en el 95%, que han adoptat de forma general els estadístics, «*this is a reasonable convention*»; és a dir, que només en un 5% dels casos la relació causa-efecte es donaria per atzar. Aquesta probabilitat es denomina «*valor-p*», i és el valor de tall per a decidir quan un resultat és significatiu i quan no ho és. Aquest valor s'expressa en una escala que va del 0 (màxima fiabilitat) a l'1 (màxim error o manca d'evidència), de manera que un resultat és vàlid fins a arribar a un *valor-p* de 0'05. Així, 5 de cada 100, o una de cada vint vegades, un experiment dona falsament un resultat, sia un fals negatiu, és a dir, que diu que una substància no és cancerígena, quan en realitat ho és, o un fals positiu, posat cas que afirmi la cancerogenicitat d'una substància quan no ho és. El *valor-p* fou creat la dècada dels anys vint del segle xx per Ronald Fisher, que proposà valorar una hipòtesi (H_0) a través d'una observació concreta (d_0), així com la construcció d'aquest número, el conegut «*valor-p*» (*p-value*).

$$p = \text{Prob} (d \geq d_0 * H_0)$$

16. Una hipòtesi nul·la (per exemple: «els subjectes no experimentaran canvis en la seva pressió sanguínia degut a l'administració de la substància experimental») és emprada per a discriminar qualsevol possibilitat excepte aquella que l'investigador està intentant demostrar, una assumpció respecte de la població de recerca que pot ser o no rebutjada com a resultat de l'experiment; aquesta és emprada a causa del fet que la major part de mètodes estadístics són menys decisius a l'hora de demostrar quelcom com a veritable, que no pas a l'hora d'aportar una forta evidència sobre la seva falsedat.

17. W. R. HAVENDER: «The science and politics of cyclamate», *The Public Interest*, USA, 1983, n. 71, p. 25.

18. LIEBERMAN & KWON (1998) proporcionen una extensa llista de falsos positius que han tingut un gran impacte en la societat nord-americana i, de retruc, en l'europea i en la mundial.

19. Dr. Jere E. Goyan, especialista en farmacocinètica, fou el primer farmacèutic que ocupà aquest càrrec en la FDA (21/10/79-20/1/81), en què es trobà amb restes de la controvèrsia del ciclamat, al mateix temps que esclatava la polèmica al voltant de la síndrome de xoc del tampó o la de la contaminació alimentària del bifenil boliclorinat. En ser escollit president Ronald Reagan, tornà al deganat de la University of California-San Francisco School of Pharmacy fins que més tard s'incorporà com a executiu dins la indústria farmacèutica.

Havender¹⁷ n'ha indicat alguns problemes. El primer de tots consisteix en el fet que la fracció de falsos positius és més alta quan la *ratio* de substàncies cancerígenes és més baixa que el *valor-p* escollit. D'aquesta manera, a mesura que decreixi el tant per cent de substàncies cancerígenes, augmentarà el nombre de falsos positius¹⁸ per a un determinat *valor-p*. Per exemple, si en un experiment amb 100 rates, 80 no desenvolupessin realment càncer i 20 sí (se suposa que a partir de la substància aplicada, i no pas a partir d'una mutació espontània, que donaria falsos positius quant als objectius de l'experiment), a causa del *valor-p* estàndard, l'experiment conclouria amb 76 rates sanes i 24 amb càncer, per la qual cosa quatre rates cancerígenes esdevindrien falsos positius. En el cas hipotètic de treballar amb 1000 espècimens i que finalment 999 fossin sans i només 1 resultés ser malalt, si hi apliquéssim el màxim *valor-p*, tindríem 50 falsos positius; és a dir, mentre que en el primer cas (població = 100) tindríem un error del 17% de falsos positius respecte de la totalitat obtinguda de rates malaltes, en el segon (població=1000), l'error seria del 5000%.

En segon lloc, durant la controvèrsia al voltant del ciclamat sòdic, un altre edulcorant artificial, el comissari de la FDA, J.E. Goyan¹⁹, argumentà que un *valor-p* de 0'05 no era tan diferent del

de 0'06, i que un de 0'02 seria també vàlid, com afirmà durant la seva acusació de difamació dels Abbot Laboratories, ja que aquests havien criticat els resultats d'un experiment que mostrava una alta incidència de tumors en els animals de laboratori tractats amb ciclamat, els resultats del qual havien estat calculats amb un *valor-p* de 0'02, essent acceptats per la FDA. Amb aquests fets, C.R. Buncher, president del Executive Committee of the Biopharmaceutical Section of the American Statistical Association²⁰, remeté una carta a Goyan, on s'expressava amb rotunditat: «*we are concerned about the extreme misrepresentation of our professional methodology [...] many of the published statements are profoundly fallacious [...] the concept expressed [referent a la interpretació dels valors-p] is foreign to everything that is taught in the statistics profession [...] we strongly encourage you to have the appropriate professionals prepare a new statement that correctly expresses the statistical principles that are involved in this issue. This new statement is needed to avoid the ridicule of knowledgeable scientists [...] WE BELIEVE THAT A REVISION MUST BE PUBLISHED IN THE FEDERAL REGISTER AS A CORRECTION*»²¹.

Quan al començament dels anys setanta l'EPA es trobà davant un altre problema sobre el grau de certesa en una controvèrsia –la de l'herbicide 2,4,5-T–, es definí obertament el següent: «*Petitioners demand sole reliance on scientific facts, on evidence that reputable scientific techniques certify as certain. Typically, a scientist will not so certify evidence, unless the probability of error, by standard statistical measurement, is less than 5%. That is, scientific fact is at least 95% certain [...] Agencies are not limited to scientific fact, to 95% certainties. Rather, they have at least the same fact-finding powers as a jury, particularly when, as here, they are engaged in rule-making*»²². Ara bé, el tant per cent d'evidència necessari o considerat com a mínim a l'hora d'acceptar un fet pot diferir enormement en funció del grup a què ens estiguem adreçant. Aquest és un punt més que cal tenir en compte per a entendre l'evolució de les controvèrsies científiques, un cop inclouen diversos agents.

El problema de l'evidència en les controvèrsies científiques que impliquen aspectes legals ha estat reconegut obertament des del començament del segle xx. L'any 1923, durant el cas *Frye v. United States*²³, ja calgué definir el valor de l'aportació científica durant un judici. Concretament, s'estrenà el test per a detectar l'engany a partir de la pressió sistòlica sanguínia, un precursor de les actuals «màquines de la veritat». La validesa de les conclusions dels experts respecte del valor de la tècnica demostrà aviat la seva polèmica, de manera que, judicialment, calgué definir el valor de l'aportació científica: «*Just when a scientific principle or discovery crosses the line between the experimental and demonstrable sta-*

20. Professor també d'Epidemiologia i Bioestadística a la University of Cincinnati.

21. W. R. HAVENDER: «The science and politics of cyclamate», *The Public Interest*, USA, 1983, n. 71 p. 27, majúscules originals del text.

22. S. JASANOFF: *The Fifth Branch*, USA: Harvard University Press, 1994, p. 51.

23. 54 App. D.C. 46, 47, 293 F. 1013, 1014 (1923).

ges is difficult to define. Somewhere in this twilight zone, the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs» (54 App. D.C., at 46, 293 F., at 1014).

És a dir, un testimoni científic és acceptat quan el seu valor és considerat com a estàndard (o evident) dins la comunitat a què pertany. Fins i tot es crearen posteriorment les *Federal Rules of Evidence*. La idea de l'acceptabilitat general, però, fou considerada posteriorment com a no determinant i el concepte d'evidència prengué nova forma a partir de la regla 702: «*If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training [509 U.S. 579,8] or education, may testify there to in the form of an opinion or otherwise.*»

Aquesta regla fou introduïda en 1975²⁴ a partir dels problemes que havia demostrat l'antiga regla de Frye durant la dècada dels anys seixanta i setanta. La Regla 702²⁵ fou emprada a fons en el polèmic cas *Daubert v. Merrell Dow Pharmaceuticals, Inc.*²⁶, l'any 1993. La polèmica aparegué quan una parella amb dos fills amb greus defectes de naixement demandà l'empresa fabricant de Bendectin, un medicament per a evitar les habituals nàusees de les embarassades i que, suposadament, provocà els terribles efectes teratològics. L'expert de l'empresa era un epidemiòleg, mentre que els de la parella eren investigadors clínics que havien realitzat tests *in vivo*, *in vitro* i estructurals. L'epidemiòleg considerà que no es podia establir una relació causal entre la ingesta de Bendectin i els efectes teratològics, i el tribunal considerà que només l'epidemiologia²⁷ podia establir relacions causals, deixant de banda l'altre tipus d'estudis, els quals mostraven evidències a favor dels efectes teratològics. Un altre criteri del tribunal fou remarcar que en altres casos²⁸ s'havia decidit que fins i tot els estudis epidemiològics no podien ser acceptats si no havien estat publicats abans i sotmesos a *peer review*.

Veiem, doncs, que els requeriments legals implícits en les controvèrsies científiques exigeixen la clarificació d'alguns conceptes, els quals són concebuts de forma diversa pels diversos grups d'agents que hi participen²⁹. D'alguna manera, puc afirmar que el dret s'ha enfrontat, a nivell pràctic i amb major contundència, amb les controvèrsies científiques que no pas amb la filosofia, la història o la sociologia de la ciència. Tanmateix, ha demostrat la riquesa de l'anàlisi del valor del testimoni dels experts: ¿es tracta d'un pacte

24. Durant el cas *Kumho Tire Co., Ltd. v. Patrick Carmichael*, 119 S. Ct. 1167.

25. El mateix any 1975 aparegué la *Federal Rule of Evidence 703* reforçant la nova línia d'evidència científica de la debatuda 702.

26. 509 US 579 (1993).

27. En l'àmbit de l'epidemiologia, disciplina que estudia les causes determinants d'una malaltia i els seus patrons de distribució, l'any 1965 Sir Austin Bradford Hill pronuncià nou criteris per decidir si una associació era més causal (que no ho pot ser mai, per definició) que espúria: (1) *strength of association*, (2) *consistency*, (3) *specificity*, (4) *temporality*, (5) *biological gradient*, (6) *plausibility*, (7) *coherence*, (8) *expertiment*, (9) *analogy*. Una explicació més detallada la trobem en FOSTER, *Ibid.*, p. 9.

28. 951 F. 2d, at 1130-1131.

29. EEA: *Late lessons from early warning: the precautionary principle 1896-2000*, Copenhagen: EEA, 2002, p. 193, recull novament la divergència de l'evidència legal i reguladora: «*One important consequence of acknowledging both scientific uncertainties (including ignorance) and the urgency of hazard reduction in situations of high stakes is the need for agreement on the sufficiency of evidence of harmful effects that is required to justify action. Such "levels of proof" can vary from the "reasonable ground for concern" of the European Commission's Communication on the Precautionary Principle, to the "beyond reasonable doubt of criminal law"*».

social (disciplinar) que té el seu fonament en variables més complexes, com ara l'experiència, el nivell de formació o el reconeixement dins la disciplina? Ara bé, cal recordar que estaríem considerant el judici legal de la polèmica, no pas el procés pròpiament científic de resolució de la controvèrsia, cosa que representa nivells diferents segons el camp d'acció de la controvèrsia.

Queda clar que no hi ha hagut acord a nivell estadístic ni respecte del valor epistèmic d'una tècnica ni respecte de la seva correcta utilització, i encara més si considerem que cadascuna d'elles ha format part d'una bateria experimental amb múltiples protocols curulls d'elements problemàtics. Hi hauria també un posicionament ideològic davant la validesa d'un apropament freqüentista i bayesià als còmput estadístic, i, més i tot, si aquest ha de ser la base per a la gestió pública del risc. Lilford & Braunholz (1996) critiquen els habituals enfocaments freqüentistes i argumenten que aquests no fan més que crear dicotomies respecte dels resultats, en considerar si aquests són significants o no, cosa que dificulta la presa de decisions. Amb un model bayesià, aquest problema sembla que pot ser resolt. Hi hauria una crítica més, a nivell teòric, quant al *valor-p* de 0.05, realitzada pels atacats bayesians. Freedman (1996) apunta que els experiments rarament poden efectuar-se sense la presència d'informació externa, com reclamen els freqüentistes en criticar la posició bayesiana. I, a més, l'elecció de $p < 0.05$ representa la inclusió d'un element subjectiu en la valoració dels resultats experimentals.

5. El dilema sobre l'elecció d'un dels dos models

La major part de filòsofs interessats a trobar una solució al problema de l'elecció d'un model estadístic, sia bayesià o freqüentista, ha apel·lat a criteris epistemològics o ontològics del tipus: «atès que la realitat és d'aquesta manera i la racionalitat científica és d'aquesta altra, l'única elecció possible és x» (digueu-ne «bayesià» o «freqüentista» segons convingui). Ara bé, ¿és correcte reduir el criteri d'elecció entre un d'ambdós models a valors purament epistèmics? ¿Ha estat, per aquest motiu, que especialistes bayesians han esdevingut freqüentistes o bé a la inversa? En tot cas, una constant en les disputes entre bayesians i freqüentistes, i que jo resumiria en la «constant epistèmica ingènua» i que es podria formular de la manera següent: «Tot científic considera que un veritable científic mai no pot pertànyer a l'escola estadística oposada.» En aquesta disputa tothom dóna per suposat que la racionalitat científica és com és, o sia, per a uns, evidentment freqüentista, i per a altres, bayesiana. Si recorrem a altres indicadors, trobem que una recerca del terme «*bayesian*» en el títol i/o resum dels articles citats al

Medline (des de 1966), ofereix en la primera dècada 29 referències del terme en els títols, però cap en els resums, mentre que entre gener de 1995 i octubre de 1999 les citacions en títols són de 235, i 577 en els resums (Silva & Muñoz, 2000). De fet, si hom hi cerca el total d'entrades sobre «*frequentism*» i «*bayesian*», n'obté 108 i 3270, respectivament.

Crec que en el procés d'adopció d'un d'ambdós models hi tenen molt a veure els valors no epistèmics, com ara la simplicitat i la facilitat d'aplicació d'ambdós models, la formació acadèmica rebuda pels especialistes, les tendències sobre els *guidelines* de les revistes acadèmiques o, fins i tot, qüestions metafísiques. Enumerem, doncs, alguns d'aquests aspectes no epistèmics que inclinen la balança dels investigadors cap al suposat dilema epistèmic.

I. Valors metafísics

En aquest punt, hom ha d'admetre que el posicionament sobre l'essència de la bona racionalitat científica es fonamenta en opcions filosòfiques de tipus metafísic que s'escolen de la pròpia avaluació racional, en ser-ne constitutives. Com pensem i com hauríem de pensar per a entendre millor el món (i quina seria la seva naturalesa) són preguntes que es remeten als orígens del propi pensament, a la reflexió sobre la filosofia primera, la metafísica. Així, segons Hacking (1972), p. 133: «*Euler at once retorted that this advice is metaphysical, not mathematical. Quite so! The choice of primitive concepts of inference is a matter of 'metaphysics'. The orthodox statistician has made one metaphysical choice and the Bayesian another*».

El criteri fonamental a l'hora de l'elecció entre un dels dos models fa referència a posicionaments metafísics o ontològics, és a dir, als principis de la realitat³⁰. La tria mateixa d'un dels models respon a una concepció prèvia sobre l'estatut de l'observador respecte del món. En aquest apartat, tindria sentit plantejar-se el grau i la qualitat de realisme o subjectivisme que defensarien els investigadors implicats.

II. Valors formacionals

Si bé se suposa que l'elecció entre un dels dos models remet a criteris purament racionals, també és cert que hi ha altres factors que hem de tenir en compte, entre els quals es troba el formacionals³¹. Bland & Altman (1998), p. 1160, ens indiquen: «*Most statisticians have become Bayesians or Frequentists as a result of their choice of university. They did not know that Bayesians and frequentists existed until it was too late and the choice had been made. There have been subsequent conversions. Some who were taught the Bayesian way discovered that when they had huge*

30. Respecte de la relació entre l'ontologia del món i la constitució lògica del nostre saber, em remeto a l'excel·lent obra de l'Estagirita, la *Metafísica*, llibre G, on descriu la metafísica o «ciència primera» com la «*episteme peri ton ontos on*». Segons Aristòtil, podem crear un discurs sobre el món, puix que el món té unes característiques que són susceptibles de ser conegudes i que determinen el nostre saber.

31. He introduït aquest concepte en un treball anterior: J. VALLVERDÚ: *Marc teòric de les controvèrsies científiques: el cas de la Sacarina*, Bellaterra: UAB [Tesi doctoral], 2002, p. 309. Allí explicava que els arguments formacionals són aquells que remeten a la *formació* (acadèmica, tècnica) i *experiència* dels científics implicats en un procés de debat per a convèncer altres participants del discurs racional. Ara bé, aquests arguments (en realitat valors tàctics de tipus no epistèmic) no incorren en les fal·làcies *ad verecundiam/ad hominem* sinó que formen part del substrat racional dels implicats.

quantities of medical data to analyse the frequentist approach was much quicker and more practical, although they remain Bayesian at heart. Some frequentists have had Damascus road conversions to the Bayesian view. Many practising staticians, however, are fairly ignorant of the methods used by the rival camp and too busy to have time to find out.»

El que en aquest paràgraf Bland & Altman ens diuen ens pot semblar un argument dèbil, tot i que una lectura atenta a la història de la ciència ens ofereix diversos exemples de preferències epistèmiques estretament relacionades amb factors nacionals: Petersen & Markle³², per exemple, han remarcat les divergències existents entre polítiques diferents respecte de la teràpia de reemplaçament d'estrògens als Estats Units i a la Gran Bretanya. L'obra general de Sheila Jasanoff³³ segueix aquesta línia de pensament i emfatitza una contraposició entre el model nord-americà i l'europeu. Atès que la ciència reguladora dels diversos països implicats s'abeura de tradicions culturals i polítiques divergents, el seu funcionament es trobarà condicionat i determinat per aquests aspectes. Sense caure en l'exageració simplista, podem afirmar l'existència d'estils nacionals diversos pel que fa al disseny de les polítiques tecnològica i científica o a la regulació governamental³⁴. En aquesta línia, és com podem explicar la dificultat no solament política sinó científica a l'hora de trobar un acord entre investigadors i instituts de recerca pertanyents a tradicions culturals i geogràfiques diferents. Altres autors, com Levidow, han adoptat un posicionament marxista per a entendre la presa de decisions científiques com a integrades completament en l'estructura social en què es manifesta. Una de les manifestacions d'aquesta ciència, a partir del cas de l'accident nuclear de Three Mile Island, serà l'aparició del que Levidow ha denominat la «*ideology of safe levels*», idea³⁵ que és comuna a més investigadors d'àrees de recerca diverses. Amb aquesta idea, s'entén el grau de manipulabilitat de què poden ser objecte els models teòrics dissenyats per a determinar el nivell màxim de tolerància a una substància, a partir de criteris científics que divergeixen en funció de la ubicació dels investigadors (estats-reguladors, privats, d'ONG's) o de la seva ubicació espacial (a nivell nacional o internacional). Abraham & Millstone (1989), p. 52 ho resumeixen d'aquesta manera: «*In Britain the government receives toxicological advice from one and only one agency, while in France, Germany and the USA governments receive advice from two or more independent agencies*». A Espanya, és el CAE (Codi Alimentari Espanyol) qui s'encarrega de regular tot allò referent als additius alimentaris, mentre que, al Regne Unit, ho fa el FAC (Food Advisory Committee). A França, se n'ocupa el CSdH (Conseil Supérieur d'Hygiène Publique de France) i

32. J. C. PETERSEN, G. F. MARKLE: «Controversies in Science and Technology», dins D.E. CHUBIN, (et al eds) *Science off the Pedestal: Social Perspectives on Science and Technology*, USA: Wadsworth, 1989, p. 16.

38. Sheila JASANOFF: *Controlling Chemicals: The Politics of Regulation in Europe and the US*, USA: Cornell, 1985, analitza les divergències entre els contextos científics europeu i nord-americà.

34. El terme «*national styles*» l'he extret del capítol 25, «Changing Policy Agendas in Science and Technology», escrit per Elzinga i Jamison, i pertanyent a S. JASANOFF [ed.]: *Handbook of Science and Technology Studies*, UK: Sage Publications, 1995, p. 576.

35. L. LEVIDOW: «Three Mile Island: The Ideology of Safe Level as a Material Force», dins *Radical Science Journal*, 1979, p. 83.

el AdM (Académie de Médecine), i, a Alemanya, el BGA (Bundesgesundheitsamt) i el DFG (Deutsche Forschungsgemeinschaft). Als EUA, conflueixen resultats de la FDA, FASEB, JEFCA i NCI³⁶. Però trobem diferències nacionals encara més pronunciades: no solament pel que fa als sistemes generals de regulació científica sinó pel que fa al valor atorgat a determinats dissenys experimentals. Cada país manifesta encara avui una preferència vers els resultats experimentals de determinades disciplines (Proctor, 1995). Sheila Jasanoff (1991b, p. 110) ho ha expressat d'una manera exacta: «*Political and legal cultures also condition the interpretation of science. Controversies over policy relevant science arise more frequently in some countries than in others. The adversarial political culture of the United States is known to provide exceptionally fertile ground for nurturing such disputes.*»

36. Segons J. ABRAHAM, E. MILLS-TONE: «Food Additive Controls: Some International Comparisons», *Food Policy*, UK, 1989, pp. 47-53; D. VOGEL: «Consumer Protection and Protectionism in Japan», *Journal of Japanese Studies*, 2002, n.18(1), Japó, p. 122.

37. Segons R. BRICKMAN: *Controlling Chemicals. The Politics of Regulation in Europe and the United States*, USA: Cornell University Press, 1985.

38. Si bé la major part dels països consideren que la mutagenicitat equival a la carcinogenicitat, les desavinences apareixen quan cal indicar quin nombre d'aquest tipus d'experiments cal realitzar per a obtenir-ne una evidència. Europa i els Estats Units manifesten una divergència considerable. Fins l'any 1985, segons Brickman *et al.* (1985), els Estats Units en requerien 12 per a minimitzar els falsos positius i negatius, seguint tres fases (sistemes bacterials, cèl·lules de mamífers *in vitro* i sistemes *in vivo* d'animals), mentre que a Europa se seguia el «*step-sequence approach*», amb dos tests inicials de base (un bacteriològic i un altre no) i dos tests verificacionals. Per a aquests autors, «*the EC's decision to prescribe only two mutagenicity tests for the "base set" for example, represented a victory for cost-cutting arguments over proposals to minimize the number of false negatives*», p. 196. Qüestions pressupostàries poden modificar el tipus de ciència i regulació final d'una contrada.

I podríem arribar encara més lluny en afirmar que, en l'elecció del model causal per al càncer, les agències reguladores opten per solucions diverses: als EUA s'escullen models lineals, mentre que a Europa es tendeix als models llindar (*threshold*). Això ens condueix al que Jasanoff (1991b) ha denominat la «*Paradox of Risk Assessment*», la qual consta de dos punts bàsics: (1) les diferències legals, culturals i institucionals de cada Estat provoquen l'existència de diferències substancials en el tipus i els resultats quant a l'assessorament de llurs riscos; (2) paradoxalment, les decisions a nivell de la gestió dels riscos manifesta una major homogeneïtat a nivell transnacional que no en l'assessorament de riscos. Això ha estat demostrat en el capítol 3r, on s'explicitava l'efecte dòmino dels països europeus en la prohibició del ciclamat i la sacarina a partir de la decisió nord-americana. Malgrat les diverses tradicions en l'AR, la gestió del risc tendeix a unificar-se, influïda també per criteris econòmics, polítics, burocràtics...

A més, es pot afirmar que les disciplines implícites en la recerca de l'anàlisi de riscos són valorades i finançades molt diversament en funció de la nació que estudiem. Així, veiem com, a Gran Bretanya, els resultats epidemiològics són considerats de gran importància, mentre que, a França i a Alemanya, aquests no gaudeixen ni de prestigi ni de subvencions tan importants³⁷. Però tampoc no és que existeixi un acord respecte dels tests de curta durada³⁸ o dels bioassaigs animals. Jasanoff (1991b) realitza un estudi comparatiu sobre el control dels productes químics segons les diverses polítiques nacionals, que manifesta aquestes divergències nacionals que acaben formant part de la formació acadèmica dels nous investigadors. En resum, podem veure com les divergències no solament es produeixen únicament a nivell general entre nacions diverses sinó entre gestors, experts i grups d'interès dins una mateixa nació.

III. Valors infraestructurals

Un dels motius pels quals avui el bayesianisme està experimentant una àmplia difusió és fonamentalment el desenvolupament de programari específic que facilita enormement l'aplicació de models bayesians a la recerca pràctica científica. En els darrers anys ha aparegut gran quantitat de programes informàtics que permeten l'ús dels complexos models bayesians, en un entorn de treball molt simple. Com a exemples, ens trobem amb els programes BACC, [B/D], BOA, BUGS (i WinBugs), MINITAB, EPIDAT, FIRST BAYES, HYDRA,... (Silva Ayçaguer i Suárez Gil, 2000); és a dir, que sense l'existència d'aquest programari, la implementació del bayesianisme fóra molt més complicada i dificultosa, per la qual cosa rebria una menor atenció per part dels investigadors. La simplicitat és un dels valors, juntament amb la comoditat, que cal tenir en compte en la pràctica científica.

IV. Valors relatius a la difusió

El cim de la dinàmica científica passa per la difusió de la recerca a través dels canals clàssics, és a dir, mitjançant congressos o premsa científica. En aquest procés de revisió o *peer review*, la informació és contrastada i criticada per tal d'avaluar-ne el valor i la validesa. Així, hem d'admetre que les normes quant a la publicació (les *guidelines*) conformen el marc epistèmic en què es pot desenvolupar el coneixement. Les normes exigides determinen la forma segons la qual³⁹ s'obté la informació i, al mateix temps, segons la qual aquesta és acceptada. L'exigència de metodologies llunyanes del freqüentisme, per exemple, reforçarà la difusió del bayesianisme. Aquest és el cas del document generat per l'International Committee of Medical Journal Editors, els quals redactaren l'*Uniform Requirements for Manuscripts Submitted to Biomedical Journals*, document disponible a Internet (<http://www.icmje.org>). En la secció destinada a l'aparell estadístic, l'informe afirma: «*Avoid relying solely on statistical hypothesis testing, such as the use of P values, which fails to convey important quantitative information.*» Ja hem vist en el punt II que els *valors-p* havien demostrat la seva conflictivitat en diversos moments de controvèrsia científica i que, per això i pel fet de constituir l'eix del freqüentisme contemporani, debilitaren el model freqüentista i la credibilitat general dels estadístics en els processos d'avaluació de riscos.

6. Conclusions

El debat entre bayesians i freqüentistes no pot ser reduït a una discussió merament epistèmica. A més, encara que fos així, caldria ampliar els valors implícits en el concepte «epistèmic». La reducció

39. Aquest és un material *online* redactat l'any 2001 i fou consultat per l'autor l'11 d'agost del 2003.

del problema a una formulació de context acadèmic merament metateòric, s'allunya del seu context d'ús real: la pràctica real científica implicada en el càlcul o el processament estadístic de les dades experimentals. Mentre els filòsofs discutim acaloradament sobre la validesa d'un dels dos models respecte de l'altre, els científics operen amb ambdós, la qual cosa redueix la resposta o la clausura de la controvèrsia⁴⁰ a la millor operativitat i a resultats més precisos; és a dir, es tracta d'una clausura gradual de controvèrsia que remet a la millor resolució dels problemes en els quals participen les metodologies freqüentistes o bayesianes. La pràctica científica, i no pas la teoria filosòfica, dirigeix actualment el debat. I si el bayesianisme està guanyant terreny dia rere dia, hem pogut veure que no és a causa del convenciment purament racional dels científics implicats sinó a causa de valors no epistèmics prou importants. L'axiologia de la pràctica científica requereix una ampliació de valors, entre els quals cal incloure el de la practicitat. Si el càlcul infinitesimal newtonià s'imposà al leibnizià, això fou possible gràcies a la seva major simplicitat. La simplificació de la metodologia possibilita que aquesta sigui capaç de treballar amb més informació per un cost humà molt més baix. I avui la capacitat de treball amb la quantitat ingent d'informació de què disposem és l'objectiu proritari dels investigadors científics. La informació és el valor màxim, però la seva síntesi és la clau de tot el procés. La polèmica entre bayesians i freqüentistes troba la seva conclusió paulatina a la llum de la pràctica, lluny de les ments isolades i llurs artificis lingüístics.

40. Quant als estudis sobre les clausures de controvèrsies científiques, em remeto a l'obra clau H. ENGELHARDT, JR H. TRISTAM [ed.]: *Scientific Controversies. Case studies in the resolution and closure of disputes in science and technology*, USA: Cambridge University Press, 1987, dins la qual trobem els interessants capítols de McMullin, Beauchamp i el propi Engelhardt

DR. JORDI VALLVERDÚ I SEGURA

Universitat Autònoma de Barcelona
[jordi.vallverdu@uab.es]

BIBLIOGRAFIA

- J. M. BERNARDO, A. F. M. SMITH: *Bayesian Theory*, Wiley: New York, 1994.
- V. CHANKONG (et al.): «The Carcinogenicity prediction and battery selection (CPBS) method: a Bayesian approach», *Mutation Research*, USA, 1985, n. 153, pp. 135-166.
- L. FREEDMAN: «Bayesian statistical methods», Editorial dins *British Medical Journal*, 1996, n. 313, pp. 569-570.
- R. N. GIERE: *Understanding Scientific Reasoning*, USA: Holt, Rinehart and Winston, 1991.
- I. HACKING: «Likelihood», *British Journal of Philosophy of Science*, UK, 1972, n.23, pp. 132-137.
- I. HACKING: *La domesticación del azar*, Barcelona: Gedisa, 1995.

A. HALD: *A History of Probability and Statistics and Their Applications before 1750*, New York: John Wiley & Sons, 1988.

C. HOWSON: «Error Probabilities in Error», *Philosophy of Science*, 64 Proceedings, USA, 1997, pp. 185-194.

C. HOWSON, URBACH: *Scientific Reasoning: The Bayesian Approach*, Chicago: Open Court, 1989. S. JASANOFF: «Cross-National Differences in Policy Implementation», *Evaluation Review*, USA, 1991b, n. 15(1), pp. 103-119.

L. LAUDAN: *Progress and its problems*, California: The University of California Press, 1978.

G. D. MAYO: *Error and the growth of experimental knowledge*, USA: The University of Chicago Press, 1996.

G. D. MAYO: «Error Statistics and Learning From Error: Making a Virtue of Necessity», *Philosophy of Science (Proceedings)*, USA, 1997a, n. 64, pp. 195-212.

G. D. MAYO: «Duhem's Problem, the Bayesian Way, and Error Statistics, or 'What's Belief Got to Do with It?», *Philosophy of Science*, USA, 1997b, n. 64, pp. 222-244.

NRC: *Evaluation of Cyclamate for Carcinogenicity*, USA: NRC Staff, 1985.

K. POPPER: *Un mundo de propensiones*, Madrid: Tecnos, 1992.

R. N. PROCTOR: *Cancer Wars. How Politics Shapes What we Know & Don't Know About Cancer*, USA: BasicBooks, 1995.

L. C. SILVA AYÇAGUER, A. MUÑOZ VILLEGAS: «Debate sobre métodos frecuentistas vs Bayesianos», *Gaceta sanitaria*, 2000, n. 14 (6), pp. 482-494.

L. C. SILVA AYÇAGUER, A. MUÑOZ VILLEGAS: «¿Qué es la inferencia bayesiana?», *JANO*, 2000, n. 58 (1338), pp. 65-66.

P. SUPPES: *A Probabilistic Theory of Causality*, Helsinki: North-Holland Publishing Company, 1970.