TEORIA BAYESIANA DELLA DECISIONE E RAGIONEVOLE DURATA DEL PROCESSO

di *Alberto Mura*

SOMMARIO: 1. Introduzione. — 2. La teoria bayesiana della decisione. — 3. La decisione giudiziale. — 4. Il valore delle fonti di prova. — 5. Non monotonia e durata del processo. — 6. La solidità pragmatica dei valori di probabilità. — 7. La durata del processo come "costo". — 8. Conclusioni.

1. *Introduzione*. — Gli sviluppi della logica deduttiva moderna (calcolo dei predicati, logica modale, logica deontica etc.) hanno arricchito in maniera significativa l'attrezzatura logica a disposizione del diritto. Tale allargamento non ha però consentito in maniera adeguata di fare i conti con un aspetto che la riflessione epistemologica e giuridica contemporanea ha mostrato essere essenziale alla teoria della prova: l'incertezza. Disporre di una logica del ragionamento in condizioni di incertezza costituisce un ulteriore arricchimento del diritto delle prove. Il bayesianesimo giuridico (1) ha il merito di aver mostrato come la dottrina delle prove e la stessa giurisprudenza possano avvalersi della nozione di probabilità matematica. Secondo il punto di vista bayesiano l'aspetto saliente della probabilità è che essa costituisce un'estensione, sino a includervi le credenze parziali, della logica deduttiva applicata alla credenza.

Tuttavia la probabilità bayesiana è parte di una più generale teoria della decisione in condizioni di incertezza e da essa non può essere separata. Viceversa il bayesianesimo giuridico, nella forma sviluppatasi negli USA a partire dagli anni '70 del secolo passato (che chiamerò *bayesianesimo* giuridico *tradizionale*), tipicamente rinuncia a un'applicazione dell'intero formalismo bayesiano, limitandosi alla considerazione delle sole probabilità. Il bayesianesimo tradizionale trae dal diritto positivo le regole per la decisione giudiziale. L'uso della nozione di probabilità è pertanto da esso condotto attraverso una riformulazione in termini probabilistici delle regole procedurali di decisione. Ad esempio il principio — tipico del diritto delle prove della *common law*, ma sostanzialmente operante anche nell'ordinamento italiano — secondo il quale la prova concludente di colpevolezza è raggiunta solo quando si è stabilito "oltre ogni ragionevole dubbio" che l'imputato ha commesso il reato di cui è accusato, è interpretato come il principio secondo il quale per la condanna dell'imputato è necessaria una probabilità di colpevolezza molto vicina alla certezza. Ma quale sia il fondamento del principio stesso è questione che il bayesianesimo tradizionale considera non pertinente al compito che la dottrina bayesiana delle prove si prefigge.

In questo lavoro cercherò in primo luogo di spiegare in che cosa consiste la teoria bayesiana della decisione e in secondo luogo cercherò di mostrare in che modo essa può riuscire utile all'interno del diritto delle prove. In primo luogo argomenterò che la teoria della decisione è in grado di gettare nuova luce sul fondamento di alcuni dei principi fondamentali come quello della "probabilità oltre ogni ragionevole dubbio" che il bayesianesimo giuridico tradizionale dà semplicemente per presupposto. In secondo luogo sosterrò che la regola della decisione giudiziale richiede un secondo parametro oltre all'alta probabilità del *thema probandum*, parametro che sarà

(1) Per una panoramica del bayesianesimo giuridico si veda AA.VV., *L'inferenza probabilistica nel diritto delle prove. Usi e limiti del bayesianesimo* (1988), a cura di TILLERS–GREEN, trad. it., Giuffrè, 2003.

chiamato “solidità probatoria” (2). Caratterizzerò questo secondo parametro alla luce della teoria bayesiana della decisione. Cercherò infine di mostrare come l’introduzione della nozione di solidità probatoria, resa possibile dall’estensione del bayesianesimo giuridico alla teoria bayesiana della decisione, consente agevolmente di affrontare tematiche che non erano alla portata della teoria bayesiana giuridica tradizionale e mostrerò ciò con un esempio importante: il principio della “ragionevole durata del processo”.

2. *La teoria bayesiana della decisione*. — Per comprendere che cosa sia e come vada usata la teoria bayesiana della decisione è bene partire con un’esemplificazione concreta. Supponiamo che una persona *X* abbia avuto un piccolo incidente stradale nel quale la sua vettura abbia colliso con quella di un altro automobilista *Y*. Per semplificare la situazione supponiamo che la vettura di *X* abbia subito un danno di € 600, mentre l’altra vettura non abbia subito danno alcuno. Supponiamo che *X* abbia il forte sospetto che la responsabilità civile dell’incidente sia da attribuire a *Y*, sì che egli avrebbe diritto a essere da questi risarcito, mentre *Y* ritiene, per parte sua, di non aver alcuna responsabilità dell’accaduto e quindi di non dover risarcire *X* di alcun danno. Il problema che si pone per *X* è se valga per lui (o per lei) la pena di avviare un’azione legale per ottenere il risarcimento.

In questa situazione le opzioni che *X* ha davanti a sé sono due: (a) avviare l’azione legale oppure (b) rinunciarvi. I *pro* e i *contra* di ciascuna opzione dipendono da una circostanza sulla quale *X* non può influire, ma su cui può avere aspettative soggettive: la decisione del giudice nel caso di azione legale (3). Per semplicità, in un caso come questo, possiamo fare coincidere i *pro* e *contra* rispettivamente con gli importi monetari (espressi in euro) dei guadagni e dei costi che, per semplicità, supporremo coincidenti con il *valore soggettivo* dei medesimi (4). Se *X* decide di fare causa fa, di fatto, una *scommessa*, che è vinta o persa a seconda che si sia vinta o persa la causa. Se *X* vince la causa guadagna € 600 e se perde la causa ha un *guadagno negativo* (cioè una perdita) uguale al costo delle spese legali, che possiamo supporre pari a € – 150. È chiaro invece che se *X* rinuncia all’azione legale non perde niente rispetto allo *status quo*, sì che questo atto ha un valore *certamente* pari allo status quo — che possiamo convenzionalmente fissare a 0 (5).

Per prendere la decisione migliore tra le due che sono possibili in questa situazione (fare causa e rinunciare alla causa), la persona *X* deve confrontare il valore delle due opzioni e scegliere quella che ha per lei il valore maggiore. In questo caso si tratta, sostanzialmente, di stabilire se la decisione di fare causa, cioè di scommettere, è preferibile o no rispetto allo *status quo* (cioè alla decisione di non intentare la causa). Per stabilire ciò non basta la considerazione dei *pro* e dei *contra* che si avrebbero, ricorrendo all’azione legale, nelle due circostanze possibili (ottenere il risarcimento oppure sopportare il costo delle spese legali senza risarcimento alcuno): è necessario integrare questi dati con le *probabilità soggettive* che si attribuiscono ai due esiti possibili della causa.

---

(2) Vi sono, in letteratura, numerosi tentativi di caratterizzare questa solidità probatoria (o peso delle prove) mediante formalismi che sono in genere incompatibili con l’impostazione bayesiana. La presente analisi dimostra che non c’è bisogno di alcun nuovo formalismo, giacché quello bayesiano appare perfettamente in grado di catturare la nozione di solidità probatoria.

(3) In realtà vi sono molti altri fattori, perlopiù incerti, da cui può dipendere la decisione di *X* (durata del processo, stress dell’azione penale, atteggiamento della controparte ecc.). Per semplicità supponiamo che *X* prenda in considerazione solo la decisione del giudice. La considerazione degli altri fattori, sebbene complicherebbe di molto i calcoli, non presenta alcuna difficoltà di principio.

(4) L’assunzione secondo la quale i valori soggettivi sono proporzionali ai valori monetari è in genere solo approssimativamente vera. Tenere conto della maniera con la quale i valori soggettivi possono dipendere dai valori monetari costituirebbe, nel presente contesto, solo un’inutile complicazione.

(5) Attribuire valore 0 allo *status quo* è una convenzione in quanto la decisione migliore resterebbe tale se si cambiasse la scala dei valori, moltiplicando ogni valore considerato per un numero positivo α e aggiungendo al risultato un secondo numero k (in maniera tale che ogni valore r sarebbe trasformato nel valore $(\alpha \times r) + k$).

È questa l'idea centrale della teoria della decisione in condizioni d'incertezza. *X* dovrà quindi valutare due valori di probabilità (che corrispondono ad altrettanti "gradi di credenza"), sotto l'ipotesi che decidesse di ricorrere alle vie legali: (i) la probabilità *p* che la causa sarebbe vinta e la probabilità *q* che essa sarebbe persa. Le leggi della probabilità (oltre che la semplice intuizione) impongono una relazione tra *p* e *q*: tanto più grande è *p* tanto più piccola è *q* e viceversa. Infatti il cosiddetto *principio delle probabilità totali* comporta l'equazione $p+q=1$, da cui discende, essendo *p* e *q* numeri compresi tra 0 e 1, che $p=1-q$ e $q=1-p$. Supponiamo che *X* attribuisca a *p* il valore ⅓ e quindi attribuisca a *q* il valore residuo ⅔. Domandiamoci: come deve fare *X* per soppesare i pro e i contra della decisione d'intentare la causa in maniera tale da tenere conto anche della probabilità che tali pro e contra hanno di verificarsi? La risposta fornita dalla teoria bayesiana è molto semplice: *si moltiplica ciascun valore per la rispettiva probabilità e si fa la somma dei prodotti così ottenuti*. Il numero così calcolato è chiamato *previsione del valore* (o anche *valore atteso*). Nel nostro caso tale valore in euro si ottiene con la seguente identità: $600\times\frac{1}{3}+(-150)\times\frac{2}{3}=100$. Quindi il valore atteso della decisione di avviare l'azione legale è pari a € 100; se si rinunciasse all'azione legale il valore atteso sarebbe invece pari a 0 e quindi minore: con questi dati la teoria bayesiana impone pertanto come scelta preferibile quella di fare causa.

Va tuttavia rimarcato che la regola che abbiamo esposto dipende completamente da valutazioni *di merito* (di valore e di probabilità) e che valutazioni di merito diverse porterebbero a decisioni differenti. La teoria non mira quindi a *prescrivere* che cosa si deve fare, non mira a impartirci *imperativi*, ma piuttosto mette a disposizione l'attrezzatura concettuale atta a (*a*) separare i vari fattori di merito coinvolti nel processo di deliberazione, (*b*) consentire la valutazione separata di ciascuno di tali fattori di merito e (*c*) ricombinare in un unico indice numerico (la previsione del valore soggettivo) tali valutazioni separate, tale da fornire una misura complessiva della desiderabilità delle diverse linee di condotta, sì che la scelta di un atto che dia luogo al valore più grande tra quelli possibili di tale indice (6) costituisce *il fine ultimo* della deliberazione (7).

Sarebbe perciò vano chiedere alla teoria bayesiana di dire da sola se la scelta di fare causa sia "logicamente" preferibile a quella di rinunciare al risarcimento. La risposta che essa fornisce dipende essenzialmente dalle valutazioni di merito di *X*. Se *X* fosse una persona dalla personalità fragile e mettesse nel conto anche il senso di acuta frustrazione che gli deriverebbe dal perdere la causa, dovrebbe correggere il valore -150 dell'eventualità di perdere la causa attribuendogli, per esempio, il valore -600, comprensivo anche del risarcimento che *X* riterrebbe adeguato per compensare tale senso di frustrazione. La regola bayesiana fornirebbe in tal caso il valore -200, raccomandando quindi a *X* di non intraprendere la causa. Analogamente se *X* attribuisse valore $1/6$ alla probabilità *p* e valore $5/6$ alla probabilità *q*, la regola fornirebbe il valore -25, invitando di nuovo *X* a non intentare causa. Da ultimo la responsabilità di decidere non può essere "delegata" alla teoria bayesiana, la quale si limita a fornire al decisore gli strumenti analitici per tener conto dei vari fattori di merito da cui la decisione dipende.

---

(6) Vi sono casi in cui il valore più grande dell'indice di desiderabilità non esiste, sebbene tipicamente essi si riferiscano a situazioni puramente ideali di nessuna rilevanza pratica, in cui sono essenzialmente coinvolte infinite alternative di condotta. Se, ad esempio, un essere infinitamente ricco si dichiarasse disposto a regalare qualunque importo finito gli fosse richiesto, non sarebbe possibile indicare il valore "più grande", perché per qualunque importo ne esisterebbe un altro maggiore.

(7) Vi sono situazioni nelle quali sembra *prima facie* che la teoria imponga una determinata linea di condotta. Secondo chi scrive tale prescrittività è solo apparente e dipende in realtà dall'incapacità della teoria di tenere in conto certi generi di preferenza. Tali fattori non sembra siano importanti nel contesto giudiziale. Per maggiori dettagli sull' interpretazione non prescrittiva della teoria bayesiana della decisione si veda MURA, *Il carattere regolativo della teoria delle decisioni*, in *Economia delle scelte pubbliche. Journal of Public Finance and Public Choice*, 1987, p. 177–193.

3. *La decisione giudiziale*. — Consideriamo ora la decisione giudiziale penale secondo il formalismo bayesiano. Fissiamo la nostra attenzione su una situazione molto semplice. Supponiamo che un singolo imputato *X* sia accusato di aver commesso un determinato reato e che, al termine del dibattimento, il giudice debba decidere se *X* sia o no colpevole di aver commesso quel reato (8). Il giudice deve valutare per ciascuna delle due decisioni a sua disposizione (assoluzione e condanna) il *valore* delle rispettive conseguenze nei due casi possibili: (a) l'imputato è colpevole e (b) l'imputato è innocente. Ciascuna conseguenza dipende sia dalla decisione adottata (assoluzione o condanna) sia dalla circostanza che l'imputato sia colpevole o no. Vi sono quindi quattro conseguenze possibili: (1) assoluzione essendo l'imputato colpevole, (2) assoluzione essendo l'imputato innocente, (3) condanna essendo l'imputato colpevole e (4) condanna essendo l'imputato innocente. Siano v_1, v_2 e v_3 e v_4 i rispettivi valori di tali conseguenze. La seguente tabella riassume la situazione:

Tabella 1

	L'imputato è colpevole	L'imputato è innocente
Assoluzione	v_1	v_2
Condanna	v_3	v_4

I valori v_1 e v_4 riguardano i due casi di *errore giudiziario* (valutato *ex post*), rispettivamente di *assoluzione di un colpevole* e di *condanna di un innocente*. I valori v_2 e v_3 riguardano invece i casi di decisione corretta (sempre secondo una valutazione *ex post*), rispettivamente di *condanna di un colpevole* e d'*assoluzione d'un innocente*. Questi valori non vanno qui visti, naturalmente, come il riflesso di *giudizi di gusto* del giudice, bensì come valori in linea di principio desumibili dall'ordinamento, che io chiamerò *valori giudiziali*. Il giudice deve, infatti, rendersi interprete delle preferenze insite nella legge (9). Sotto questo profilo la decisione giudiziale differisce sostanzialmente da quella della decisione individuale che abbiamo visto nella sezione precedente: la decisione del giudice è sempre fatta da un punto di vista *disinteressato*, quindi secondo valori che idealmente sarebbero identici per qualsiasi altro giudice dovesse decidere al suo posto.

In secondo luogo il giudice deve valutare le probabilità di ogni conseguenza possibile di ciascuna linea di condotta. Accanto alla tabella dei valori possiamo avere quindi una parallela tabella delle probabilità:

Tabella 2

	L'imputato è colpevole	L'imputato è innocente
Assoluzione	p_1	p_2
Condanna	p_3	p_4

In generale può accadere che la probabilità di ciascuno dei due eventi corrispondenti alle due colonne della tabella dipenda dalla decisione adottata (rappresentata da una delle righe della tabella)

(8) Escludiamo così tutti i sottocasi che complicano la casistica reale e che rendono assai più numerose le opzioni a disposizione del giudice.

(9) Naturalmente sarebbe impossibile desumere valori *puntuali* dall'ordinamento. Tuttavia, come si vedrà, alcune equazioni e disequazioni, che stabiliscono relazioni puramente qualitative, sono perfettamente argomentabili con considerazioni dottrinali e sono sufficienti per trarre significative conclusioni.

(10). Nel nostro caso, tuttavia, una siffatta influenza della decisione del giudice sulla probabilità di colpevolezza o innocenza dell'imputato deve essere ovviamente esclusa, perciò i valori di probabilità sono soltanto due: l'uno corrispondente alla probabilità che l'imputato sia colpevole, l'altro corrispondente alla probabilità che l'imputato sia innocente. Denotiamo con le lettere p e q, rispettivamente, queste probabilità. La tabella 2 può pertanto essere riscritta così:

Tabella 3

	L'imputato è colpevole	L'imputato è innocente
Assoluzione	p	q
Condanna	p	q

Il primo passo nel calcolo bayesiano consiste nel moltiplicare il valore contenuto in ogni casella della tabella 1 con il numero contenuto nella cella corrispondente della tabella 3. Possiamo riassumere la situazione mediante la seguente tabella:

Tabella 4

	L'imputato è colpevole	L'imputato è innocente
Assoluzione	$v_1 \times p$	$v_2 \times q$
Condanna	$v_3 \times p$	$v_4 \times q$

Con le moltiplicazioni della tabella 4 si "ponderano" i valori delle diverse possibili conseguenze delle due decisioni con le rispettive probabilità che hanno di verificarsi. Il passo successivo consiste nel sommare i valori della tabella 4 che stanno nella medesima riga, sì da ottenere, per ciascuna delle due decisioni possibili, un unico indice numerico, il quale dipende sia dai valori delle conseguenze sia dalle rispettive probabilità. Si ottengono così due numeri, secondo le formule della seguente tabella:

Tabella 5

	Previsione del valore
Assoluzione	$(v_1 \times p) + (v_2 \times q)$
Condanna	$(v_3 \times p) + (v_4 \times q)$

L'ultimo passo consiste nel confrontare il valore numerico risultante dalla formula della prima riga della tabella 5 (previsione del valore della assoluzione) con quello della seconda riga della medesima tabella (previsione del valore della condanna): l'azione preferibile è quella che ha il valore più grande.

È facile mostrare come questa semplice impostazione consente un'elementare spiegazione del principio secondo il quale la decisione di condanna può essere la scelta migliore solo se la probabilità di colpevolezza è molto elevata. La chiave sta nella scelta dei valori v_1, v_2, v_3 e v_4. Sembra naturale assumere che il giudice attribuirebbe alle circostanze che danno luogo a decisioni corrette il medesimo valore v (la qual cosa implicherebbe la relazione $v = v_2 = v_3$) e anche che egli

(10) Si pensi a uno studente che si trovi a decidere se trascorrere la vigilia di un esame ripassando i testi di studio oppure a passeggio con gli amici. In questo caso le circostanze rilevanti sarebbero relative all'esito della prova (se allo studente importasse solo superare l'esame sarebbero "promosso" e "respinto"). In una tale situazione la probabilità di ciascuna delle circostanze potrebbe dipendere dalla decisione dello studente.

attribuirebbe alle circostanze che danno luogo a decisioni errate un valore *minore* di v (sì che sarebbe $v_1 < v$ e $v_4 < v$).

Ora, se il giudice attribuisse anche a queste ultime risultanze il medesimo valore giudiziale, sì che valesse anche $v_1 = v_4$, allora basterebbe che *p*, cioè la probabilità di colpevolezza, fosse anche di un'inezia maggiore di ½ (cioè del 50%) per imporre al giudice, secondo la teoria bayesiana, la condanna dell'imputato. Infatti, come mostrato nella tabella 5, la previsione del valore giudiziale della decisione di condannare l'imputato sarebbe data dalla formula $v_3 \times p + v_4 \times q$ che, in virtù delle assunzioni $v_2 = v_3$ e $v_1 = v_4$ sarebbe pari a $v_2 \times p + v_1 \times q$, mentre la previsione del valore giudiziale della decisione di assolvere l'imputato sarebbe pari a $v_1 \times p + v_2 \times q$. Poiché si è supposto v_1 minore di v_2, se *p* è maggiore di *q* anche $v_2 \times p + v_1 \times q$ è maggiore di $v_1 \times p + v_2 \times q$, dato che la differenza tra la prima e la seconda quantità sarebbe pari (come un facile calcolo mostra) a $(p - q)(v_2 - v_1)$ che, essendo il prodotto di due numeri maggiori di 0, sarebbe a sua volta maggiore di 0. Quindi, sotto queste ipotesi, il giudice dovrebbe condannare l'imputato se *p* fosse maggiore di *q* (o, il che è equivalente, se *p* fosse maggiore di ½) e assolverlo in caso diverso (11).

È chiaro quindi che c'è una relazione tra la scelta dei valori e la soglia di probabilità necessaria per la condanna dell'imputato. Solo alla condizione che v_4 sia molto minore di v_1, cioè solo alla condizione che si consideri molto più grave l'errore insito nella condanna d'un innocente rispetto all'errore insito nell'assoluzione d'un colpevole, la soglia della probabilità oltrepassata la quale il giudice dovrebbe condannare l'imputato può essere molto elevata. La teoria bayesiana è in grado di quantificare in maniera esatta la relazione che sussiste tra il punto di soglia oltre il quale si avrebbe la probabilità sufficiente per la condanna (chiamiamo p_0 tale probabilità) e i valori v_1 e v_4. Sotto l'ipotesi $v_2 = v_3$ (che sentenze "corrette" abbiano ugual valore indipendentemente dal fatto che esse siano di assoluzione o di condanna) tale relazione (di cui ometto la dimostrazione) è la seguente:

(1.1)
$$p_0 = \frac{v_4}{v_1 + v_4}.$$

Dall'equazione (1.1) discende che ciò che determina p_0 è, sotto l'ipotesi $v_2 = v_3$, il rapporto v_4 / v_1. Ciò appare chiaro se denotiamo con *r* tale rapporto, poiché allora vale la seguente identità (della quale anche ometto la dimostrazione):

(1.2)
$$p_0 = \frac{r}{r + 1}.$$

Così, se si ponesse $v_1 = -1$ e $v_4 = -10$, p_0 risulterebbe essere pari a circa il 90%; se invece $v_1 = -1$ ma $v_4 = -100$, p_0 sarebbe pari a circa il 99%, e così via.

Chi scrive ritiene molto importante che sia gli studiosi del diritto delle prove sia il giudice sia il legislatore siano consapevoli del nesso logico che connette il modo con cui sono valutati gli errori giudiziari alla soglia oltre la quale il dubbio diviene "irragionevole". Per esprimerci in maniera un po' approssimativa (ma sostanzialmente corretta) tale soglia dipende da quante volte si ritiene più grave la condanna d'un innocente rispetto all'assoluzione d'un colpevole. È perciò chiaro, alla luce dell'analisi che abbiamo sviluppato, che considerare, *ceteris paribus*, più grave l'errore giudiziario

(11) A rigore, nel caso in cui fosse $p =$ ½ la decisione di condannare avrebbe il medesimo valore della decisione di assolvere. Per evitare che il giudice si trovi paralizzato come l'asino di Buridano, si potrebbe convenire che, in tal caso, egli debba scegliere sempre la decisione di assolvere.

consistente nell’assoluzione dei colpevoli implica la scelta di un valore proporzionalmente più basso per p_0 e quindi la disposizione a condannare con un grado di probabilità della colpevolezza dell’imputato più bassa. Il risultato secondo cui la soglia di probabilità dipende dal *rapporto* v_4/v_1 rende vana la pretesa di contrastare entrambi gli errori giudiziari attribuendo “un grandissimo valore negativo” a ciascuno di essi: decrementando, infatti, in maniera proporzionale sia v_1 sia v_4 (ad esempio moltiplicandoli entrambi per –100), il rapporto v_4/v_1 rimarrebbe invariato. E se si sottraesse sia a v_1 sia a v_4 una medesima quantità positiva *b* molto più grande di ciascuno di essi, allora il risultante rapporto $r=(v_4-b)/(v_1-b)$ diverrebbe prossimo a 1, sì che p_0 — in virtù dell’equazione (1.2) — diverrebbe prossimo al valore ½.

Per contrastare entrambi i tipi di errore l’unica via possibile è quella di ridurre l’incertezza, alterando i valori di probabilità (12). Nel caso-limite in cui *p* oppure *q* avessero valore 1 (si sapesse cioè con certezza se l’imputato è colpevole oppure innocente), non potrebbe esserci alcun errore e la previsione del valore giudiziale sarebbe massima. È chiaro che più ci si approssima a questa situazione ideale più piccola diviene la probabilità di errore. Tuttavia, come si vedrà più avanti, al termine del processo il giudice non dispone più di alcuna fonte di prova che possa promettere di farlo avvicinare in maniera significativa alla certezza. E allora, una volta fissati i valori “finali” di probabilità, se questi non sono vicinissimi a quel caso ideale, si determina una “tensione essenziale” tra l’istanza della certezza della pena per i colpevoli e l’istanza della certezza della colpevolezza per i condannati, istanze che sono irrimediabilmente in competizione, sì che un “compromesso” tra di esse riesce inevitabile. La determinazione del valore minimo per la probabilità della colpevolezza sufficiente per la condanna rispecchia questo compromesso (esprimibile con il rapporto $r=v_4/v_1$) tra le due istanze (13).

4. *Il valore delle fonti di prova.* — Un altro vantaggio del ricorso alla teoria bayesiana della decisione è quello di spiegare perché sia meglio basarsi su molti elementi di prova piuttosto che su pochi. Questo dato è solitamente dato per scontato, ma una spiegazione rigorosa della sua ragionevolezza è meno banale di quanto si possa ritenere *prima facie*. Ciò che va dimostrato è che, allargando l’insieme degli elementi di prova, la previsione del valore aumenta e comunque non può mai diminuire. La prova che le cose stanno in questi termini è stata ottenuta dal matematico statunitense Irving John Good (14). Ciò che consegue dal risultato di Good è che tutte le volte che si è in grado di acquisire a costo zero nuove informazioni potenzialmente rilevanti (nel senso che possono portare, con probabilità maggiore di 0, a una diversa decisione giudiziale), la previsione del valore dell’acquisizione della nuova informazione è strettamente maggiore della previsione del valore giudiziale che si avrebbe rinunciando alla sua acquisizione.

Supponiamo di avere a disposizione una fonte di prova (ad esempio un teste o un documento non ancora esaminato) dalla quale potrebbero derivare nuovi elementi di prova rilevanti. Il giudice

---

(12) Il grado di incertezza può essere misurato mediante la formula $4\times(p-p^2)$, che dà luogo a una parabola convessa che ha valore massimo pari a 1 quando $p=q=½$ e valore minimo pari a 0 quando $p=1$ e $q=0$ oppure $p=0$ e $q=1$.

(13) Attraverso uno studio empirico sulle espressioni qualitative dei gradi di convinzione adoperate nella giurisprudenza si potrebbe cercare di stimare, sia pure in via molto approssimativa, il valore di fatto adoperato di p_0. Sarebbe interessante in particolare vedere se p_0 è invariante oppure no al variare della fattispecie penale. Per un tentativo di determinare le probabilità soggettive con metodi statistici, si veda ad esempio NESS, *Ordinal Positions and Scale Values of Probability Terms as Estimated by Three Methods*, in *Measurement and Evaluation in Counseling and Development*, XXVIII, 1995, p. 152–161.

(14) GOOD, *On the Principle of Total Evidence*, in *British Journal for the Philosophy of Science*, XVII, 1967, p. 319–321.

deve decidere se rinunciare o no all'esame della fonte. La decisione, naturalmente, deve *precedere* la conoscenza del contenuto informativo che da quella fonte di prova può provenire. Supponiamo (ricorrendo a un'idealizzazione che più avanti discuteremo criticamente) che il *costo* dell'esame della fonte sia nullo o (il che in fondo è lo stesso) che di tale costo non si tenga conto ai fini della decisione se esaminare o no la fonte. In questa situazione l'esame della fonte è da escludersi solo se si ha la certezza che le informazioni che se ne trarrebbero sarebbero irrilevanti o, perlomeno, se si ha la certezza che esse non possono alterare la decisione giudiziale che si avrebbe rinunciando a esse. Purtroppo una tale certezza raramente si raggiunge, e a rigore, non si può *mai* raggiungere in maniera perfetta.

5. *Non monotonia e durata del processo*. — Si pone allora la domanda: quando deve concludersi il processo? La risposta canonica è che esso deve concludersi allorché sia stata raggiunta la prova di colpevolezza dell'imputato "oltre ogni ragionevole dubbio", ovvero quando il giudice si sia convinto che una tale prova non può essere raggiunta. Alla base di tale dottrina v'è la convinzione che una "solida" prova di colpevolezza consista non solo in un elevato grado di probabilità (maggiore di p_0) ma anche nel fatto che tale grado sia stato oltrepassato in maniera *definitiva*. Purtroppo, dal punto di vista bayesiano, tale pretesa entra in conflitto con la cosiddetta *non monotonia* delle inferenze probabilistiche.

Spiegherò brevemente che cosa sia la monotonia e la ragione per la quale essa, mentre vale nella logica deduttiva, non può valere ove vi sia anche la minima incertezza. Le inferenze della logica *deduttiva* sono tali per cui se una proposizione *q* è validamente deducibile a partire da certe premesse $p_1, p_2, \ldots, p_n$ essa è deducibile anche da un insieme *più ampio* di premesse che comprenda, oltre a $p_1, p_2, \ldots, p_n$, ulteriori premesse. Ciò reca con sé la conseguenza che se si è in possesso della prova certa di $p_1, p_2, \ldots, p_n$, la verità di *q* deve considerarsi acquisita definitivamente, senza che nuove prove possano farla revocare in dubbio.

Purtroppo, nell'ambito del diritto delle prove, l'ideale di una prova concludente che discenda deduttivamente da elementi di prova certi è illusorio. Il giudizio concludente va tipicamente *oltre* le prove ed è per ciò stesso incerto. Il significato dell'espressione "oltre ogni ragionevole dubbio" *non* ha nulla a che vedere con quella certezza assoluta che, purtroppo, non ha cittadinanza alcuna al di fuori della matematica pura. Il bayesianesimo giuridico muove esattamente da questa constatazione. Tuttavia, se l'inferenza giudiziale non è rigorosamente deduttiva ed esiste quindi un sia pur minimo margine d'incertezza o possibilità d'errore, allora essa è non monotonica. In altri termini, essa non è *mai* raggiunta con un grado di certezza che in linea di principio non possa diminuire in seguito all'acquisizione di nuove prove. Ciò ha come conseguenza che il superamento della soglia p_0 della probabilità di colpevolezza non è di per sé motivo sufficiente per considerare definitivamente "raggiunta" la prova concludente e quindi conclusa la fase di acquisizione degli elementi di prova. Se, dopo aver superato p_0, la difesa esibisse un nuovo alibi e lo sostanziasse di prove cogenti, la probabilità di colpevolezza potrebbe scendere ben al di sotto del valore p_0. E un siffatto "colpo di scena probatorio" è sempre *possibile.*

Vediamo quindi che il carattere di non monotonia delle inferenze probabilistiche, presa insieme al valore positivo di qualsiasi fonte di prova potenzialmente rilevante, implica che la teoria bayesiana raccomanderebbe sempre di proseguire *ad infinitum* nella raccolta degli elementi di prova, senza mai arrivare a considerare le acquisizioni probatorie tali da ritenere maturi i tempi per l'emanazione della sentenza. In tal modo, però, il processo non avrebbe mai termine.

6. *La solidità pragmatica dei valori di probabilità*. — Dal punto di vista puramente *logico* non v'è soluzione al paradosso sollevato dalla non monotonia delle inferenze probabilistiche. Infatti, in qualsiasi momento si potrebbe sperare che, rinviando *ad infinitum* la conclusione del processo, emergerebbe prima o poi una fonte da cui deriverebbe una prova tale da condurre a un

rovesciamento della decisione giudiziale apparentemente ormai stabilita in maniera definitiva. Anche una *minima* probabilità che ciò avvenisse dovrebbe suggerire al giudice il rinvio della sentenza. Ciò è naturalmente contrario tanto al buon senso quanto al principio della ragionevole durata del processo. Per questo ritengo che sia di vitale importanza per il bayesianesimo giuridico trovare la soluzione di questo paradosso.

Intuitivamente una valutazione di probabilità, a parità del valore numerico che essa esprime, può essere più o meno *solida* a seconda che essa si fondi su molte o poche prove. Ad esempio, sembra ragionevole giudicare pari a ½ la probabilità che, lanciando una moneta mai lanciata in precedenza, si presenti il lato “testa” in un determinato lancio futuro *L*; ma tale valutazione appare molto più solida o fondata se si è prima di *L* lanciata la moneta un milione di volte ed esattamente nella metà dei casi si è presentato il lato “testa”. La differenza tra i due casi consiste in questo: che nel primo caso la convinzione subirebbe ragionevolmente una significativa modifica se si facesse, prima di *L*, una serie – poniamo – di dieci lanci “preliminari” ed essi risultassero tutti croce o tutti testa, laddove nel secondo caso una tale serie di lanci preliminari modificherebbe in maniera insignificante la probabilità che *L* dia luogo al risultato “testa”.

È importante osservare che la maggiore solidità probatoria del secondo giudizio *è relativo alle fonti di prova di cui si dispone*. Se fosse dato di poter conoscere con molta precisione tutti i parametri fisici del lancio *L*, si potrebbe predire con altrettanta precisione il risultato di *L*, la qual cosa significa attribuire a uno dei due risultati possibili una probabilità molto alta e all’altro una probabilità molto bassa. Se tali informazioni fossero realmente accessibili e il calcolo della probabilità subordinatamente a tali informazioni potesse essere portato a termine, allora il milione di lanci precedenti non avrebbe presumibilmente influenza alcuna sulla valutazione di probabilità (essi sarebbero pressoché irrilevanti alla presenza della nuova fonte di prova), sì che il valore complessivo di tali lanci sarebbe pressoché nullo. In altri termini: dal punto di vista di chi avesse a disposizione quei parametri fisici come fonte di prova, la valutazione basata su un milione di lanci sarebbe ben poco più solida di quella iniziale basata solo sull’ignoranza completa. Tuttavia, se le sole informazioni accessibili fossero quelle statistiche, allora il giudizio basato su un milione di prove sarebbe molto più solido. La conclusione che se ne deve trarre è che la solidità probatoria di una valutazione probabilistica non è qualcosa di “intrinseco”, una proprietà della valutazione di probabilità che conferisca ad essa una maggiore “oggettività”, bensì dipende dall’insieme delle fonti di prova a disposizione. Si tratta quindi di una nozione *pragmatica*, non *logica*. Essa, cioè, dipende essenzialmente dal *contesto* e, segnatamente, dal contesto delle fonti di prova accessibili nella concreta situazione processuale.

Queste considerazioni suggeriscono di misurare la solidità dei gradi di probabilità inferiti dalle prove sulla base del valore complessivo delle fonti di prova a disposizione. Chiamo *contesto probatorio* l’insieme delle fonti di prova rilevanti che in un dato momento del processo sono consultabili. Il calcolo simultaneo della previsione del valore della consultazione di più prove non presenta di per sé alcuna novità concettuale rispetto a quanto si è già detto rispetto a una singola fonte (anche se le formule mediante le quali tale calcolo è compiuto sono alquanto più complesse). Possiamo allora considerare il valore complessivo della consultazione delle fonti che appartengono al contesto probatorio come una misura della *labilità probatoria* del valore di probabilità rispetto alla decisione giudiziale raggiunto sino a quel momento. Una bassa labilità equivale naturalmente a un’alta solidità probatoria. Quindi il valore dell’informazione che può provenire dal contesto probatorio fornisce una misura della solidità del valore di probabilità del fatto su cui verte la decisione giudiziale.

Se includiamo nell’idea di “prova oltre ogni ragionevole dubbio” non solo la richiesta di un elevato grado di probabilità, ma anche quella di un elevato grado di solidità probatoria di quest’ultimo (nel senso specificato), ne consegue che si dovrebbe proseguire l’acquisizione di nuove risultanze probatorie *sino a quando il valore della consultazione delle fonti a disposizione non divenisse minore di un certo valore-soglia* v_0.

7. *La durata del processo come "costo"*. — Si pone tuttavia il problema di come far discendere il valore v_0 dai valori giudiziali v_1, v_2, v_3, v_4, in modo analogo a quello con cui avevamo potuto desumere da tali valori p_0. È facile però vedere che una tale deduzione è impossibile nel quadro della teoria bayesiana della decisione, in quanto quest'ultima prescrive sempre come razionale la consultazione di qualsiasi fonte di informazione ottenibile a costo nullo il cui valore sia maggiore di 0. E qualsiasi informazione potenzialmente in grado di incidere sulla decisione del giudice ha un valore maggiore di 0.

La soluzione di questo rompicapo consiste nel considerare la durata stessa del processo come un "costo", posto che sia nell'interesse dell'ordinamento la minimizzazione, *ceteris paribus*, della durata del processo. Possiamo rappresentare il costo della durata con una funzione $c(t)$ che soddisfa le tre seguenti proprietà: (a) che $c(t)$ abbia valore 0 per $t = 0$; (b) che $c(t)$ sia strettamente crescente al crescere di t e (c) che $c(t)$ sia *limitata superiormente*, assumendo un valore in ogni caso strettamente minore del valore di un'ideale informazione perfetta che attribuisca probabilità 0 oppure 1 alle possibilità (innocenza e colpevolezza) da cui dipende la decisione giudiziale.

Le proprietà (a)–(c) sono sufficienti per rendere praticamente certo che prima o poi si arriverebbe a un punto in cui, secondo la teoria bayesiana, sarebbe irrazionale l'esame di nuove fonti di prova al fine di acquisire ulteriori informazioni. Il valore previsto della consultazione della fonte sarebbe, infatti, pari alla differenza tra il valore calcolato come spiegato più sopra e la previsione di $c(t)$ supposto che si decida di consultare la fonte. Quando $c(t)$ è maggiore del valore della consultazione della prova, tale consultazione non dovrebbe essere effettuata. E quando ciò vale per tutte le fonti di prova disponibili, la fase di acquisizione probatoria deve considerarsi conclusa (15).

8. *Conclusioni*. — Abbiamo mostrato che il diritto bayesiano delle prove potrebbe giovarsi in maniera significativa del formalismo della teoria bayesiana della decisione e non solo, come accade nella letteratura del bayesianesimo giuridico, di quello della probabilità. Il modello teorico generale richiede l'introduzione di *valori giudiziali*, i quali dovrebbero essere implicitamente contenuti nell'ordinamento e idealmente da esso desunti. Si è illustrato come attraverso questa impostazione sia possibile catturare in modo esatto il nesso che sussiste tra la richiesta di un grado di probabilità "oltre ogni ragionevole dubbio" per la condanna penale e la diversa valutazione dei due possibili errori giudiziari: assolvere un colpevole e condannare un innocente. In questo modo il ricorso alla teoria della decisione fornisce una trattazione più profonda rispetto a quella fornita dal bayesianesimo giuridico tradizionale, perché consente di spiegare la razionalità di quella richiesta desumendola dai valori giudiziali.

Inoltre l'applicazione al diritto delle prove della teoria della decisione non conduce solamente a una trattazione *più profonda*, bensì consente anche di risolvere difficoltà e trattare questioni che l'impostazione tradizionale del bayesianesimo giuridico non riesce a modellare in maniera soddisfacente. Tra i problemi sostanzialmente irrisolvibili nell'ambito del bayesianesimo giuridico tradizionale v'è il problema di una caratterizzazione dell'idea della *solidità probatoria* delle valutazioni probabilistiche, nozione questa che sembra essere indispensabile per una caratterizzazione della prova concludente. Abbiamo mostrato che da un punto di vista pragmatico,

---

(15) In questa prospettiva è possibile precisare le nozioni di *non verosimiglianza*, di *non pertinenza* e di *irrilevanza*, considerate dal diritto delle prove come motivo di inammissibilità di nuove acquisizioni probatorie (Cfr. UBERTIS, *La prova penale*, Utet, 1995, p. 58–65). Nessuna di tali condizioni va interpretata nel senso di un'*assoluta* non informatività del materiale probatorio giudicato inammissibile, bensì nel senso che il valore informativo di tale materiale è, in previsione, talmente basso da non eccedere il costo previsto dell'allungamento del processo che la sua acquisizione recherebbe con sé.

tale cioè da fare riferimento essenziale a quello che abbiamo chiamato il *contesto probatorio*, questo problema può essere risolto nell'ambito della teoria bayesiana della decisione. Abbiamo infine fatto vedere che le stesse idee conducono a una soddisfacente chiarificazione dell'idea di "ragionevole durata del processo" che sembra difficilmente catturabile con gli strumenti limitati del bayesianesimo tradizionale.

La teoria bayesiana, nella sua versione completa, basata sull'idea di *decisione in condizioni d'incertezza,* ha le carte in regola per dar luogo a una nuova stagione del bayesianesimo giuridico, tale da consentire il superamento di quelle difficoltà che a molti sono sembrate fatali a questa impostazione. Non si vuole con questo sostenere che la via qui prospettata sia quella definitiva. Essa ha a sua volta i suoi limiti. Ad esempio, essa non tiene nel dovuto conto il fatto che le parti sono persone razionali, ciascuna con i suoi fini, le quali *interagiscono* reciprocamente. La teoria bayesiana della decisione è adatta a modellare tutto ciò che attiene alla *decisione individuale* (16), qual è ad esempio quella del giudice. Una modellazione più completa del processo penale, che includa anche quella di tutte le *dramatis personae* del processo, potrà venire solo da un'applicazione della *teoria dei giochi*, la quale costituisce un ulteriore ampliamento della teoria della decisione, in cui si considerano le decisioni di ciascun decisore in rapporto a quelle degli altri decisori. Come ognuno può vedere, il bayesianesimo giuridico è un programma di ricerca vitale e promettente, ma è ancora nello stadio iniziale del suo sviluppo.

(16) O, tutt'al più, alla decisione collettiva da parte di un collegio di decisori tutti idealmente portatori di un medesimo sistema di preferenze.