ISSN 1125-856X

Cassazione penale

direttore scientifico **Domenico Carcano**
condirettore **Mario D'Andria**
LVIII - ottobre 20**18**, n° **10**

10 2018

estratto

LA VALUTAZIONE DEGLI INDIZI SECONDO IL CARDINALE NEWMAN: DALLA TEOLOGIA AL DIRITTO

di **Bruno Arrigoni, Cristina Di Girolami, Lorenzo Peccati**

MVLTA PAVCIS AG GIUFFRÈ EDITORE

470 LA VALUTAZIONE DEGL'INDIZI SECONDO IL CARDINALE NEWMAN: DALLA TEOLOGIA AL DIRITTO

The evaluation of circumstantial evidence according to Cardinal Newman's principle: from Theology to Law

Partiamo dal problema teologico dell'esistenza di Dio e mostriamo come lo schema proposto secoli fa dal Cardinale Newman per trattarlo possa essere utile in un processo indiziario per valutare la significatività d'un insieme d'indizi e nel disegno d'una strategia processuale in presenza d'eccezioni.

Starting from the theological problem of the existence of God, we show that the scheme proposed centuries ago by Cardinal Newman can be fruitfully used in a circumstantial legal trial to assess the significance of the available information and in order to design a trial strategy in the presence of possible objections.

(Traduzione in inglese a cura degli Autori)

di **Bruno Arrigoni**
Avvocato Cassazionista, Docente nella Scuola Forense di Monza

di **Cristina Di Girolami**
Ricercatrice all'Università degli Studi di Chieti-Pescara

di **Lorenzo Peccati**
Professore Emerito di Matematica all'Università Bocconi, Milano

1. INTRODUZIONE

Nel mondo giuridico è prepotentemente entrata la nozione di probabilità. Un interessante contributo di P. Garbolino [1] è un buono specchio di naturali problemi che nascono quando mondi, magari pensati disgiunti, si trovano a dover condividere ... il pianerottolo.

Lo stile di questo contributo è molto lontano dal normale stile della letteratura giuridica, ma la maggioranza degli Autori non è di giuristi e la minoranza si concede qualche licenza [2].

Useremo lo spazio a disposizione per illustrare una rilettura, che speriamo interessante, d'un problema classico, all'incrocio tra logica, incertezza e diritto. Già il fatto che vi siano tre direzioni, fa presagire le difficoltà al semaforo: meglio due, ma, purtroppo, sono tre.

———

(1) P. Garbolino, *Dall'effetto probabile alla causa probabile. La valutazione del nesso causale*, in *questa rivista*, 2004, p. 300 ss.

(2) Se lo scritto risulterà gradito al Lettore giurista, gran parte del merito è del prof. Giulio Ubertis, responsabile di questa rubrica, che ci ha aiutato con competenza e pazienza a migliorare l'esposizione.

Due AA. hanno già cercato d'aprire un dialogo sui rapporti tra diritto e matematica, in vari modi (3).

Illustreremo il problema classico, noto come il "principio del cardinale Newman", che, forse, a suo tempo, non immaginava fosse così rilevante, in tema d'aggregazione dei valori probanti d'indizi plurimi. Lo schema sagace del cardinale Newman è suscettibile d'una reinterpretazione (4), per certi versi, inaspettata, in termini di strategie processuali.

Quanto l'eventuale lettore legge è frutto d'un processo non facile.

Il linguaggio giuridico è formale, ma a suo modo scarica debiti logici su nozioni come la "certezza giuridica" e la garanzia «al di là d'ogni ragionevole dubbio», che meritano totale apprezzamento per chi *cotidie* deve decidere, ma che hanno bisogno di corroborazione logica, quando s'esca dal quotidiano.

Anche perché il domani sarà diverso.

Non è stato facile giungere a questo contributo: alla fin fine, si tratta d'un problema a cavallo tra due mondi (diritto e probabilità), che non parlano la stessa lingua, ancorché entrambi s'occupino spesso degli stessi problemi.

Il problema linguistico sta nell'uso del linguaggio matematico, ostico ai più, ma talora necessario.

Gli AA. conoscono il problema.

Normalmente, i giuristi non vedono di buon occhio i numeri, forse poiché pensano ch'essi stiano alle loro abilità come la livella di Totò alla disuguaglianza tra gli uomini.

Dunque gli AA. sanno che i giuristi, prevalenti lettori di questo articolo invitati ad usare i numeri, porranno probabilmente l'obiezione: «discorso affascinante, ma fallato in partenza: come si può misurare la probabilità di successo d'un'eccezione (5), se, ad esempio, 40% o 70%?».

Ricorriamo a un esempio affatto banale.

Qualunque giurista dovesse incontrare gli AA. di persona, saprebbe certo dire, senza errare, chi fra loro sia il più basso di statura, senza però necessariamente essere in grado di riferirne l'altezza esatta.

Se poi lo misurasse, dovrebbe opportunamente precisare che il metro e settantatré accertato è tale se, e solo se ("*iff*", per "*if and only if*", scriverebbe un matematico, impiegando 3

———

(3) B. Arrigoni, L. Peccati, *Diritto e Matematica: due mondi disgiunti?*, in *Annali del Dipartimento di Metodi e Modelli per l'Economia del Territorio e la Finanza (Perspectives on Behavioural Sciences)*, Patron, 2016, p. 33-53, ove il problema è affrontato in generale. Quest'articolo approfondisce una questione specifica.

(4) Fenomeno comune in matematica, ove l'astrazione porta a scoprire similarità strutturali inaspettate, meno in altri àmbiti, in cui il concreto ancoraggio spesso costringe.

(5) Un esempio elementare. Si pensi alla condizione di procedibilità (querela) prevista dal codice penale per alcune fattispecie di reato. Ove la querela appaia sporta non tempestivamente (ordinariamente, 3 mesi dalla conoscenza del fatto), il difensore del querelato avrà buon diritto a ritenere che un'eccezione di difetto di procedibilità sia "più fondata che non", benché non possa assegnare al successo di detta eccezione probabilità 1 (cioè certezza); il querelante potrebbe infatti dimostrare, per esempio, d'aver avuto conoscenza del fatto delittuoso successivamente alla sua commissione (così posticipando il *dies a quo* per il calcolo dei 3 mesi), ovvero il giudice o il pm potrebbero riqualificare il fatto o rinvenire ulteriori fattispecie di reato tali da rendere l'azione procedibile d'ufficio. Per contro, se la querela fosse tempestiva, il difensore del querelante potrebbe ritenere soddisfatta la condizione di procedibilità con probabilità prossima a 1 (ma non così anche l'esito del giudizio, stante l'incertezza in capo al querelante circa la decisione nel merito per lui favorevole, che impedisce d'assegnare probabilità 1 al successo nell'esito complessivo del giudizio). In entrambi i casi, il difensore (del querelante, anziché del querelato) sul tema specifico dell'eccezione di procedibilità è senz'altro in grado di esprimere un ordine di grandezza prognostico circa l'esito favorevole oppure sfavorevole al proprio assistito dell'eccezione.

lettere anziché le 9 d'un giurista ...), la statura è espressa col sistema metrico decimale. Infatti, altrettanto giustamente, qualcun altro potrebbe invece dire che egli è alto 5,675853 (usando il piede anglosassone).

Numeri diversi per una stessa grandezza, perché, ciascuno, frutto d'una convenzione tra le (spesso) infinite possibili.

Accomodando la *vexata quaestio* delle diverse stature degli AA. alle nostre eccezioni, non dubitiamo che i giuristi hanno senz'altro una percezione dell'ordine di grandezza della probabilità di accoglimento di un'eccezione (6). Ciò nonostante, molti s'ostinano a ritenerne impossibile la misurazione, e s'affidano dunque a mutevoli (e confuse) definizioni del "grado" di probabilità (7), laddove il "grado" pur somiglia a una scala di misura, solo che non ha il vantaggio di potersi esprimere con un numero.

Se i numeri sono frutto d'una convenzione, forse allora converrà ai giuristi valutare l'opportunità d'adottarne una, a loro piacimento, che consenta d'esprimere – appunto – un numero: decidano loro se il sistema metrico decimale o i citati piedi, se i chilogrammi o le libbre, se Celsius o Fahrenheit (8).

Il fondamento d'una convenzione siffatta, però, non può che radicarsi *in primis* nella necessità di riconoscere che la probabilità è ciò che è: un numero compreso fra 0 e 1, ossia tra

———

(6) Un esempio elementare. Si pensi a un'eccezione di decadenza dalla garanzia per vizi riconoscibili *ex* art. 1495 c.c., a seconda che risulti effettuata oppure no, la denunzia entro il termine di 8 giorni dalla consegna. Se la denunzia sarà stata effettuata, l'eccezione avrà un "peso" (non risolutivo comunque, poiché l'incertezza residua, che impedisce d'assegnare immediatamente probabilità di successo 1 all'eccezione, risiede nella prova della sussistenza del vizio); se la denunzia non sarà stata effettuata entro gli 8 giorni, ma solo il 9°, l'eccezione avrà un altro peso (anche qui, comunque, minore di 1, poiché il giudice potrebbe optare per una diversa qualificazione del contratto – appalto, anziché vendita – con diverso termine di decadenza: 60 gg. e non 8). Un'assegnazione immediata di una probabilità richiederebbe un ulteriore e non banale uso della probabilità condizionata. Generalmente essa non sarebbe 1.

(7) Raduniamo alcune indicazioni sul "grado di probabilità" come appaiono in diritto in tema di sussistenza del nesso causale. Nel solo anno 2015, la Corte di cassazione penale ha descritto il grado di probabilità utile per integrare la prova dell'esistenza del nesso causale impiegando almeno 10 differenti locuzioni: *certezza* o *alta probabilità* (Sez. III, 17 dicembre 2015, n. 842); alta probabilità (Sez. III, 10 novembre 2015, n. 50454); *significativa probabilità* (Sez. II, 24 settembre 2015, n. 40912, in *C.E.D. Cass.*,n. 264589); *certezza o elevata probabilità* (Sez. III, 15 settembre 2015, n. 43113, K.M. ed altro, in *Dir. & giust.*, 2015); *alto grado di probabilità logica* (Sez. fer., 25 agosto 2015, n. 41158, P.G. in proc. E. ed altri, in *C.E.D. Cass.*, n. 264883); *ragionevole, concreta ed attuale probabilità* (Sez. II, 10 giugno 2015, n. 30939, C., *ivi*, n. 264401); *significativamente probabile* (Sez. III, 19 maggio 2015, n. 37092, S., *ivi*, n. 265177); *elevata e qualificata probabilità* (Sez. I, 12 maggio 2015, n. 31421, L., *ivi*, n. 264757); *ragionevole probabilità di colpevolezza nella fase cautelare o certezza non incisa dal ragionevole dubbio nella fase di merito* (Sez. II, 6 maggio 2015, n. 28602, P., *ivi*, n. 264138); *elevata "probabilità logica"* (Sez. IV, 4 febbraio 2015, n. 11136, C.M., in *Riv. it. med. leg.*, 2015, n. 2, p. 665). Le cose non vanno meglio in àmbito civile, ove la S.C., benché in arco temporale più ampio di quello sopra analizzato, è ricorsa alle seguenti locuzioni: *probabilità prevalente* (Cass. civ., sez. III, 5 maggio 2009, n. 10285, Soc. Aerolinee Itavia c. Min. difesa e altro, in *Resp. civ. prev.*, 2010, n. 2, p. 461); *elevata probabilità* (Cass. civ., Sez. III, 11 giugno 2009, n. 13530, Z. c. P., in *Giust. civ. Mass.*, 2009, 6, p. 904); *altamente probabile e verosimile – serio e ragionevole criterio di probabilità scientifica* (Cass. civ., Sez. III, 30 ottobre 2009, n. 23059, Canzanella c. De Paola ed altro, in *Giust. civ. Mass.*, 2009, 10, p. 1522); *adeguata probabilità* (Cass. civ., sez. lav., 26 marzo 2010, n. 7352, Liverani ed altro c. Inps ed altro, *ivi*, 2010, 3, 446); *probabilità qualificata* (Cass. civ., sez. lav., 5 agosto 2010, n. 18270, Nalin c. Inail ed altro, *ivi*, 2010, 9, 1186); *notevole grado di probabilità* (Cass. civ., Sez. lav., 10 febbraio 2011, n. 3227, Inail c. N.S., *ivi*, 2011, 2, 213); *rilevante grado di probabilità* (Cass. civ., sez. lav., 11/10/2012, n. 17349); *certezza probabilistica, desumibile dal preponderante criterio dell'evidenza logica* (Cass. civ., Sez. III, 11 febbraio 2014, n. 3010, R.B. c. Duomo assicur. - *online* 2014). Ci pare evidente che del problema della probabilità vi sia contezza. Meno del fatto che siano noti gli strumenti appropriati per la soluzione dei vari problemi.

(8) Suggeriamo uno sguardo alle regole fissate da *ISO*: Organizzazione internazionale per la normazione (*International Organization for Standardization*) in tema di standardizzazione di unità di misura. *ISO* è la più importante organizzazione a livello mondiale per la definizione di norme tecniche.

la probabilità d'un evento impossibile e la probabilità d'un evento certo. A ben vedere, non v'è nulla di più pertinente, coi giuristi, che s'affannano a descrivere il loro mestiere come costellato da un'infinità di possibili soluzioni diverse.

Tradurre la percezione dell'ordine di grandezza in una misura è la sfida che gli Autori pensano si possa risolvere, con vantaggio per le professioni legali, proprio partendo dall'esempio offerto dal cardinale Newman, e facendo tesoro dell'insegnamento della teoria della Probabilità dovuto a Bruno de Finetti: la soluzione sta nell'approccio operativo della teoria "soggettiva", che, val la pena ricordare, non significa affatto "arbitraria".

Le implicazioni che possono scendere dall'adozione d'una siffatta scala di misura sono straordinarie, come gli AA. già stanno vedendo nel trattare il tema del nesso causale (sia in àmbito civile, sia in penale), ove la selva di concetti espressi s'aggroviglia sempre più, addirittura incorrendo in svarioni di calcolo (9), proprio perché non sempre s'accetta di considerare la probabilità solo come lo strumento operativo che essa è.

Ne daremo qui brevemente conto.

Un osservatore qualsiasi, anche distratto, si rende anzitutto conto che la solidità dell'italica diffidenza nei confronti del dialogo fra matematica e diritto, allorché si discuta di probabilità (10), è presto incrinata guardando all'estero, e specialmente agli Stati Uniti (11), ove si registra da tempo una vivace scuola di pensiero che cerca di costruire sistemi giuridici improntati al "numero". Noi non siamo convinti che riducrci ai numeri sia sempre vincente, ma siamo altrettanto convinti che tentare di rifiutarli a priori sia perdente.

La bibliografia di lingua inglese consultabile in argomento è poderosa, e spazia in àmbiti molto diversi fra loro, accomunati però dal tentativo di teorizzare il processo giuridico di

———

(9) In una sentenza della Corte di cassazione civile risulta mal applicata la regola del 50% *plus unum*. Cass. civ., Sez. III, 21 luglio 2011, n. 15991 (Osp. S. Giovanni di Dio Fatebenefratelli c. C.T. e altro, in *Giust. civ. Mass*., 2011, 7-8, p. 1098): «la concorrenza di cause di diversa incidenza probabilistica non conduca ipso facto alla aberrante regola del 50% *plus unum*, bensì alla compiuta valutazione dell'evidenza del probabile; così, esemplificando, se, in tema di danni da trasfusione di sangue infetto, le possibili concause appaiono plurime e quantificabili in misura di dieci, ciascuna con un'incidenza probabilistica pari al 3%, mentre la trasfusione attinge al grado di probabilità pari al 40%, non per questo la domanda risarcitoria sarà per ciò solo rigettata». Ragionevolmente si intuisce che l'evento più probabile è quello trasfusionale, dal momento che 30% < 40%. Per utilizzare la regola del 50% *plus unum* in modo corretto si devono dapprima normalizzare i valori delle probabilità a 1. Infatti, analizzando il problema, risulta che la somma delle probabilità di tutti gli eventi considerati risulta 3% x 10 + 40% = 70%. Da cui, le probabilità di ognuna delle 10 possibili concause ha probabilità 3%/70% = 4.285% e la probabilità dell'evento trasfusionale vale 40%/70% = 57.15%. Ora la somma delle probabilità risulta 4.285% x 10 + 57.15% = 100%: con i nuovi valori opportunamente normalizzati è ora possibile utilizzare la regola del 50% *plus unum* confermando lo stesso risultato.

(10) V., ad es. e fra gli altri, M. Taruffo, *Note sparse su probabilità e logica della prova*, in *Riv. trim. dir. e proc. civ*., 2014, n. 4, p. 1507 ss., ove s'afferma: «rimango infatti convinto – e mi confortano le opinioni di numerosi altri studiosi del tema – che siano gli strumenti della probabilità logica, e non quelli del calcolo della probabilità quantitativa, a consentire una comprensione e un'analisi adeguate della logica della prova».

(11) Inserendo in un qualsiasi motore di ricerca sul *web* i termini "*math*" e "*law*", e con un poco di pazienza, si scoprirà che si danno riscontri in quantità inimmaginabile per un giurista italiano (il quale troverebbe assai meno, digitando "matematica" e "diritto"). Si potrebbe pensare che la maggiore propensione statunitense a quel tipo di studi e di tecniche non sia solo il frutto d'un atteggiamento culturale di fondo, inteso in senso lato, votato all'empirismo e al pragmatismo tipici del mondo anglosassone. Si potrebbe sostenere che tale propensione è anche il portato d'un sistema giudiziario che si regge sull'analisi dei case, in virtù del principio dello *stare decisis*, che senz'altro ben si presta, per sua natura, ad una analisi di tipo statistico, e quindi ad una riduzione a "numero". Ma il sistema giudiziario non sembra altro che un'ulteriore foglia di fico dietro la quale la cultura dell'insanabile nostrana dicotomia Norme/Numeri tenta di nascondersi: anche nei Paesi di *Civil Law*, e in Italia in particolare, il principio dello *stare decisis* non è infatti del tutto estraneo in via sostanziale al sistema, come ben sanno gli avvocati quando s'imbattono in una sentenza della Corte di cassazione a sezioni unite.

decisione come frutto d'una valutazione sostanzialmente probabilistica, con studi che si spingono fino alla previsione delle decisioni della corte in funzione dei singoli giudici facenti parte del collegio [12].

Un recente esempio in letteratura [13] affronta, nella carrellata dei case rilevanti, anche l'omicidio di Meredith Kercher [14], anticipando l'assoluzione di Amanda Knox e Raffaele Sollecito, poi effettivamente pronunziata il 27.03.2015 dalla V sezione penale della Corte di cassazione.

La citazione d'un "caso" giudiziario noto, trattato dalla dottrina straniera con criteri matematici, ci offre la possibilità d'illustrare una non banale applicazione concreta della probabilità a un altro [15] episodio di cronaca giudiziaria italiana, quando l'avv. Raffaele Della Valle si cimentò in un esperimento giudiziale saltando da un balcone all'altro d'una casa di ringhiera, nel contesto d'un processo penale per omicidio.

L'accusa riteneva che non esistesse una via d'accesso/fuga diversa dall'unico portone affacciato sulla pubblica via. Poiché un testimone (intento a fumare sigarette sul balcone di casa che dava giusto su quel portone) aveva riferito che, né prima, né dopo l'orario di consumazione dell'omicidio, qualcuno era entrato e/o ne era uscito, la procura aveva imbastito il processo nei confronti del figlio, convivente della vittima, trovato in casa col morto, ma che urlava la propria innocenza dicendo che un estraneo s'era introdotto nell'abitazione.

A distanza d'anni, ricostruendo il processo in termini probabilistici, e richiamata la nozione d'evento complementare, si potrebbe dire che:

E: c'è una sola via d'accesso fuga, secondo la procura;

non E: ve ne sono altre.

L'opinione della procura era che Pr (E) = 1, mentre Pr (*non* E) = 0.

L'avv. Della Valle, affacciatosi sul retro della casa di ringhiera, saltò dal ballatoio [16] dell'appartamento ove s'era consumato l'omicidio, riuscendo a raggiungere il balcone d'una casa dirimpetto, dalla quale ebbe poi buon giuoco a uscire attraverso un portone ovviamente diverso da quello d'accesso "normale" all'abitazione luogo del delitto.

L'esperimento dimostrò l'esistenza d'una via d'ingresso/uscita diversa dalla "controllata" dal testimone. A parte alcuni aspetti tecnici, provare che un evento è possibile equivale a provare che il suo evento contrario non è certo. La necessità, per la condanna, della verità, al di là d'ogni ragionevole dubbio, condusse all'assoluzione. Ma tutto ciò conferma la nostra sospettata ingenuità nel rapporto tra il giurista e la probabilità.

Gli AA. cercheranno d'esprimere il loro ragionamento probabilistico verbalmente o, se del caso, attraverso semplici esempi numerici.

La formalizzazione è rigorosamente confinata in appendice.

Presenteremo dapprima molto sinteticamente (par. 2) il problema "storico", per poi rileggerlo nell'altra prospettiva (par. 3). Alcune righe conclusive segnaleranno (par. 4) l'interesse

———

(12) V., uno su tutti, L.H. Tribe, *Trial by Mathematics: precision and ritual of a legal process*, in *Harvard Law Review*, 1971. Qui l'incertezza percepita è endogena rispetto al sistema, mentre l'incertezza "vera" è esogena e pervasiva.

(13) L. Schneps, C. Colmez, *Math on Trial: How Numbers Get Used and Abused in the Courtroom*, Basic Books, 2013.

(14) Cap. IV: "*Math error number 4: double experiments. The case of Meredith Kercher: the test that wasn't done*".

(15) Passato e senza *pathos* corrente.

(16) Suggeriamo il *link http://www.studiolegaledellavalle.com/svbuilder_rdv/gallery_rdv.html*.

della questione dal punto di vista delle professioni legali. Confiniamo in coda (par. 5) alcuni aspetti che richiedono una trattazione più formalizzata.

2. IL PROBLEMA DEL CARDINALE NEWMAN E LA SUA PRIMA RIVISITAZIONE

2.1. Come nasce il problema

Che la matematica possa aver a che fare col diritto è cosa di cui, come ben sanno gli AA., parecchi dubitano. Che essa sia impiegata addirittura in Teologia, potrebbe perfino apparire sacrilego.

Eppure v'è stato un campione, proprio nella storia della teologia cattolica, che ha usato la probabilità per argomentare la prova (niente meno che) dell'esistenza di Dio: John Henry Newman [17].

A 16 anni (ancora anglicano, benché agli AA. non interessi quali fedi abbia professato in vita), egli, come avrebbe poi scritto, ebbe «*coscienza di ciò che è un dogma*» (in questo prendendo assai poco le distanze da un matematico, se non pel fatto che quest'ultimo preferisce termini come *assioma* o *postulato* ...) e ciò servì a farlo «*riposare nel pensiero di due esseri, entrambi unici, supremi, entrambi testimoniati da un complesso di prove abbaglianti: l'io e il suo Creatore*».

Prove abbaglianti? Sì: esattamente, come quelle che ogni giurista spera d'avere a propria disposizione, quando scende nell'agone di carneluttiana memoria.

Ma, come càpita ai giuristi, non di rado si hanno tra le mani (solo) indizi, anziché prove; e, così dicono i meglio informati, la bravura di chi è talentuoso risiederebbe nel trasformare i primi nelle seconde, grazie – amano sempre dire i giuristi – all'uso della logica [18].

———

[17] Londra, 21 febbraio 1801 - Edgbaston, 11 agosto 1890. Cardinale, teologo e filosofo inglese, che si convertì in età matura, dall'anglicanesimo al cattolicesimo. È sorprendente il numero di corrispondenze che qualsiasi motore di ricerca restituisce per il suo nome, ai più sconosciuto.

[18] V., per esempio, Sez. IV, 17 settembre 2010 (dep. 13 dicembre 2010), n. 43786: «L'idea di causa è parte essenziale dell'immagine che l'uomo ha del mondo ed è uno strumento concettuale insostituibile per la soluzione di problemi conoscitivi che continuamente sorgono nella vita e, quindi, nella giurisprudenza. Per conseguenza, non può essere accettata l'idea che la soluzione d'un problema causale debba sempre sottendere una base scientifica precostituita e certa. Occorre invece considerare che, sia pure non frequentemente, le emergenze fattuali sono tante e tanto chiare da consentire d'articolare il ragionamento probatorio in chiave induttiva, cioè sulla base della mera analisi logica dei fatti. *Si tratta del baconiano ragionamento induttivo, i cui canoni sono sviluppati pure da J.S. Mill.* L'esempio classico è dell'agente che patisce quando assume una determinata sostanza ed è invece immune da sofferenza quando non l'assume. Il danno è maggiore quando l'assunzione è abbondante. È escluso che la lesione sia determinata da altre sostanze assunte in concomitanza, poiché esse fanno parte delle consolidate ed innocue abitudini del soggetto». Non mancano esempi anche in diritto civile; v., per es., Trib. Milano, Sez. X, 20 ottobre 2009, n. 12415, in *Giustizia a Milano*, 2009, n. 11, p. 80, per la quale «lo standard di "certezza probabilistica" in materia civile va verificato riconducendone il grado di fondatezza all'ambito degli elementi di conferma disponibili in relazione al caso concreto (cd. *probabilità logica* o *baconiana*). Il criterio della probabilità prevalente fonda anche il sistema logico-operativo della prova presuntiva che è essenzialmente un ragionamento probabilistico per giungere alla conclusione "più probabile" (fatto ignoto) tra quante possono essere ipoteticamente tratte dalla premessa e cioè dal fatto noto». V. anche Cass. civ., Sez. III, 20 febbraio 2015, n. 3390; Cass. civ., sez. III, 26 febbraio 2013, n. 4792; Cass. civ., Sez. III, 14 giugno 2011, n. 12961. In dottrina, v. G. Canzio, *La valutazione della prova scientifica fra verità e ragionevole dubbio*, in AA.VV., *Scienza e processo penale. Nuove frontiere e vecchi pregiudizi*, a cura di C. Conti, 2011, p. 61 ss., per il quale: «La probabilità – s'intende "logica", *baconiana*, qualitativa – nell'indicare il grado di conferma oggettiva dell'ipotesi

Invece, il nostro Newman nella *Grammatica dell'assenso* (19):

– inizia, separando le proposizioni "*razionali*", dalle reali (parendo evocare la distinzione pascaliana tra *esprit de géométrie* ed *esprit de finesse*);

– prosegue esaltando il valore del "*senso dell'illazione*", con sorprendente riferimento a giuristi (20);

– quindi, infine e soprattutto (e qui sta il motivo dell'interesse degli AA. nei suoi confronti), egli conclude con un *coup de théâtre*, usando la probabilità, poiché sentenzia che «*a darci la certezza d'una cosa basta la convergenza di probabilità diverse, le quali, presa una a una e singolarmente, non sono nulla più che mere probabilità*» (21).

Il cuore del ragionamento di Newman è il seguente: se si hanno più indizi, nessuno con peso autonomo adeguato, nessuno d'essi, da solo, potrà convincere la nostra ragione.

Se invece consideriamo tutti gl'indizi congiuntamente, se rammentiamo la loro indipendenza, se ne aggreghiamo opportunamente i pesi di ciascuno, si può realizzare la trasformazione da indizi in prova.

Ricordiamo l'adagio reso popolare dal grande probabilista Bruno de Finetti (22): «*tre indizi fanno una prova*». Diciamo anche sùbito che, ragionando in termini di probabilità, la sintesi d'indizi non è un *optional*, ma mera conseguenza della logica dell'incertezza.

Come vedremo meglio nel cuore dell'articolo, il cardinale Newman suggerisce una strada in alternativa per la valutazione della probabilità di colpevolezza utilizzando la probabilità di innocenza, complice una valutazione notevolmente più semplice della prima, per infine ottenere la probabilità di colpevolezza per complemento.

Forse J.H. Newman non poteva immaginare che la sua dissertazione sulla prova dell'esistenza di Dio sarebbe stata oggetto d'attenzione matematica in dettaglio, da parte di de Finetti, che coniò l'espressione "Problema del cardinale Newman" e aprì la porta alla rivisitazione dal teologico al ... penale, passando, come vedremo, dal civile.

O forse il nostro era consapevole d'aver appoggiato il piede su un punto cruciale (23).

Passiamo alla rivisitazione.

———

formulata in ordine allo specifico fatto da provare, a fronte della pluralità dei possibili schemi esplicativi, contiene la verifica aggiuntiva, "sulla base dell'intera evidenza disponibile", dell'attendibilità dell'impiego della legge scientifica o statistica per il singolo evento e della persuasiva, elevata credibilità razionale dell'accertamento, che deve reggere alla prospettiva di falsificazione e agli elementi di prova antagonisti a sostegno delle controipotesi».

(19) J.H. Newman, *An Essay in Aid of a Grammar of Assent*, Burns, Oates & Co., 1870; trad. it.: *Grammatica dell'assenso*, Jaca Book, 2005.

(20) «*We often hear of the exploits of some great lawyer, judge or advocate, who is able in perplexed cases, when common minds see not but a hopeless heap of facts, foreign or contrary to each other, to detect the principle which rightly interprets the riddle, and, to the admiration of all hearers, converts a chaos into an orderly and luminous whole. This is what is meant by originality in thinking: it is discovery of an aspect of a subject-matter, simpler, it may be, and more intelligible than any hitherto taken*» (J.H. Newman, *op. cit.*, cap. 9, § 3, *The Range of the Illative Sense*).

(21) Analogamente: «*This portion of the work finished, I proceed to justify certitude as exercised upon a cumulation of proofs, short of demonstration separately*» (J.H. Newman, *op. cit.*, note 2, *On the alternative intellectually between Atheism and Catholicity*.

(22) B. de Finetti, *Teoria delle probabilità*, vol. I, Einaudi, 1970, cap. 4.15.4, p. 187 ss.

(23) «*And whereas the certitude is viewed by the Judge as following on converging probabilities, which constitute a real, though only reasonable, not an argumentative, proof, so it will be observed in this particular instance, that, in illustration of the general doctrine which I have laid down, the process is one of "line upon line, and letter upon letter," of various details accumulating and of deductions fitting into each other; [...] just as we see that two straight lines are meeting, and are certain they will meet at a given distance, though we do not see the junction*» (J.H. Newman, *op. cit.*, cap. 8, § 2, *Informal Reference*).

Disponiamo d'una serie d'indizi di colpevolezza, ciascuno con un suo "valore probante". Ci chiediamo quale sia il "valore probante" dell'insieme degli indizi.

Esempio 1 *Disponiamo di due indizi di colpevolezza, il primo con valore probante 40%, il secondo con valore probante 70%. Quanto prova l'insieme dei due?*

Bisogna subito chiarire la nozione, un po' elusiva, di "valore probante".

Visto che:

– si tratta di numeri;
– stiamo cercando di ragionare in condizioni d'incertezza;
– lo strumento appropriato è la probabilità (24);

potremmo provare a usare proprio la probabilità.

Il tema è delicato perché ci si può avvicinare a quest'oggetto da diverse direzioni:

– *Classica*, P.S. de Laplace (25):

$$\text{probabilità} = \frac{\text{numero dei casi favorevoli}}{\text{numero dei casi equipossibili}}$$

e qui si sente puzza di bruciato, perché nel definire la probabilità, s'usa la probabilità stessa attraverso la nozione di casi equipossibili, cioè, con uguale probabilità (26);

– *Statistica*, R. von Mises (27):

$$\text{probabilità} \approx \frac{\text{numero di successi su tante prove}}{\text{numero di prove}}$$

ovvero il rapporto fra il numero di successi quando ripeto "tante" volte le prove e il numero totale di prove effettuate. La domanda è: "Quando possiamo definire il numero delle prove 'tante'?", e per questo usiamo il ≈ e non l'=. Ma, anche qui, c'è uno sgradevole sentore: infatti la domanda "E se il numero di prove che abbiamo non è elevato?" inchioda;

– *Soggettiva*, de Finetti, Ramsey (28): la probabilità "vera" non esiste: essa è un semplice riassunto per scopi operativi dello stato d'informazione che un soggetto ha. Soggetti diversi, con stati d'informazione diversi, danno probabilità diverse. Sorprendente? Speriamo di no. C'è una tecnologia per far dichiarare opinioni non distorte:

———

(24) Avremmo voluto scrivere che, di fronte all'incertezza, la Probabilità è il *top* di gamma.

(25) P.S. de Laplace, *Théorie Analytique des Probabilités*, 1812.

(26) La versione latina di definizione di Aristotele suona: *definitio fit per genus proximum et differentiam specificam*, e non concorda granché con l'approccio classico alla probabilità.

(27) R. von Mises, *Wahrscheinlichkeit, Statistik und Wahrheit: Einführung in die neue Wahrscheinlichkeitslehre und ihre Anwendung*, Springer-Verlag, 1928.

(28) B. de Finetti, *op. cit.*: F.P. Ramsey, *Truth and Probability*, 1926, in F.P. Ramsey, *The Foundations of Mathematics and other Logical Essays*, 1931, Ch. VII, p.156 ss., edited by R.B. *Braithwaite*, Kegan, Paul, Trench, Trubner & Co., Harcourt, Brace and Company. Si trova un importante riscontro in G. Ubertis, *Profili di Epistemologia Giudiziaria*, Giuffrè, 2015, p. 3, in tema di verità giudiziale, ne precisa la dipendenza dallo stato d'informazione. È esattamente la stessa piattaforma dell'approccio soggettivista.

$$\text{probabilità} = \frac{\text{posta giudicata equa}}{\text{ammontare della vincita}}$$

senza sapere prima se si sarà banco o scommettitore (29).

Nel contesto che c'interessa, il pericolo di valutazioni distorte non rileva granché, perché le probabilità che interessano, sono normalmente fornite *ab extra* e (sperabilmente) affidabili (30).

Gli approcci classico e statistico sovrappongono colpevolmente due questioni diverse:

– come posso, in certi casi, valutare una probabilità?

– che cosa è la probabilità?

C'interessa sottolineare che valutazioni di tipo probabilistico (31) fanno parte della vita quotidiana di ognuno di noi. Entrano in scena inconsapevolmente in tante scelte che effettuiamo tutti i giorni. Studiosi e matematici si sono occupati di dare un "ordine" al mondo aleatorio e ad ora ne esiste una modellazione robusta capace di contribuire significativamente a molti campi della ricerca scientifica. In questo contesto ci limitiamo a richiamare solo gli elementi essenziali, il lettore può riferirsi all'appendice per nozioni di base di probabilità. Ricordiamo che la probabilità è sempre un numero compreso fra 0 e 1, solitamente indicato in percentuale, e che la probabilità di un evento è uguale a 1 meno la probabilità dell'evento *complementare* (32) (detto anche *opposto*).

I tre approcci classico, statistico, soggettivista sono detti *completi* perché, in maniera differente, suggeriscono sia come valutare le probabilità, sia quali calcoli si possano fare sulle probabilità. I tre approcci completi generano, in sostanza, la stessa aritmetica sulle probabilità, differiscono essenzialmente nel come valutare le probabilità. È molto popolare un quarto approccio, detto *assiomatico*. L'approccio assiomatico, proposto da un grandissimo matematico russo del secolo scorso, A.N. Kolmogorov (33), si lava le mani del problema della valutazione delle probabilità: semplicemente codifica il calcolo sulle probabilità. Per i nostri scopi questo punto di vista *incompleto* non è di speciale interesse, appunto perché non gestisce il problema di "come dare le probabilità".

Per completezza e consapevolezza, se ne trova traccia in appendice.

2.2. Probabilità dinamiche

In quest'articolo siamo particolarmente interessati a studiare il valore probante degli indizi, ovvero alla probabilità di colpevolezza o innocenza avendo in mano un certo numero di indizi. Matematicamente questo *valore probante* corrisponde ad una quantità ben definita che vogliamo adoperare. Si tratta di *probabilità condizionata* della colpevolezza o innocenza sapen-

———

(29) È un espediente che viene dalle *Mille e una notte*, per evitare che siano fatte dichiarazioni distorte per motivi d'interesse.

(30) Nella letteratura *Law and Economics*, ma non solo, sono assegnate dalle relazioni peritali.

(31) La nostra posizione, un po' da *outsider*, ci fa parere estranei, ma, anche all'interno delle discipline giuridiche sulla questione rileviamo un chiaro e condivisibile "taglio": G. Ubertis, *op. cit.*, p. 4 col magistrale titolo della sezione 2: *L'illusione della "conoscenza giudiziaria oggettiva"*. Al di là del titolo, condividiamo la lucida sintesi di questioni niente affatto banali.

(32) Per esempio la probabilità d'ottenere al lancio di un dado un numero pari sarà uguale a 1 meno la probabilità d'ottenere un numero dispari. Se il dado non è truccato, le probabilità di pari e dispari saranno entrambe 1/2, cioè 50%.

(33) A.N. Kolmogorov, *Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung*, in *Mathematische Annalen*, 1931, n. 104, p. 415 ss.

do che si sono raccolti un certo numero di indizi. La maniera soggettiva d'affrontare la probabilità è la migliore, *in primis*, perché non esclude gli altri procedimenti di valutazione, ma li confina operativamente. Essa porta con sé, come cruciale, la nozione di *probabilità condizionata*, che è nel fulcro del nostro contributo.

L'idea basilare è che le valutazioni di probabilità dipendono dallo stato d'informazione. Se esso evolve, non è affatto innaturale che la probabilità evolva.

La probabilità misura il grado di fiducia che meritano gli eventi. Per es., l'evento "Tizio è colpevole", che chiamiamo C (come "colpevole"), è denotata dai matematici con:

$$\Pr(C)$$

simbolo che dovrebbe autospiegarsi.

Se uno non sa nulla, circa la colpevolezza, può tradurre l'ovvia descrizione del suo punto di vista così: "Non ho motivi per ritenere colpevole C più probabile di I, come 'innocente', quindi, per me:

$$\Pr(C) = \Pr(I) = 50\%\text{"}$$

Cambia lo stato d'informazione, arriva il referto d'un perito, che chiamiamo R, come "referto". Lo sviluppo del discorso non muta, anche se assumiamo che lo stato di informazione cambi in séguito dell'escussione di testimoni e di qualsiasi (T) e di qualsiasi altra acquisizione rilevante (si pensi ad esempio alle immagini registrate di una videocamera di sorveglianza). Ci possiamo aspettare che la valutazione probabilistica cambi. Anzi, in linea di principio, ce lo auguriamo. Al centro di questo articolo è la valutazione di colpevolezza, dati uno o una serie di referti. Per la colpevolezza, tale probabilità è denotata con:

$$\Pr(C \mid R)$$

per l'innocenza, con:

$$\Pr(I \mid R)$$

Il primo simbolo si legge "*probabilità (condizionata)* di C dato R", l'evento C si dice *condizionato*, l'evento R si dice *condizionante*. Per il secondo analogamente.

Torniamo sùbito al punto dell'evoluzione dello stato d'informazione. Un approccio ingenuo alla probabilità, peraltro affatto naturale, di fronte alla gestione della transizione dalla *probabilità iniziale* alla *probabilità finale*:

$$\Pr(C) \rightarrow \Pr(C \mid R)$$

si rifugia nel "Mah! Chissà!" e deputa a inconoscibili processi mentali tale revisione d'opinioni. Il terzo Autore ha empiricamente osservato tale atteggiamento mentale, durante un seminario, tranquillamente dichiarato da parte d'un docente di discipline giuridiche.

La logica binaria del vero e del falso (ossia la logica dello 0, 1) ha una sua reputazione rispettata anche in àmbito giuridico, mentre la sua naturale estensione all'incerto (ossia la probabilità) appare spesso un po' stregonesca.

Il punto, non sempre tenuto presente, è che la "nuova probabilità", per es. $\Pr(C \mid R)$, non è da valutare *ad libitum*, ma è semplice frutto di calcoli elementari su "vecchie probabilità" (ovvero elementari probabilità non condizionate).

Detto diversamente: mescolare informazioni probabilizzate in partenza con nuova evidenza empirica non abita nel regno della creatività, o tra processi mentali inconoscibili, ma nell'aritmetica elementare.

Illustriamo la questione attraverso il semplice:

Esempio 2 *Pensiamo a un dado non truccato: a ciascuna delle sei facce*

$$1, 2, 3, 4, 5, 6$$

attribuiamo la stessa probabilità d'uscita, pari a 1/6. *Indichiamo con* X *il punteggio ottenuto buttando il dado. A noi interessa lo specifico evento (risultato)* $X = 4$, *cui, poco sopra, abbiamo dato probabilità* 1/6*:*

$$\Pr(X = 4) = \frac{1}{6}$$

Il dado è gettato, il risultato X *è registrato, ma non lo conosciamo. Sul risultato X abbiamo solo un'informazione parziale: per esempio, sappiamo che il risultato è maggiore di 2, che si scrive* $X > 2$. *Ossia* X *è* 3 *o* 4 *o* 5 *o* 6. *L'intuizione suggerisce che l'informazione, ancorché parziale, dovrebbe modificare la nostra valutazione di probabilità, passando dalla "probabilità che* X *sia uguale a* 4*" alla "probabilità che* X *sia uguale a* 4*, sapendo anche che* $X > 2$*". In altre parole passiamo*

$$\textit{da } \Pr(X = 4) = \frac{1}{6} \;\; \textit{a} \;\; \Pr(X = 4 \mid X > 2) = ?$$

Per calcolare $\Pr(X = 4 \mid X > 2)$, *in questo contesto, è naturale ragionare dicendo che dei 6 casi possibili ne restano in ballo solo* 4, *onde:*

$$\Pr(X = 4 \mid X > 2) = \frac{1}{4}$$

Avremmo potuto raggiungere lo stesso risultato con uno strumento di ragionamento più generale (34).

La sua importanza sta nel fatto che la probabilità condizionata di rilievo:

$$\Pr(C \mid R)$$

---

(34) Da cui nasce il teorema di Bayes, ormai empiricamente rilevante in àmbito giuridico.

è semplicemente quoziente o rapporto di due probabilità non condizionate. Scriviamo con una formula il punto principale:

$$\Pr(C \mid R) = \frac{\Pr(C \cap R)}{\Pr(R)} \qquad (1)$$

dove la scrittura "$C \cap R$" sta per l'occorrenza congiunta della colpevolezza di Tizio e dell'esito R del referto. Al numeratore sta la probabilità attribuita al verificarsi congiuntamente di due eventi: la colpevolezza C di Tizio e l'esito R del referto; al denominatore sta la probabilità attribuita all'esito del referto R. Vediamo sùbito come la valutazione di quest'ultima si possa appropriatamente articolare.

Moltiplicando ambo i membri dell'uguaglianza sopra scritta per la probabilità del referto, s'ottiene una formula d'interesse generale:

$$\Pr(C \mid R) \times \Pr(R) = \Pr(C \cap R) \qquad (2)$$

riscrivibile, per comodità espositiva, come:

$$\Pr(C \cap R) = \Pr(C \mid R) \times \Pr(R)$$

che possiamo spiegare a parole, anche abbandonando l'interpretazione che qui ci ha condotto di C e di R.

La formula ci dice che la probabilità che due eventi si presentino congiuntamente è uguale al prodotto della probabilità che uno d'essi (per es. R) s'avveri per la probabilità $\Pr(C \mid R)$, che l'altro s'avveri, condizionata al fatto che il primo si sia avverato [35].

Questa formula è la versione vera e genuina del c.d. *principio della probabilità composta*. Riportiamo, ma segnaliamo al lettore che la nozione di "principio", per es., adorata in Fisica, è nozione scivolosa: non si sa bene se un principio sia vero perché può essere dimostrato (e allora meglio sarebbe chiamarlo teorema) o perché se ne postula la sussistenza (e allora sarebbe meglio chiamarlo assioma). Nel nostro contesto viaggiamo tranquilli, perché riesce essere un teorema.

Vediamo subito un esempio:

Esempio 3 *Riprendiamo l'esempio 2. Richiedere congiuntamente il verificarsi di due eventi "il punteggio è 4" con "il punteggio è maggiore di 2" equivale al solo evento "il punteggio è 4" perché il primo è più restrittivo del secondo e, quindi, la loro intersezione coincide col primo. La probabilità che entrambi si manifestino è allora la probabilità che esca 4, cioè 1/6. La probabilità che si manifesti l'evento condizionante (punteggio maggiore di 2) è 4/6, come risulta applicando banalmente la formula (1), che ci dice che il rapporto tra le due probabilità in giuoco è 1/4. In formule:*

———

(35) Un *punctum dolens* a questo proposito è se si possa parlare di $\Pr(A \mid B)$, solo una volta che B si sia verificato. Tali valutazioni di probabilità sono meramente congetturali. A parole: qual è la probabilità da dare ad A, supponendo (che si sia provato adeguatamente) che B si è verificato.

$$\Pr (X = 4 \mid X > 2) = \frac{\Pr [(X = 4) \cap (X > 2)]}{\Pr (X > 2)} = \frac{1/6}{4/6} = \frac{1}{4}$$

come prima ottenuto.

Il lavoro svolto sopra ci consente di fare due passi ulteriori: (1) *articolare il calcolo della probabilità d'ottenere un dato risultato* da un referto; (2) introdurre la nozione d'*indipendenza* (stocastica) *d'eventi*, introdotta oltre in 2.4.

2.3. Probabilità d'un risultato

Sopra, nella formula (1), abbiamo usato il simbolo Pr (R) per la probabilità d'ottenere un certo risultato R da un referto, indipendentemente da colpevolezza o innocenza di Tizio. Vediamo ora come tale probabilità possa essere valutata.

Il risultato R è compatibile sia con C sia con I. La sua probabilità può essere ottenuta, usando la formula (2), sommando le due probabilità di "R ∩ C" e "R ∩ I" (36):

$$\begin{aligned} \Pr (R) &= \Pr (R \cap C) + \Pr (R \cap I) = \\ &= \Pr (R \mid C) \times \Pr (C) + \Pr (R \mid I) \times \Pr (I) \end{aligned} \tag{3}$$

Esempio 4 *Una certa malattia non endemica è diffusa nella popolazione con una composizione tra affetti (A)*(37) e sani (S) (38) . *Solo 1% della popolazione è affetto. In formule* (39):

$$\Pr (A) = 1\% \text{ e } \Pr (S) = 99\%$$

Si è sviluppato un test clinico per individuare la presenza o meno di tale malattia. L'esame clinico dà il risultato positivo P: si sa che il test non funziona perfettamente ma è un buon test. La probabilità che il test sia P nel caso in cui Tizio sia effettivamente affetto (A) è:

$$\Pr (P \mid A) = 98\%$$

La probabilità che il test dica P nel caso in cui Tizio sia sano (S) è valutata 2% (40):

$$\Pr (P \mid S) = 2\%$$

Immaginiamo di pescare a caso un soggetto in una popolazione mista e di testarlo. Qual è la probabilità che il test riesca positivo, indipendentemente dalla correttezza dell'indicazione? Facciamo un piccolo stop e chiediamo al lettore di dare un numero. Nell'esperienza del terzo autore la risposta 98% è prediletta ed è basata su un ben preciso pseudo-ragionamento

———

(36) Per i lettori a conoscenza, qui si giocano i "falsi positivi", ormai ricorrenti in molti procedimenti con rilievo penale.

(37) *A* come "affetto".

(38) *S* come "sano".

(39) Qui la somma a 1 delle probabilità *necesse est*.

(40) Non vi è nessun obbligo che la somma delle due ultime probabilità sia 1 perché non sono eventi complementari, benché si tratti di probabilità condizionate a eventi complementari, cosa ben diversa. Qualora si verificasse una somma unitaria essa sarebbe meramente casuale, come nei *disclaimer* dei titoli di coda dei *film*.

intuitivo: "Il test funziona al 98%, sia nell'accertare, sia nel non accertare, perciò la probabilità che il test funzioni è 98%". In contesti similari gl'ingegneri – sempre testimone il terzo Autore – dicono 96.04%, perché sono ingegneri e perché 98% x 98% = 96.04%. Sbagliano tutti e non di poco, perché il risultato cui giungiamo usando la formula (3), sconvolgente per gl'ingenui, è:

$$\Pr(P) = \Pr(P \mid A) \times \Pr(A) + \Pr(P \mid S) \times \Pr(S) =$$
$$= 98\% \times 1\% + 2\% \times 99\% = 2{,}96\%$$

Con la probabilità non si scherza. La lista d'aggettivi di cui essa è corredata in dottrina, ma, soprattutto, in giurisprudenza, pensiamo possa essere semplicemente spia di malcerta famigliarità.

Non vanno esenti da ciò nemmeno i matematici: rammentiamo il popolare paradosso delle tre porte, o di Monty Hall (41). Ma qui esso non è cruciale. Esso rileva maggiormente, per es., nel classico problema di "riconoscimento"del parlatore attraverso una perizia fonica (42).

2.4. Indipendenza (stocastica)

Estremizziamo per spiegarci. Al solito: C sta per Tizio è colpevole. Attribuiamo all'evento una certa probabilità Pr (C). Ci arriva il risultato R d'un referto e siamo chiamati a rivedere (eventualmente) la valutazione di probabilità:

$$\Pr(C) \rightarrow \Pr(C \mid R)$$

Il risultato del referto è: "uno zio di Tizio ha fatto il morbillo e ne è uscito". Siamo contenti per lo zio, ma per noi:

$$\Pr(C \mid R) = \Pr(C) \tag{4}$$

Quando ciò accade, si dice che l'evento R è giudicato stocasticamente (43) indipendente da C.

Sensazionale è che, se le probabilità date rivelano questo giudizio di indipendenza espresso in (4), allora, il sedicente *principio delle probabilità composte* riesce enormemente semplificato, infatti la (2) si riduce a:

$$\Pr(C \cap R) = \Pr(C) \times \Pr(R) \tag{5}$$

———

(41) Ci sono tre porte chiuse. Due di esse vuote e una contenente un tesoro. Si sceglie una delle tre porte inizialmente. La domanda è se convenga o meno cambiare la scelta fatta una volta che si è a conoscenza che una delle due porte non scelte è vuota. Si tratta abbastanza facilmente la questione di valutare la probabilità di scelta del tesoro usando la probabilità condizionata. La risposta è sì. Anche se controintuitiva, si passa da una scelta equiprobabile quindi di 1/3, a una scelta con probabilità condizionata di 2/3. La probabilità di vincita raddoppia quando si condiziona al fatto di sapere che una delle porte non aperte è vuota.

(42) Una buona sintesi si trova in M. Biral, *L'identificazione della voce nel processo penale: modelli, forme di accertamento, tutela dei diritti individuali*, in *Riv. it. dir. e proc. pen.*, 2015, p. 1842 ss. Per alcuni aspetti rilevanti in processi nel tribunale di Torino, sono risultati significativi: L. Peccati, R. Piazza, *Applicazione del Teorema di Bayes a un problema giudiziario*, in *Atti del quattordicesimo Convegno AMASES*, 1990 e L. Peccati, R. Piazza, *Applicazione del teorema di Bayes al problema del riconoscimento del parlatore*, in *Atti del XIX Convegno Annuale dell'Associazione Italiana di Acustica*, 1991.

(43) Vuol dire solo "in probabilità".

che è la versione più popolare del principio: la probabilità che due eventi si presentino insieme è, banalmente, il prodotto delle probabilità di ciascuno dei due eventi.

La nozione di probabilità condizionata è cruciale, anche perché conferma l'appartenenza della probabilità al suo mondo consolidato: è il riassunto operativo d'un dato stato d'informazione, che può variare nel tempo.

In generale, nei problemi d'interesse per questo articolo, ci concentriamo sulla valutazione della probabilità che un certo evento si sia verificato, condizionata a un dato stato d'informazione, che è un po' la situazione nella quale quotidianamente ci troviamo.

La probabilità che due eventi A, B si verifichino insieme riesce dunque (44):

$$\Pr(A \cap B) = \Pr(A \mid B) \times \Pr(B) = \Pr(B \mid A) \times \Pr(A)$$

A parole, la probabilità che due eventi A, B si verifichino insieme è prodotto di due probabilità:

– la probabilità che uno di essi – mettiamo – B si verifichi;

– la probabilità condizionata che, dato B, si verifichi l'altro.

Consideriamo il seguente:

Esempio 5 *Brutalmente, qual è la probabilità che un soggetto S, in una data popolazione, fumi. C'interessa valutare la probabilità che il soggetto, femmina, fumi. Chiamiamo B l'evento "il soggetto è una soggetta, cioè femmina". Abbiamo statistiche sulla composizione per sesso di quella popolazione. Sulla base delle stesse attribuiamo probabilità* Pr *(B) al sesso femminile del soggetto. Per es.,* Pr *(B)* = 45%. *Disponiamo anche di statistiche sull'abitudine di fumare delle signore. Denotiamo con A l'evento "Tizio fuma", cosicché l'evento A | B è l'evento "Tizia fuma". Sulla base delle statistiche di cui sopra, valutiamo* Pr *(A | B)* = 50%. *La probabilità che un soggetto sia femmina e che fumi è strettamente vincolata alle probabilità precedenti:*

$$\Pr(A \cap B) = \Pr(A \mid B) \times \Pr(B) \tag{6}$$

ossia: *la probabilità che un individuo/a sia femmina (B) e che fumi (A) deve essere:*

$$\Pr(A \cap B) = 50\% \times 45\% = 22.5\%$$

———

(44) Dall'uguaglianza:

$$\Pr(A \mid B) \times \Pr(B) = \Pr(B \mid A) \times \Pr(A)$$

risolvendo rispetto a Pr ($A \mid B$), si trova:

$$\Pr(A \mid B) = \Pr(A) \times \frac{\Pr(B \mid A)}{\Pr(B)}$$

che è la formula di Bayes. Nel primo membro c'è la "nuova probabilità" Pr ($A \mid B$), nel secondo la "vecchia" Pr (A) moltiplicata per un appropriato fattore d'adeguamento che rapporta la probabilità d'osservare B subordinatamente all'evento d'interesse e la probabilità d'osservare B comunque.

Una situazione interessante riguarda il caso d'*indipendenza stocastica*.

Con riferimento alle righe sopra, continuiamo a indicare con A l'evento che Tizio fumi. Cambiamo però B. Esso è ora ridefinito come "Tizio è nato in un giorno dispari del suo mese di nascita". Non ci risultano oroscopi basati sulla parità del giorno di nascita. L'attribuzione a un segno zodiacale sarebbe ben più grossolana, ancorché goda d'ampio seguito.

Assumiamo coraggiosamente che il lettore ritenga irrilevante la parità del giorno di nascita. In tal caso, l'informazione aggiuntiva sulla parità mensile del giorno di nascita non dovrebbe implicare modificazioni nelle valutazioni di probabilità e, come nella formula (4), s'avrebbe:

$$\Pr(A \mid B) \times \Pr(A) \qquad (7)$$

onde, come nella (5)

$$\Pr(A \cap B) = \Pr(A) \times \Pr(B) \qquad (8)$$

Riprendiamo l'Esempio 1 e consideriamo questo:

Esempio *Chiamiamo i_1 e i_2 i due indizi. Al primo è assegnato valore probante $Pr(C \mid i_1) = 40\%$; al secondo $Pr(C \mid i_2) = 70\%$. Possiamo facilmente dedurre che: $Pr(I \mid i_1) = 1 — 40\% = 60\%$ e $Pr(I \mid i_2) = 1 — 70\% = 30\%$. Poiché, in ipotesi d'innocenza, i due indizi sono stocasticamente indipendenti, la probabilità del loro verificarsi congiuntamente è il prodotto delle loro probabilità di verificarsi: $Pr(I \mid i_1, i_2) = Pr(I \mid i_1)$ x $Pr(I \mid i_2) = 60\%$ x $30\% = 18\%$. Concludiamo che il valore probante dell'insieme dei due indizi è $Pr(C \mid i_1, i_2) = 1 — Pr(I \mid i_1, i_2) = 1 — 18\% = 82\%$.*

I lettori più attenti avranno notato che in tale esempio il prodotto delle probabilità è effettuabile solo per eventi tra loro (stocasticamente) indipendenti. Ciò accade per l'ipotesi dell'innocenza ma non per quella della colpevolezza. Non sarebbe dunque affatto corretto calcolare $\Pr(C \mid i_1, i_2) = 40\%$ x $70\% = 28\%$.

La riproponiamo a parole. Un indizio ha valore probante *tot*%, un altro ha valore probante *tit*%. Quanto valgono i due indizi insieme? Sotto l'ipotesi di colpevolezza, per combinare le probabilità generate occorrerebbe tener conto delle probabilità di uno, dato l'altro, e viceversa.

Il principio del cardinale Newman suggerisce una strada alternativa: valutare la probabilità d'innocenza, dati degl'indizi indipendenti, e poi, per complemento, ottenere la probabilità di colpevolezza.

Partiamo da un semplice:

Esempio Il solo percorso che da una strada pubblica porta al capannone ove è stato trovato un assassinato comporta il passaggio su due parti d'uno spiazzo. La prima è quasi coperta di segatura, la seconda di gesso. I RIS analizzano le suole delle scarpe d'un signore e vi rinvengono sia segatura, sia gesso.

Come valutare la probabilità di colpevolezza di questo signore dalla congiunta presenza di segatura e gesso? I due eventi sono dipendenti se quel signore è colpevole, perché è dovuto passare su ambo le parti dello spiazzo, ma sono indipendenti se quel signore è innocente ed è andato a trovare, cosa che fa raramente, anche se in tempi decentemente lontani, le sue due "adorate" zie, un po' maniache, di cui la prima adora la falegnameria e produce montagne di segatura, mentre l'altra si diletta di sculture in gesso, con conseguente gessatura del vialetto d'accesso.

Le presenze di segatura e gesso, nel secondo caso, sono indipendenti. Possiamo pertanto moltiplicare le loro probabilità d'occorrenza sotto innocenza.

Cerchiamo di spiegarci. Siamo interessati a calcolare la probabilità di colpevolezza dell'imputato Tizio sapendo che abbiamo due indizi i_1 e i_2. Tale valore non risulta di facile valutazione. Al contrario la probabilità di innocenza risulta relativamente facile da stabilire. Concludiamo deducendone per differenza la probabilità di colpevolezza che aveva motivato la nostra analisi. Infatti vale la seguente formula:

$$\Pr(C \mid i_1, i_2) = 1 - [1 - \Pr(I \mid i_1)] \times [1 - \Pr(I \mid i_2)]$$

Qui sta il cuore del ragionamento del cardinale Newman. Proviamo a convincercene riprendendo l'argomento (quasi) a parole. Pensiamo a un caso in cui, contro un innocente (evento I), siano presentati due indizi giudicati indipendenti i_1, i_2. Ciò significa che gli eventi:

– contro l'innocente si presenta l'indizio i_1, evento che un matematico indicherebbe con $i_1 \mid I$;

– contro l'innocente si presenta l'indizio $i2$, evento che un matematico indicherebbe con $i_2 \mid I$;

sono stocasticamente indipendenti.

Segue che $I \mid i_1$ e $I \mid i_2$ lo sono del pari e, quindi, la probabilità d'innocenza dati i due indizi è

$$\Pr(I \mid i_1, i_2) = \Pr(I \mid i_1)] \times \Pr(I \mid i_2)] \tag{9}$$

Poiché la probabilità d'innocenza è il complemento a 1 della probabilità di colpevolezza, allora, da:

$$\begin{cases} \Pr(I \mid i_1, i_2) = 1 - \Pr(C \mid i_1, i_2) \\ \Pr(C \mid i_1) = 1 - \Pr(I \mid i_1) \\ \Pr(C \mid i_2) = 1 - \Pr(I \mid i_2) \end{cases}$$

segue:

$$\Pr(C \mid i_1, i_2) = 1 - [1 - \Pr(I \mid i_1)] \times [1 - \Pr(I \mid i_2)] \tag{10}$$

Se chiamiamo, per brevità $p1$, $p2$ le due probabilità d'innocenza:

$$\Pr(I \mid i_1); \Pr(I \mid i_2)$$

otteniamo:

$$\Pr(C \mid i_1, i_2) = 1 - (1 - p_1) \times (1 - p_2)$$

Quest'argomentazione può essere estesa al caso generale di n indizi:.

In presenza di n indizi indipendenti, ciascuno con un suo valore probante, otteniamo la celebrata *formula del cardinale Newman*:

$$\Pr(C \mid i_1, i_2, \ldots, i_n) = 1 - (1 - p_1) \times (1 - p_2) \times \ldots \times (1 - p_n) \tag{11}$$

Non tediamo il lettore con i dettagli tecnici, ma, nel caso di numerosi indizi, tutti con lo stesso valore probante, c'è un fatto interessante. La probabilità di colpevolezza cresce col numero d'indizi (com'è ovvio), ma meno che proporzionalmente rispetto al numero d'indizi. Tutto ciò è naturale e concorda con la letteratura (45).

Un'ultima osservazione: se un indizio, mettiamo, il numero 2 è inchiodante, ossia ha valore probante $p_2 = 1$, allora la formula del cardinale ci dà:

$$\begin{aligned}\Pr(C \mid i_1, i_2, \ldots, i_n) &= 1 - (1 - p_1) \times (1 - 1) \times \ldots \times (1 - p_n) = \\ &= 1 - (1 - p_1) \times 0 \times \ldots \times (1 - p_n) = 1 - 0 = 1\end{aligned}$$

come deve essere (46).

3. UN'INTERPRETAZIONE GIUDIZIARIA

Lo schema logico del cardinale Newman può anche interessare la scelta di strategie processuali.

Tratteggiamo il problema.

In apertura d'un processo, un avvocato può porre mettiamo 2 eccezioni e_1, e_2 alla continuazione dello stesso.

Ciascuna è indipendente dall'altra e ha probabilità di successo W (W come "*Win*") (47) assegnata:

$$\begin{cases} p_1 = \Pr(W \mid e_1) \\ p_2 = \Pr(W \mid e_2) \end{cases}$$

Vogliamo valutare la probabilità di successo per l'avvocato d'un paniere di due eccezioni.

Conviene dire subito che l'idea di sommare i *valori frenanti* (48) è peregrina.

Esempio 6 *L'avvocato può presentare due eccezioni* e_1, e_2, *con* valore frenante *rispettivo*:

$$p_1 = \Pr(W \mid e_1) = 40\% \;\text{ e }\; p_2 = \Pr(W \mid e_2) = 70\%$$

Qual è la probabilità di successo $\Pr(W \mid e_1, e_2)$*? È evidente che per essa:*

———

(45) V. M.O. Finkelstein, *Quantitative Methods in Law*, The Free Press (Mac-Millan), 1978, e/o R.D. Friedman, *The Science of Conjecture*, Johns-Hopkins University Press, 2001. Il volume più recente sul tema è: M.O. Finkelstein, B. Levin, *Statistics for Lawyers*, 3rd ed., Springer, 2015. Si trova un'introduzione elementare in M. D'Amico, L. Peccati, *Metodi matematici, statistici e finanziari per giuristi*, 7^ ed., EGEA, 2016.

(46) Se in un prodotto, un fattore è nullo, anche il prodotto lo è. Quest'osservazione supporta un punto di vista generale: il ragionamento probabilistico non è contra il ragionamento logico, basato sul vero e sul falso, ma ne è semplicemente un'estensione. Quando, per avventura, un problema che nasce nel mondo dell'incertezza si derubrica nel mondo della certezza, quanto la logica "estesa" dell'incertezza fornisce coincide con quanto fornisce la logica della certezza.

(47) Avremmo potuto scegliere S, come "successo", già usato per "sano", e sarebbe analogamente seguito I per "insuccesso", lettera già buttata per "innocente", onde il passaggio all'inglese.

(48) Locuzione di nuovo conio, differentemente valutati da accusa e difesa. Ci mettiamo dal lato della difesa, che deve scegliere una strategia.

$$p_1 + p_2 = 40\% + 70\% = 110\%$$

non sta in piedi.

L'aggregazione degli stessi passa fatalmente attraverso lo schema del cardinale Newman.

Si passa alle probabilità d'insuccesso L (L come *Loss*) delle singole eccezioni:

$$\Pr(L \mid e_1) = 1 - \Pr(W \mid e_1) = 1 - p_1$$

e, analogamente, per la seconda eccezione:

$$\Pr(L \mid e_2) = 1 - \Pr(W \mid e_2) = 1 - p_2$$

L'indipendenza delle eccezioni ci porta alla probabilità d'insuccesso del paniere d'eccezioni. Quindi, nel caso dell'es. precedente, abbiamo:

$$\Pr(W \mid e_1, e_2) = 1 - (1 - p_1) \times (1 - p_2) = 1 - (1 - 40\%) \times (1 - 70\%) = 82\%$$

Le implicazioni della rilettura del principio, dal punto di vista strategico, sono di qualche interesse.

Gli AA. si riservano di studiare la conseguente strategia ottima, tenendo conto del costo del numero d'eccezioni.

Considerando un caso non inusuale ma ragionevole, che potrebbe difatti capitare, si può fornire un suggerimento che non è lontano da certe pratiche professionali. Non vogliamo dire che questo che si debba fare ma che in generale è di uso comune che si faccia.

Consideriamo il caso (non inusuale) d'indifferenza, anche se – magari – ignota, tra il valore frenante delle singole eccezioni. In tal caso, la probabilità di successo cresce col numero delle eccezioni, ma meno che proporzionalmente, onde il suggerimento strategico (che viene dai numeri, e non dal naso): eccepire poco, ma bene.

4. CONCLUSIONI

Siamo convinti che quanto precede possa essere d'interesse per il ragionamento giuridico in condizioni d'incertezza.

Speriamo d'avere stimolato pensieri sul tema, anche se non abbiamo insistito sul costrutto formale (nell'Appendice illustriamo qualche passo ulteriore).

Ci preme segnare un ultimo punto: le caratteristiche di accumulazione indiziaria implicite nello schema del cardinale Newman.

Accumulare informazioni è normale in ogni professione, in particolare nelle legali.

L'idea intuitiva d'un processo d'accumulazione è strettamente legata a una proprietà di operazioni aritmetiche elementari (addizione e moltiplicazione) che ciascun lettore ha incontrato, magari non sempre serenamente.

La prima si chiama *proprietà associativa*:

$$a + (b + c) = (a + b) + c \quad \text{e} \quad a\,(bc) = (ab)\,c$$

che legittima le scritture:

$$a + b + c \text{ e } abc$$

Per es.:

$$3 + 2 + 1 = 6 = (3 + 2) + 1 = 6 = 3 + (2 + 1)$$

o:

$$3 \times 2 \times 1 = (3 \times 2) \times 1 = 6 = 3 \times (2 \times 1)$$

Potremmo rappresentare la situazione in generale scrivendo:

$$a * b * c = (a * b) * c = a * (b * c)$$

ove "*" sta per "+" o "×".

Una domanda naturale (almeno per un matematico) è se esistano altre operazioni "associative", diverse da addizione e moltiplicazione, che potrebbero ben rappresentare un processo d'accumulazione. La risposta è che ne esistono infinite e che l'operazione che accumula i valori probanti p_1, p_2:

$$p_1 * p_2 = 1 - (1 - p_1) * (1 - p_2)$$

è una di esse.

Un ulteriore aspetto meritevole d'interesse va sotto il nome di *commutatività*: nell'addizione e nella moltiplicazione, l'ordine degli ingredienti non influisce sul risultato:

$$a + b = b + a \text{ e } ab = ba$$

per esempio:

$$5 + 1 = 1 + 5 = 6 \text{ e } 5 \times 2 = 2 \times 5 = 10$$

in sintesi:

$$a * b = b * a$$

L'operazione che "accumula" p_1 e p_2 è anche commutativa:

$$p_1 * p_2 = 1 - (1 - p_1) \times (1 - p_2) = 1 - (1 - p_2) \times (1 - p_1) = p_2 * p_1$$

È un'altra proprietà affatto naturale in un processo d'accumulazione.

Concludendo: la formula del cardinale Newman garantisce, per così dire, democrazia nell'accumulazione degl'indizi e irrilevanza logica dell'ordine con cui essi sono presentati.

5. APPENDICE: PILLOLE DI PROBABILITÀ (MATEMATICA)

In questa appendice ricordiamo sinteticamente alcune proprietà di base della probabilità utili ai fini della comprensione dell'articolo. Qualunque sia il tipo di probabilità cui ci si riferisce, la probabilità gode d'alcune proprietà, che rilevano anche per il nostro problema.

Riassumiamo:

– Indichiamo con X un esperimento aleatorio, con vari possibili esiti (per es., gettiamo un dado, i possibili esiti sono 1, 2, 3, 4, 5, 6).

– Chiamiamo Ω l'insieme dei possibili esiti (49). Chiamiamo *risultati elementari* i singoli possibili esiti.

– I sottoinsiemi di Ω li chiamiamo eventi, li denotiamo con lettere maiuscole e di essi ci interessa valutare la probabilità (50).

– Se chiamiamo *E* un evento, che può essere vero o falso, l'incertezza può dipendere dal fatto che *E* è futuro (51) o, anche se passato, le informazioni disponibili non sono sufficienti per stabilirne la veridicità (52). Ad ogni evento *E* posso associare *la* sua probabilità (53):

$$\Pr(E)$$

– Tutto l'insieme Ω è l'evento certo. Il suo complementare è l'insieme vuoto Ø, ed è l'evento impossibile.

– Un evento impossibile (54) ha probabilità 0.

– Un evento certo (55) ha probabilità 1.

– La maggior parte degli eventi interessanti non è né certa, né impossibile e ha probabilità tra 0 e 1.

– L'evento *A* "e" *B* corrisponde congiuntamente all'insieme intersezione dei due insiemi $A \cap B$.

– L'evento A "o" B corrisponde all'insieme unione dei due insiemi *A* U *B*, senza esaurire necessariamente lo spettro degli eventi che possono interessare.

– Frequentemente le probabilità sono espresse in percentuale: invece di dire $\Pr(E) = 0.30$ è comune dire che la probabilità di *E* è 30%. Niente di rotto: basta rammentare che "%" vuol solo dire "diviso 100", cioè 30% vuol solo dire 30/100 = 0.3.

– Quando lo stato d'informazione è costante nel tempo, la probabilità deve obbedire (56) a un'ulteriore regola, detta d'*additività*. Se due eventi $E1$, $E2$ sono *incompatibili*, ossia se non possono verificarsi congiuntamente (57), allora la probabilità che uno almeno d'essi si verifichi è la somma delle loro probabilità, in formula:

———

(49) Nel caso del dado Ω = {1, 2, 3, 4, 5, 6}.

(50) Nel caso del dado un esempio di evento potrebbe essere "Che esca un numero pari" che, tradotto come sottoinsieme di Ω, risulta $E1$ = {2, 4, 6}. Un altro esempio potrebbe essere "Che esca un numero minore di 3", che, tradotto come sottoinsieme di Ω, risulta $E2$ = {1, 2}.

(51) Pioverà o no tra una settimana?

(52) Mr. Hyde avrà davvero accoltellato Mr. X?

(53) Ciò delegittima locuzioni d'uso comune del tipo "Quante probabilità ho di riuscire a prendere il treno, se vado in stazione con un taxi?". Come se ciò che viene denominato probabilità fosse un gregge di pecore, eventualmente da contare.

(54) Per es., lanciamo una moneta con le due facce "testa e croce" ed esce ... un coniglio.

(55) Per es., questo articolo è pubblicato su *Cassazione Penale* ha probabilità 1.

(56) Qualunque ne sia l'interpretazione, classica, frequentista/statistica o soggettivista.

(57) Per es.: $E1$: "domani a mezzogiorno la temperatura in un luogo precisato sarà sotto 10°", ed $E2$: "domani a mezzogiorno la temperatura nello stesso luogo precisato sarà sopra 12°".

$$E_1, E_2 \text{ incompatibili} \rightarrow \Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$$

– Una conseguenza di questo fatto, molto importante per noi, riguarda un caso specialissimo. Traduce il classico principio logico del *tertium non datur*. Consideriamo due eventi E e "non E". L'evento "non E" se e solo se E è falso [58]. Esso è detto *contrario* di E e lo denotiamo con $\bar{E}$. L'evento certo, generalmente indicato con Ω, può pensarsi così:

$$E \cup \bar{E} = \Omega$$

Se ne ottiene:

$$\Pr(E \cup \bar{E}) = \Pr(\Omega)$$

ossia:

$$\Pr(E) + \Pr(\bar{E}) = 1$$

onde:

$$\Pr(\bar{E}) = 1 - \Pr(E)$$

in parole:

$$\Pr(\text{evento contrario}) = 1 - \Pr(\text{evento}).$$

———

(58) Se, per es., l'evento E è, per es.: E: "domani a mezzogiorno la temperatura in un luogo precisato sarà sotto 10°", l'evento $\bar{E}$ è "domani a mezzogiorno la temperatura in quel luogo precisato sarà da 10° in su".

www.giuffre.it/riviste/penale