

## STATISTICA E CONOSCENZA GIURISDIZIONALE: ALCUNI ELEMENTI PER UN DISCORSO\*.

di Giovanni Camillo Porzio\*\*

**Sommario.** 1. In premessa. – 2. Trappole conoscitive. – 3. Il livello di incertezza e la variabilità naturale. – 4. Il confronto tra proposizioni e il rapporto di verosimiglianza. – 4.1. Una scala di equivalenza verbale per il rapporto di verosimiglianza. – 5. In conclusione.

1

### 1. In premessa.

Nel corso dei miei studi statunitensi, ormai diversi anni fa, ebbi l'occasione di incontrare una studentessa americana, dottoranda di un corso di dottorato in statistica a anni alterni. Nell'anno in cui non frequentava i corsi di statistica studiava per un titolo accademico in scienze giuridiche. Mi spiegò che era noto che *il mercato* cercava figure professionali con questa doppia competenza. Stiamo parlando degli ultimi anni del secolo scorso. Ne rimasi particolarmente colpito. Mi sembrò il segno di una distanza tra due mondi, quello statunitense e quello italiano, che al tempo non avevo idea di quanto, quando e se si sarebbe colmata. E mai avrei pensato di ritrovarmi qui, anni dopo, a affrontare il tema del rapporto tra statistica e conoscenza giurisdizionale in un Dipartimento di Economia e Giurisprudenza. Tema che, per la mia storia personale e professionale, affronterò dal punto di vista della statistica.

Che statistica e diritto si parlino, d'altra parte, sembra essere opportuno e all'ordine del giorno non solo per paesi come l'Italia, ma anche e ancora per paesi con una storia del diritto molto diversa come la Gran Bretagna: è del 2017 la pubblicazione di linee guida per avvocati sull'uso della statistica e della probabilità nei tribunali a cura della britannica *The Inns of Court College of Advocacy* in collaborazione con la Royal Statistical Society, la prestigiosa società degli statistici della Corona Britannica<sup>1</sup>.

---

\* Sottoposto a referaggio.

\*\* Professore ordinario di Statistica – Università degli studi di Cassino e del Lazio Meridionale.

<sup>1</sup> *Statistics and Probability for Advocates: Understanding the Use of Statistical Evidence in Courts and Tribunals*; <https://www.icca.ac.uk/statistics-and-probability-for-advocates/>

Per questo motivo, questo intervento si concentra su alcuni elementi che ritengo possano facilitare un discorso tra queste discipline. L'auspicio è quello di poter offrire degli spunti di riflessione a chi, senza particolari competenze di statistica in partenza, desiderasse comprendere meglio come oggi statistica e conoscenza giurisdizionale possano interagire tra di loro.

Detto questo, vorrei sottolineare in premessa quanto il pensare arguto e - perché no - creativo, nonché la capacità di leggere le differenze sottili e il non fermarsi all'apparenza, siano a mio giudizio elementi fondanti dei processi di conoscenza che caratterizzano le scienze statistiche. Ben al di là forse dell'immagine che questa scienza offre di sé, spesso percepita troppo vicina all'oggettività e al rigore logico che i numeri per definizione sembrano possedere nell'immaginario collettivo. Mi permetto di credere che queste caratteristiche dei processi di apprendimento che si sviluppano all'interno di un'analisi statistica non siano troppo dissimili da quelle che si sviluppano tra le scienze giuridiche e nella conoscenza giurisdizionale in particolare.

## 2. Trappole conoscitive.

Non è dunque un caso che troviamo insidie e trappole conoscitive sia nel campo della conoscenza basata sui dati sia nell'ambito dell'ermeneutica giuridica. Da questa prospettiva, sia il ragionamento giuridico che quello statistico, potremmo dire come ogni tentativo umano di conoscere, affrontano gli stessi problemi.

In proposito, vorrei citare due situazioni nelle quali si vede come una lettura superficiale di semplici dati possa portare a una loro cattiva interpretazione e dunque a un loro cattivo utilizzo.

Il primo esempio riguarda la lettura che si può fare dei dati relativi alle recidive. Ho avuto modo di leggere l'opinione di chi sostiene che, poiché le recidive sono più basse tra coloro che sperimentano misure alternative al carcere, se ne conclude che tali misure siano più efficaci in termini di riduzione della probabilità di recidiva.

A pensarci bene, si tratta della semplice lettura di una "frequenza", cioè di quanto spesso un fenomeno accade in un ambito rispetto a un altro. A uno statistico, ma probabilmente a chiunque desideri andare oltre la superficie delle cose, appare evidente che questo elemento

(la frequenza di recidive) non può sostenere l'una tesi o l'altra. Il motivo è semplice: nella valutazione che porta un giudice a optare per una misura alternativa al carcere è già presente una valutazione sul soggetto che ne dovrebbe usufruire. Concederà cioè con più facilità queste misure a coloro che riterrà più propensi a un comportamento futuro meno o per nulla di tipo *criminale*.

Tecnicamente, noi statistici diremmo che si tratta di un caso di campionamento non casuale (vale a dire, i dati provengono da un insieme di persone che già in fase di *selezione* avevano una minor probabilità di recidiva). I risultati di un'analisi basata su tali dati sono da considerarsi *distorti*<sup>2</sup> in quanto siamo in presenza di un problema di selezione del campione e in particolare di un campione che non è rappresentativo della popolazione di riferimento. Bene aggiungere che questo errore nell'analisi del dato non porta acqua a nessun mulino: semplicemente non è questo un dato da valutare se si desidera fare una riflessione sull'utilità o meno di questi dispositivi di legge.

Come secondo esempio, sempre con riferimento a un concetto semplice quale è l'idea di frequenza, ovvero al numero di casi che avvengono sul totale, farò riferimento a un caso complesso che ha avuto una certa eco sulla stampa qualche anno fa: quello di un'infermiera dell'ospedale di Lugo in Romagna che avrebbe assassinato nel sonno alcuni dei pazienti dell'ospedale in cui lavorava.

Apprendo dai giornali<sup>3</sup>, che l'accusa abbia voluto rafforzare la propria posizione portando come elemento di giudizio il fatto che durante i turni di lavoro della suddetta infermiera il numero di morti in reparto fosse di gran lunga superiore alla media. Si tratta, per l'appunto di un dato di frequenza (decessi nelle 24h).

Alla difesa è stato possibile smontare questo elemento di giudizio: si dava il caso che l'infermiera avesse spesso il turno di notte, e è noto che la frequenza dei decessi notturni supera quella dei decessi durante il giorno. In breve, nella lettura di una quantità, di un numero, si era dato eccessivo valore al numero in sé piuttosto che al contesto in cui quel numero era maturato. Tecnicamente, noi statistici diremmo che non si era tenuto conto delle

---

<sup>2</sup> In statistica, il termine "distorto" ha un significato ben preciso. Per il livello di questo discorso, basti dire che la rappresentazione della realtà che si ottiene tramite un campione ottenuto in maniera non casuale non corrisponde generalmente alla realtà come dovrebbe emergere utilizzando una metodologia di campionamento "corretta".

<sup>3</sup> Corriere di Romagna, 19 dicembre 2015, così come riprodotto nella Rassegna Stampa dell'Unione dei Comuni della Bassa Romagna; <http://www.labassaromagna.it/content/download/57165/900716/file/UnionedeiComuniBassaRomagna19042015.pdf>

variabili di disturbo (quelle che, appunto, *disturbano* una corretta lettura del dato), ovvero di quelli che vengono anche chiamati *fattori confondenti*.

In breve, (a) la statistica, la *scienza del dato* per antonomasia, non ha tecnicamente la possibilità di dire tutto ciò che ci piacerebbe potesse dire (nel primo esempio sopra discusso, non c'è modo di sapere in base a questi dati se le misure alternative al carcere hanno o meno un impatto sulla possibilità di recidiva); (b) la lettura di dati anche semplici deve essere accorta (nel secondo esempio, il non prendere in considerazione altri elementi rispetto a quelli che saltano subito agli occhi porta a una visione distorta della realtà dei fatti).

### 3. Il livello di incertezza e la variabilità naturale.

Ciò affermato, e cioè che è necessaria esperienza e competenza anche per avvicinarsi correttamente a informazioni numeriche apparentemente di facile lettura, facciamo un passo avanti nel discorso e valutiamo, nell'ipotesi che il dato statistico sia stato letto correttamente, quale altro tipo di conoscenza è in grado di aggiungere la scienza statistica. Come ricordato da Adrian Bowman nella sua relazione di apertura in un recente convegno che si è tenuto a Cassino<sup>4</sup>, lo specifico della statistica è la capacità di valutare il livello di incertezza presente nei dati. Un'affermazione importante che proverò a spiegare con qualche ulteriore esempio e che ci permette di fare un altro passo avanti nella comprensione del contributo che le scienze statistiche possono dare alla conoscenza giurisdizionale.

Il primo esempio è banale. Posso affermare, senza paura di essere smentito, che l'età media dei componenti la mia famiglia a oggi è di 30,2 anni (cioè di circa 30 anni e due mesi e mezzo). Ora, è evidente a chi mi ascolta che io non ho 30 anni, ma non credo che ciò meravigli. Infatti, non è qualcosa di particolarmente strano: non è strano che - laddove c'è una media - ci sono anche valori superiori e valori inferiori rispetto alla media stessa.

Traslando il discorso in termini di conoscenza giurisdizionale, ne consegue che non è sufficiente affermare che un certo fenomeno è superiore (o inferiore) alla media, né è

---

<sup>4</sup> Adrian Bowman, Professore di Statistica presso l'Università di Glasgow, ha tenuto una "Keynote Lecture" al Convegno internazionale "Classification and Data Analysis" tenutosi presso l'Università di Cassino e del Lazio Meridionale nel settembre 2019; <http://cladag2019.unicas.it/keynote-speakers/>

sufficiente affermare quante volte lo sia (due volte, tre volte, cento volte) per individuare potenziali anomalie e a loro volta possibili segnali di comportamenti o situazioni al di fuori della legge.

Piuttosto, è necessario valutare quante volte lo sia rispetto ai casi simili disponibili. Tornando al caso dell'infermiera, e nell'ipotesi che la frequenza relativa dei decessi fosse superiore alla media dei decessi notturni in quel dato periodo dell'anno (è noto l'effetto stagionale sui decessi), il dato osservato, quello portato in giudizio dall'accusa, andrebbe confrontato con quanto in media le frequenze dei decessi differiscono dalla loro media. Sembra un gioco di parole ma non lo è.

Solo per illustrare meglio l'idea, supponiamo che la media dei decessi di notte sia del 2% dei ricoverati. Supponiamo che, nelle notti in cui era in servizio l'infermiera, si fosse osservata una frequenza del 5%. È sufficiente questo dato per ritenere questo evento come altamente improbabile e quindi come elemento di giudizio a sfavore dell'infermiera? La risposta è semplicemente *no*. Tale dato va infatti confrontato con le frequenze dello stesso fenomeno in ambiti simili. La media può essere 2% laddove le frequenze osservate in casi analoghi siano state, ad esempio, 0%, 0%, 0%, 1%, %, 1%, 2%, 3%, 3%, 4%, 5%, 5%, 5%, 4%, 4%, 0%, 0%, 1%, 1%, 1%, 0%, 3%.

Se così fosse, ciò che a prima vista sembrava *anomalo* in realtà non lo è. Saremmo infatti in presenza di quella che alcuni statistici chiamano *variabilità naturale*: il fenomeno ha una sua variabilità interna (dovuta per esempio alle diverse attrezzature disponibili nelle diverse sale della clinica, alla capacità d'uso dei singoli infermieri delle stesse attrezzature, e a molti possibili altri fattori) per cui la differenza dalla media che osserviamo, per quanto ampia (il 5% è più del doppio del 2%) è in realtà *nella norma*.

Vale ovviamente la possibilità del caso inverso, ovvero che in presenza di una ipotetica frequenza media pari al 2%, quella osservata nel caso di specie (nelle notti in cui era in servizio l'infermiera) sia stata di non molto più alta, per esempio del 2,5%, e tuttavia abbastanza alta rispetto alla *variabilità naturale* da diventare elemento di giudizio a sfavore dell'imputato. Si pensi alla sequenza 2.01%, 2.04%, 2.00%, 2.13%, 2.17%, 2.11%, 2.20%, 2.03%, 2.12%, 2.02%, 1.90%, 2.19%, 2.01%, 1.62%, 1.45%, che ha sempre media 2%, ma con una variabilità (una diversità tra i valori) di gran lunga inferiore rispetto alla sequenza di percentuali riportati nel caso precedente. In tal caso, il valore 2,5% apparirebbe particolarmente anomalo.

Lo statistico si forma e impara nel tempo a misurare questa variabilità, la variabilità naturale<sup>5</sup>, eventualmente a enuclearla da variabilità spiegabili in altro modo tramite l'uso di modelli più o meno complessi e infine a utilizzare questa variabilità per vagliare la veridicità di una o più ipotesi, o proposizioni.

#### 4. Il confronto tra proposizioni e il rapporto di verosimiglianza.

Un elemento che vorrei mettere in evidenza in questo discorso è la probabilità in quanto misura del grado di incertezza associata a una certa conoscenza. In particolare, si va a discutere quale spazio la probabilità può trovare – dal punto di vista della statistica – all'interno di un processo di conoscenza giurisdizionale.

Dato un elemento concreto, la cosiddetta *evidenza*, cioè ciò che viene effettivamente osservato, è possibile infatti ritrovarsi a esprimere un giudizio “in probabilità” sulla veridicità di una o più proposizioni<sup>6</sup>.

Si immagini, a esempio, che venga individuata una traccia di sangue sul luogo del delitto e che ci si chieda se questa traccia appartenga o meno a un certo indiziato. L'ipotesi, o proposizione, è che la traccia appartenga all'indiziato, mentre l'evidenza è quanto emerge dall'analisi del DNA. Analisi che tipicamente fornisce una misura non certa e dunque una misura in probabilità. È possibile infatti determinare quanto è probabile che osserviamo l'evidenza (la traccia di DNA sotto esame) sotto l'assunto che sia vera l'ipotesi o proposizione che desideriamo valutare.

In questo contesto, va ben distinto ciò che è possibile da ciò che non è possibile affermare. Non è infatti assolutamente possibile *valutare la probabilità che l'ipotesi sia vera (o falsa)*, ma solo – come detto – la probabilità di osservare quanto osservato (l'evidenza) sotto l'assunto che l'ipotesi sia vera. Una apparente piccola distinzione linguistica che è invece una sostanziale differenza di approccio che dovrebbe essere ben presente a chi si interroga

---

<sup>5</sup> Si veda in proposito quanto scrive in maniera divulgativa, in merito alla “*Natural Variability*”, J. M. UTTS, *Seeing Through Statistics*, 4th Edition, Cengage Learning, 2014.

<sup>6</sup> Si seguiranno qui le Linee Guida elaborate dalla “Rete europea degli istituti di statistica forense” (ENFSI) per la valutazione delle relazioni tecniche nella scienza forense (*ENFSI guideline for evaluative reporting in forensic science*, <http://enfsi.eu/documents/forensic-guidelines/>); l'ENFSI (*European Network of Forensic Science Institutes*) è l'organismo tecnico di riferimento europeo, sia della Comunità Europea, sia del Gruppo di Cooperazione Europeo delle Polizie, nonché dell'Europol e dell'Interpol, riguardo la definizione degli standard tecnici utilizzati dai Laboratori di Polizia Scientifica.

sul contributo che la statistica può fornire alla conoscenza giurisdizionale.

Il discorso sulla valutazione delle ipotesi non si conclude però così semplicemente perché, nella dinamica di un processo, vanno confrontate tra loro (almeno) due ipotesi: quella della difesa e quella dell'accusa.

Di conseguenza, e continuando con l'esempio, avremo due probabilità da mettere a confronto: la probabilità dell'evidenza sotto l'ipotesi che la traccia appartenga all'indiziato e la probabilità della stessa evidenza sotto l'ipotesi che la traccia non appartenga all'indiziato.

Logica vorrebbe che si confrontassero queste due probabilità e che laddove l'una fosse maggiore dell'altra si propenda per l'una delle due ipotesi piuttosto che per l'altra. Per fare ciò, è opportuno ricorrere al Rapporto di Verosimiglianza (RV), cioè al rapporto tra queste due probabilità.

Il rapporto di verosimiglianza<sup>7</sup> misura infatti il valore dell'evidenza e viene utilizzato per decidere quale tra due proposizioni alternative in un dibattimento processuale debba prevalere. È rapporto tra due probabilità per definizione: la probabilità di osservare l'evidenza sotto l'ipotesi che l'imputato sia colpevole (la proposizione dell'accusa) e la probabilità di osservare la stessa evidenza sotto l'ipotesi che l'imputato non sia colpevole (proposizione della difesa). Ognuna delle probabilità presa in modo isolato sarebbe infatti un'informazione che porterebbe a una visione distorta della realtà.

Come è stato infatti sottolineato<sup>8</sup>, è sempre vero che, anche se l'evidenza fosse poco probabile sotto l'ipotesi che l'imputato è innocente, potrebbe essere ancora meno probabile sotto l'ipotesi che l'imputato sia colpevole. Non può essere dunque utilizzata come prova l'analisi della sola probabilità dell'evidenza sotto una delle due ipotesi.

Evidentemente, se in un dato caso specifico il RV avesse un valore vicino o esattamente uguale a uno, allora le due probabilità sono equivalenti e non c'è motivo per preferire un'ipotesi rispetto all'altra.

Se il RV dovesse essere maggiore di uno, allora bisognerebbe concludere che ciò che si è osservato è più probabile se la proposizione (l'ipotesi) al numeratore è vera piuttosto che la proposizione specificata al denominatore, viceversa se il RV è minore di uno. In breve, il

---

<sup>7</sup> Il rapporto di verosimiglianza è anche noto in letteratura, compresi in alcuni lavori in lingua italiana, come *Likelihood Ratio* (LR).

<sup>8</sup> A. Colin, R. Paul e J. Graham, *Fundamentals of probability and statistical evidence in criminal proceedings*, Royal Statistical Society, 2010.

RV fornisce una misura di quanto il risultato proveniente dalle evidenze sia più probabile sotto l'una o l'altra ipotesi.

D'altra parte, la forza dell'argomentazione, dell'essere a favore dell'una o dell'altra preposizione è legata anche all'ampiezza del rapporto, cioè a quanto sia maggiore (o minore) di uno.

#### 4.1. Una scala di equivalenza verbale per il rapporto di verosimiglianza.

È interessante in proposito recuperare la *scala di equivalenza verbale* tra il valore del RV e la forza argomentativa della prova, così come elaborata dapprima da Nordgaard *et al.*<sup>9</sup>, e poi ripresa a fini illustrativi dalle succitate Linee Guida ENFSI<sup>10</sup>.

L'idea di questa scala di equivalenza nasce dall'ipotesi che sia necessario in qualche modo tradurre i valori del RV in una scala ordinale e in un'equivalente espressione verbale: una *equivalenza verbale* tra numeri e parole può certamente accorciare le distanze tra i tecnici dei numeri e chi si trova a lavorare quotidianamente nelle aule di un tribunale. Mentre infatti il valore assunto da un RV è facilmente comprensibile dagli scienziati forensi con una formazione quantitativa, può risultare ostico a altri operatori del settore.

Essendo quindi un'idea in linea con lo spirito che anima questo lavoro, la scala di equivalenza viene qui riportata in Tabella 1.

Come avvertenza, va detto che gli intervalli di valori del RV in Tabella 1 sono indicativi, visto che il RV assume valori nel continuo: è evidente che un RV di 999 è sostanzialmente equivalente a un RV di 1001.

Aggiungo che, se a prima vista questo discorso può sembrare un po' astratto, e non utile nella prassi della conoscenza giurisdizionale, il rapporto di verosimiglianza è stato utilizzato per mettere in evidenza come, in un particolare caso di specie, la prova del DNA non poteva essere considerata risolutiva da un punto di vista scientifico, in quanto la corretta valutazione del rapporto di verosimiglianza portava a un valore vicino all'unità<sup>11</sup>.

<sup>9</sup> A. Nordgaard, R. Ansell, W. Drotz e L. Jaeger, *Scale of conclusions for the value of evidence*, Law, Probability & Risk, 2012, pp. 11:1–24.

<sup>10</sup> Di cui alla nota 6 di questo contributo.

<sup>11</sup> F. Taroni, S. Bozza e P. Garbolino, *Contaminazioni di un reperto con il DNA. Quando la prova genetica porta direttamente alla condanna*, in *Diritto penale contemporaneo*, 12 febbraio 2018.



Tabella 1.

Valore del rapporto di verosimiglianza (RV) e sua equivalenza verbale<sup>12</sup>. Per ogni valore o insieme di valori vengono suggerite due frasi equivalenti.

<i>Valore del Rapporto di Verosimiglianza (RV)</i>	<i>Equivalenza verbale (per prima ipotesi o proposizione si intende quella di cui la probabilità dell'evidenza è al numeratore)</i>
1	Il valore del RV non favorisce una ipotesi rispetto all'altra. Il valore del RV non fornisce elementi utili per risolvere la questione
2 – 10	Il valore del RV favorisce debolmente la prima ipotesi rispetto a quella alternativa. Il dato osservato è leggermente più probabile se la prima preposizione è vera piuttosto che non la seconda.
10 - 100	Il valore del RV favorisce moderatamente la prima ipotesi rispetto a quella alternativa. Il dato osservato è più probabile se la prima preposizione è vera piuttosto che non la seconda.
100 - 1.000	Il valore del RV favorisce con una certa forza la prima ipotesi rispetto a quella alternativa. Il dato osservato è apprezzabilmente più probabile se la prima preposizione è

<sup>12</sup> Fonte: Linee Guida ENSFI, di cui alla nota 6. Traduzione dall'inglese a cura dell'autore.

	vera piuttosto che non la seconda.
1.000 - 10.000	Il valore del RV favorisce con forza la prima ipotesi rispetto a quella alternativa. Il dato osservato è molto più probabile se la prima proposizione è vera piuttosto che non la seconda.
10.000 - 1.000.000	Il valore del RV favorisce con particolare forza la prima ipotesi rispetto a quella alternativa. Il dato osservato è di gran lunga molto più probabile se la prima proposizione è vera piuttosto che non la seconda.
Più di 1.000.000	Il valore del RV favorisce in maniera estremamente robusta la prima ipotesi rispetto a quella alternativa. Il dato osservato è straordinariamente e di gran lunga molto più probabile se la prima proposizione è vera piuttosto che non la seconda.

In breve, il RV ci può aiutare a discriminare tra due diverse ipotesi o proposizioni. E, tuttavia, ritengo che anche qui vada introdotta un'ulteriore riflessione.

Se infatti è vero che un rapporto di verosimiglianza pari a 10 ci dice che una delle due proposizioni è 10 volte più probabile dell'altra, dovremmo anche e sempre chiederci quanto vale tale probabilità. Se il rapporto è uguale a 10, e la probabilità di un'evidenza empirica data la prima ipotesi è 0,7 (70%) contro una probabilità di 0,07 (7%), allora la forza dell'evidenza in favore della prima proposizione è innegabile. Cosa succede se invece un rapporto di verosimiglianza pari a 10 corrisponde a una probabilità dell'evidenza pari a 0,001 (1 su 1.000) con riferimento alla prima proposizione a fronte di una probabilità di 0,0001 (1 su 10.000) con riferimento alla seconda? Evidentemente, l'una ipotesi ha più forza dell'altra, ma entrambe sono molto deboli.

Se ne conclude che, quando si lavora con questo tipo di probabilità, tutti e tre gli elementi di questo discorso (le probabilità dell'evidenza sotto le due proposizioni e il loro rapporto) vanno presi in considerazione.

### 5. In conclusione.

Per concludere, anche solo a partire da questi piccoli cenni, mi appare che gli spazi per un dialogo tra la scienza statistica e la conoscenza giurisdizionale siano ampi e le prospettive di prassi e di ricerca ancora molte e da esplorare. Senza voler essere esaustivo, penso alle problematiche legate ai casi di discriminazione sui luoghi di lavoro, alle valutazioni del danno per colpa medica, ai danni da inquinamento ambientale verso gruppi o comunità territoriali (l'amianto, l'ILVA di Taranto, gli sversamenti nella terra dei fuochi), al problema dell'incertezza che è propria degli strumenti di misura (per esempio all'uso delle informazioni provenienti dalle scatole nere che vengono montate sulle automobili nelle cause con incidenti stradali), al problema della individuazione della causa nelle malattie professionali.

E tuttavia, a tutti è chiesto uno sforzo. Al *tecnico* va chiesto di rendere il proprio linguaggio il più vicino possibile a quello del giurista. Al giurista, di comprendere i limiti della conoscenza tecnica. Non si chieda al tecnico di dire ciò che non può dire.

**Abstract:** La necessità di far parlare tra loro le scienze statistiche e quelle giuridiche è sempre più evidente, sia in Italia che all'estero. Per questo motivo, questo contributo offre a un pubblico non specialista alcuni elementi propri delle scienze statistiche potenzialmente utili per districarsi tra termini e problemi che si possono incontrare nella prassi e nella ricerca legate alla conoscenza giurisdizionale. Si presentano alcuni esempi di trappole conoscitive; si discute il livello di incertezza e il concetto di variabilità naturale; si ripercorre l'uso corretto del rapporto di verosimiglianza; si offre una scala di equivalenza verbale per tale rapporto. Si delineano infine alcuni spazi aperti per una collaborazione tra queste due discipline, sottolineando l'importanza di riconoscere i limiti propri della conoscenza che nasce a partire dall'analisi dei dati e dell'evidenza empirica.

**Abstract:** The need for statistics and law to meet somewhere and to talk to each other is widely recognized, both in Italy and abroad. For this reason, this contribution provides some basic elements of statistics potentially useful within a judicial knowledge process. Some examples of cognitive traps are presented; the level of uncertainty and the concept of natural variability are discussed; the correct use of the likelihood ratio is highlighted; a verbal equivalence scale is offered for such a ratio. Finally, a list of issues where these two disciplines can cooperate is offered, while the importance of recognizing the limits of the knowledge arising from the analysis of data and empirical evidence is eventually underlined.

**Parole chiave:** frequenza relativa – variabilità naturale – probabilità – rapporto di verosimiglianza – scala di equivalenza verbale.

**Key words:** relative frequency – natural variability – probability - likelihood ratio – verbal equivalents.