

DATI E METODI PER LA STATISTICA GIUDIZIARIA



Maria Filomeno,
Irene Rocchetti

Studies in Law
and Social Sciences

9



Università degli Studi Roma Tre
Dipartimento di Giurisprudenza

NELLA STESSA COLLANA

1. G. ROJAS ELGUETA, N. VARDI (a cura di), *Oltre il soggetto razionale*, 2014
2. F. MEZZANOTTE (a cura di), *Le «libertà fondamentali» dell'Unione europea e il diritto privato*, 2016
3. C.A. D'ALESSANDRO, C. MARCHESE (a cura di), *Ius dicere in a globalized world. A comparative overview*, 2018
4. A. ZOPPINI, P. SIRENA (a cura di), *I poteri privati e il diritto della regolazione*, 2018
5. F. CAGGIA, G. RESTA (a cura di), *I diritti fondamentali in Europa e il diritto privato*, 2019
6. A. SOMMA, V. ZENO-ZENCOVICH (a cura di), *Comparazione e diritto positivo. Un dialogo tra saperi giuridici*, 2021
7. R. LUPI, *Studi sociali e diritto*, 2022
8. G. GIANNONE CODIGLIONE, L. PIERDOMINICI (a cura di), *Comparative Law in Times of Emergencies*, 2022

Università degli Studi Roma Tre
Dipartimento di Giurisprudenza
Collana “Studies in Law and Social Sciences”

9

Maria Filomeno, Irene Rocchetti

DATI E METODI PER LA STATISTICA GIUDIZIARIA

Prefazione di
Antonella Ciriello

Postfazione di
Raffaello Lupi



Roma TrE-PRESS
2023

Collana pubblicata nel rispetto del Codice etico adottato dal Dipartimento di Giurisprudenza dell'Università degli Studi Roma Tre, in data 22 aprile 2020.

Il volume pubblicato è stato sottoposto a previa e positiva valutazione nella modalità di referaggio *double-blind peer review*.

Coordinamento editoriale:
Gruppo di Lavoro *RomaTrE-Press*

Elaborazione grafica della copertina: **MOSQUITO**, mosquitoroma.it

Impaginazione: Colitti-Roma colitti.it

Caratteri tipografici utilizzati:
Brandon Grotesque(copertina e frontespizio)
Adobe Garamond Pro (testo)

Edizioni: RomaTrE-Press
Roma, giugno 2023
ISBN: 979-12-5977-191-9

<http://romatrepress.uniroma3.it>

Quest'opera è assoggettata alla disciplina *Creative Commons attribution 4.0 International License* (CC BY-NC-ND 4.0) che impone l'attribuzione della paternità dell'opera, proibisce di alterarla, trasformarla o usarla per produrre un'altra opera, e ne esclude l'uso per ricavarne un profitto commerciale.



L'attività della *RomaTrE-Press* : svolta nell'ambito della
Fondazione Roma Tre-Education, piazza della Repubblica 10, 00185 Roma

Indice Generale

<i>Prefazione di ANTONELLA CIRIELLO</i>	III
<i>Introduzione</i>	V

CAPITOLO I

Le rilevazioni giudiziarie in Italia: organi e fonti

1. <i>La Commissione per l'efficienza della giustizia del Consiglio d'Europa (Cepej)</i>	1
2. <i>L'Eurostat</i>	2
3. <i>L'Istituto Nazionale di Statistica (Istat)</i>	3
4. <i>Il Ministero della Giustizia e la Direzione Generale di Statistica (DGStat)</i>	11
5. <i>Il Consiglio Superiore della Magistratura e l'Ufficio statistico</i>	18

CAPITOLO II

Tutela della riservatezza e titolarità dei dati statistici giudiziari

1. <i>Il trattamento dei dati personali per scopi statistici</i>	21
2. <i>La titolarità dei dati</i>	22
3. <i>L'accesso ai dati – Dati pubblici e dati aperti</i>	26

CAPITOLO III

Misure statistiche e indicatori utilizzati in ambito giudiziario

1. <i>Le misure statistiche</i>	27
2. <i>Gli indicatori statistici</i>	29
3. <i>I principali indicatori di performance quantitativi e qualitativi in giustizia</i>	30

CAPITOLO IV

Statistica per la giustizia: come descrivere e sintetizzare i dati

1. <i>Unità statistiche e caratteri</i>	39
2. <i>Distribuzione statistica e frequenza</i>	40
3. <i>Le rappresentazioni grafiche di variabili qualitative e quantitative discrete</i>	42
4. <i>Le rappresentazioni grafiche di variabili quantitative continue</i>	46
5. <i>Indici di posizione</i>	53

6. <i>Il concetto di variabilità e gli indici relativi</i>	61
7. <i>Il rapporto di concentrazione</i>	64
8. <i>La Mutabilità</i>	69
9. <i>Distribuzioni doppie di frequenze</i>	71

CAPITOLO V

Statistica per la giustizia: come analizzare i dati

1. <i>La connessione statistica</i>	75
2. <i>La Concordanza</i>	77
3. <i>La relazione tra variabili</i>	79
4. <i>Cenni alla verifica delle Ipotesi</i>	82
5. <i>Cenni al campionamento e alle sue tecniche</i>	85
6. <i>La qualità dei dati</i>	93
7. <i>Cenni alle principali tecniche di analisi multivariata</i>	95
8. <i>Cenni alle Serie Storiche</i>	96
9. <i>Software per la Statistica</i>	99

CAPITOLO VI

La statistica nell'organizzazione degli uffici giudiziari

1. <i>Le Tabelle di organizzazione degli uffici giudiziari giudicanti</i>	101
2. <i>L'organizzazione degli uffici di Procura</i>	104
3. <i>Programmi di gestione</i>	106
4. <i>Valutazioni di professionalità e standard di rendimento</i>	108
5. <i>Il Piano Nazionale di Ripresa e Resilienza</i>	110
6. <i>L'ufficio del Processo</i>	113
7. <i>Digitalizzazione e data lake</i>	114
<i>Postfazione di RAFFAELLO LUPI</i>	117
<i>Appendice</i>	121
<i>Riferimenti bibliografici</i>	127

Antonella Ciriello

Prefazione

In un momento storico in cui non si parla altro che di digitalizzazione, intelligenza artificiale, riduzione dell'arretrato e misurazione della performance, non vi è chi dubiti dell'importanza dei dati e della statistica, come scienza che trasforma i dati in informazioni.

E, tuttavia, se i dati e la loro lettura attraverso la scienza statistica costituiscono indispensabili oggetti di conoscenza per il giurista essi, nonostante se ne faccia un gran parlare, non sono comunemente patrimonio di conoscenza comune.

Vuoi per la complessità del fenomeno, precipitato logico della eterogeneità e complessità della articolata macchina della giustizia, vuoi per le difficoltà concrete di rendere i dati omogenei, indagabili, intellegibili, finora la conoscenza di questa scienza (considerata peraltro, a torto e anacronisticamente non indispensabile e centrale) è appartenuta a pochi.

E, invece, il bagaglio di conoscenze statistiche compone la competenza qualificata degli amministratori della giustizia, a tutti i livelli, oltre che degli studiosi del fenomeno ed anche del personale amministrativo e la scienza statistica assume in questo momento un ruolo assolutamente centrale nella gestione e organizzazione della Giustizia, come chiarito dalle autrici.

Il volume "Dati e metodi per la statistica giudiziaria" di Maria Filomeno e Irene Rocchetti, accompagna per mano il lettore attraverso un viaggio che affronta il fenomeno in maniera completa, da tutti i possibili punti di vista.

Oltre agli aspetti storici e organizzativi, con la descrizione dei soggetti responsabili della Statistica, è affrontato in maniera piana l'annoso problema della titolarità dei dati statistici, della loro accessibilità, della responsabilità per la loro qualità (aspetto, quest'ultimo, di particolare rilevanza per il magistrato che a tale qualità può concretamente contribuire con attività specifiche).

Anche quando sono descritti i concetti statistici, essi sono spiegati in maniera semplice, e ogni spiegazione è sempre accompagnata da esempi concreti che consentano al lettore di calare le formule nella realtà, comprendendone l'utilizzo.

Le autrici, forti di una esperienza pluridecennale al servizio della

giustizia, non si sottraggono alla esauriente trattazione di istituti normativi e organizzativi assai discussi e complicati, come il Documento organizzativo generale, le Tabelle organizzative degli uffici, anche requirenti, i Programmi di gestione e gli Standard di rendimento (ai fini della valutazione di professionalità del magistrato), offrendo le loro equilibrate e obiettive descrizioni alla valutazione critica del lettore, relativamente ad istituti per i quali il Consiglio superiore ha percorso nell'ultimo decennio un delicato cammino evolutivo (in alcuni casi non ancora completato, come per gli Standard di rendimento).

L'attualità e modernità dell'opera, che la rende unica nel suo genere, si può apprezzare quando si legge la parte relativa al PNRR, piano Nazionale di Ripresa e Resilienza, con la chiara descrizione degli obiettivi perseguiti, nonché le pagine dedicate all'ufficio del processo (istituto nuovo e tanto discusso per le sue ricadute pratiche anche in relazione alle prospettive europee) e al tema complesso della digitalizzazione e del data lake (e all'enorme ampliamento della capacità informativa e di analisi degli strumenti che, attraverso le tecnologie dell'intelligenza artificiale, tale nuovo approccio produrrà, rendendo accessibili i dati in formato nativo).

Un'analitica appendice contenente la distribuzione della dimensione degli uffici (tabella 21), completa l'opera.

È questo un volume indispensabile per chi, senza temere i risultati della scienza che raccoglie i dati per trarne informazioni, voglia con essi confrontarsi in maniera matura e consapevole, come appare doveroso, poiché solo in questo modo la giustizia può assumere davvero lo scettro e governare sé stessa.

Si tratta – dunque - di una operazione conoscitiva indispensabile, in questi tempi di innovazioni tecnologiche, cui i giuristi non possono sottrarsi, per non correre il rischio di subire scelte e valutazioni, ma per governarle attraverso una illuminata gestione delle informazioni, alimentando così, attraverso la conoscenza e l'elevata professionalità conseguita, i fondamentali valori dell'imparzialità e dell'indipendenza, che la Costituzione repubblicana affida all'ordine giudiziario.

Introduzione

La statistica giudiziaria è diventata supporto fondamentale per l'organizzazione degli uffici giudiziari. Essa consente sia di rilevare gli effetti complessivi diretti e indiretti delle leggi sia di ricavare informazioni preziose per l'organizzazione degli uffici.

Questa duplice funzione, interna ed esterna, deve oggi necessariamente avvalersi delle rilevazioni statistiche quantitative che 'popolano' le basi dati che costituiscono l'ecosistema informativo della Giustizia.

"La statistica giudiziaria comprende sia la statistica processuale che osserva la mole dei processi in carico al sistema giustizia sia le rilevazioni sul merito di tale attività come litigiosità e criminalità. Queste ultime, riportando le caratteristiche salienti dei procedimenti, risultano anche più ricche sotto il profilo della analisi sociale del fenomeno.

Le statistiche giudiziarie derivano prevalentemente da fonti amministrative e sono rimaste fondamentalmente invariate dal 1861 fino agli ultimi anni del secolo scorso, quando vennero introdotte numerose innovazioni nel processo di raccolta delle informazioni. I dati comunque sono da sempre trasmessi dagli uffici giudiziari competenti per materia e l'unità minima di riferimento è rimasta il singolo procedimento giudiziario.

Anche i contenuti sono rimasti sostanzialmente stabili nel tempo. Le modifiche apportate nel periodo postbellico ai modelli di rilevazione si limitano infatti a tener conto dei nuovi istituti giuridici, senza mutare i contenuti informativi delle indagini.

Negli anni si sono registrati, invece, alcuni cambiamenti nella titolarità delle rilevazioni statistiche. Le prime rilevazioni in ambito giudiziario penale si rintracciano nel Granducato di Toscana e Regno di Sardegna, dove si svolgevano già rilevazioni statistiche sull'attività dell'apparato giudiziario, con fini essenzialmente amministrativi. Con il completamento del processo di unità nazionale, le varie statistiche giudiziarie furono ricondotte nell'ambito del Ministero di grazia e giustizia, presso il quale venne creato un Ufficio centrale di statistica (r.d. del 22 dicembre 1872), poi soppresso nel 1881. Successivamente le statistiche giudiziarie passarono nella competenza del Ministero dell'agricoltura, presso il quale era funzionante la Divisione di statistica generale, antesignana del futuro Istat.¹ Nel 1908 le statistiche giudiziarie tornarono al Ministero di grazia e giustizia e vi rimasero fino al 1938, quando furono trasferite all'allora Istituto centrale di statistica con la legge n. 402 del 24 marzo 1938.

È soltanto a partire dal 1997, in base ai principi di decentramento della funzione statistica introdotti dal d.lgs. n. 322 del 1989, che le rilevazioni sul movimento dei procedimenti civili e penali sono state trasferite nuovamente dall'Istat al Ministero della giustizia; mentre la rilevazione sui delitti denunciati alle Forze dell'ordine diventa di titolarità del Ministero dell'interno. Resta invece tuttora di competenza dell'Istat l'indagine sulla criminalità.¹”

Alla luce del volume di informazioni presenti nelle varie basi dati che costituiscono il sistema informativo della Giustizia diventa allora indispensabile padroneggiare le tecniche statistiche maggiormente adottate per l'analisi delle stesse e che risultano cruciali nelle fasi di programmazione, al fine di produrre conoscenza di supporto per le decisioni.

L'impiego della statistica attualmente non può più prescindere dalla capacità di calcolo necessaria ad elaborare la crescente quantità di dati prodotti e immessi nelle varie piattaforme digitali.

Per comprendere questo aspetto citeremo un esempio storico. “Nel tardo ottocento la popolazione americana cresceva del 30% ogni dieci anni, e l'elaborazione dei dati del censimento del 1880 richiese otto anni: la sua ultimazione si trovava quasi a coincidere con l'inizio del censimento successivo, quello del 1890. L'intero progetto minacciava di crollare per via delle dimensioni che stava assumendo. Ma il rimedio non tardò a venire: prendendo ispirazione dalla tecnica della foratura dei biglietti ferroviari del tempo, Herman Hollerith progettò un sistema di carte perforabili come supporto fisico per le informazioni assieme ad una macchina per contarle e ordinarle.

In seguito, nel 1911 si associò ad altri per fondare la CTRC che pochi anni più tardi fu ribattezzata come IBM.” (cfr. P. ZELLINI, *La dittatura del calcolo*)

Una sfida analoga è quella che il comparto della Giustizia si trova ad affrontare oggi. L'introduzione del processo civile e penale telematico e i numerosi registri informatizzati costituiscono preziosi giacimenti informativi - data lake - da cui attingere per produrre conoscenza. I datawarehouse (DWH) civile e penale sono gli evoluti sistemi di intelligence che consentono di effettuare analisi strutturate e approfondite sui dati provenienti da tali fonti che occorre correttamente interpretare.

Nel contesto delineato la lettura del dato giudiziario, per sua natura complesso e multidimensionale, risulterà tanto più efficace quanto più la relazione semantica che sussiste tra le fattispecie giuridiche astratte e le

¹ Tratto da 'L'Italia in 150 anni. Sommario di statistiche storiche 1861-2010'.

informazioni digitali codificate garantirà un *rapporto di indicazione*² statisticamente robusto che consentirà di tradurre i concetti in indicatori di sintesi.

A tal fine la diffusione di una cultura del dato giudiziario risulta di fondamentale importanza per migliorare la qualità delle informazioni, ottenere statistiche attendibili e interpretarne i risultati.

È questo il filo conduttore del compendio. La conoscenza del dato, infatti, parte dai principali enti che si occupano della raccolta e, mediante l'ausilio di misure, indicatori di sintesi e strumenti statistici, ne consente la concreta utilizzazione nella predisposizione di procedure e documenti organizzativi all'interno degli uffici giudiziari.

² Nelle scienze sociali indica il rapporto che sussiste tra concetti o fenomeni astratti e gli indicatori empirici utilizzati per rilevare indirettamente tali fenomeni attraverso delle misurazioni statistiche.

Capitolo I

Le rilevazioni giudiziarie in Italia: organi e fonti

La conoscenza delle fonti dei dati e della loro tipologia è fondamentale per avviare un percorso conoscitivo in ambito statistico. In questo capitolo vengono presentati i principali enti, italiani ed europei, che si occupano di produrre statistiche in ambito giudiziario: la Commissione per l'efficienza della giustizia del Consiglio d'Europa, l'Eurostat, l'Istat, il Ministero della Giustizia, il Consiglio Superiore della Magistratura.

Vengono quindi illustrate le statistiche a disposizione di cittadini e interlocutori specializzati.

1. La Commissione per l'efficienza della giustizia del Consiglio d'Europa (Cepej)

Il Consiglio d'Europa ha come obiettivo quello di promuovere la democrazia, i diritti umani, l'identità culturale europea e la ricerca di soluzioni ai problemi sociali in Europa.

All'interno del Consiglio d'Europa è stata istituita nel 2002 la Commissione per l'efficienza della giustizia del Consiglio d'Europa (Cepej), un organismo con l'obiettivo di migliorare l'efficienza e il funzionamento della giustizia negli Stati membri nonché di sviluppare e implementare gli strumenti adottati dal Consiglio d'Europa. Con l'istituzione della Cepej il Consiglio d'Europa ha espresso la volontà di promuovere lo Stato di diritto e i Diritti Fondamentali in Europa, sulla base della Convenzione Europea sui diritti umani, e specialmente nei suoi seguenti articoli: 5 (Diritto alla libertà e sicurezza), 6 (Diritto al processo equo), 13 (Diritto al ricorso effettivo), 14 (Divieto di discriminazione).

A tal proposito la CEPEJ individua benchmark, raccoglie e analizza dati, definisce strumenti di misura e metodi di valutazione, adotta documenti (reports, linee guida, etc.), sviluppa contatti con personalità qualificate, organizzazioni non governative, istituti di ricerca e centri di informazione,

organizza udienze, promuove networks di professionisti legali.

La Cepej è composta da tecnici ed esperti dei 46¹ Paesi membri supportati da un apparato di segreteria e i suoi compiti sono quelli di:

- analizzare i risultati dei sistemi giudiziari,
- identificare le difficoltà che essi incontrano,
- definire modi concreti di miglioramento della valutazione dei loro risultati e del funzionamento dei sistemi giudiziari,
- fornire assistenza agli Stati membri se richiesto,
- proporre alle istanze competenti del Consiglio d'Europa i campi in cui sarebbe desiderabile elaborare un nuovo strumento legale (informazioni tratte dal seguente sito <<https://www.coe.int/en/web/cepej/about-cepej>>).

I dati presenti sul sito della Cepej sono inerenti ai casi di litigiosità civile (e commerciale) dal 2010 e loro variazione % in vari Paesi, al personale in organico presso gli uffici giudiziari dei vari paesi europei e non, al budget totale dei Tribunali dei vari Paesi, all'indice di sviluppo ICT (Information and Communication Technology) totale e per categoria (tasso di distribuzione, comunicazione con i Tribunali, Tribunale e gestione del caso, supporto alle decisioni) e materia (civile, penale, amministrativa), al numero di norme legislative, all'impatto dei progetti ICT. Sono inoltre presenti dei report con analisi più dettagliate anche sui profili giudiziari dei Paesi.

2. *L'Eurostat*

L'Ufficio Statistico dell'Unione Europea è una direzione generale della Commissione Europea, raccoglie ed elabora i dati dagli Stati membri dell'Unione europea e promuove il processo di armonizzazione della metodologia statistica tra i Paesi.

Sul sito ufficiale dell'Eurostat (<<https://ec.europa.eu/eurostat/web/main/data/statistics-by-theme>>) sono disponibili statistiche generali e regionali, relative all'economia e alla finanza, alla popolazione e alle condizioni sociali, all'industria e ai servizi, agricoltura e pesca, commercio internazionale, trasporto, energia e ambiente, scienze, tecnologia e società digitale.

Sullo stesso sito sono presenti indicatori europei che forniscono informazioni economiche generali sull'area dell'euro, dell'Unione europea e degli Stati membri. La maggior parte degli indicatori sono prodotti

¹ Dal 16/03/2022 la Federazione russa non è più un membro del Consiglio d'Europa.

da Eurostat e sono integrati da alcuni indicatori monetari e finanziari prodotti dalla Banca Centrale Europea e dai dati delle indagini di tendenza economica prodotte dalla Commissione Europea.

3. *L'Istituto Nazionale di Statistica (Istat)*

L'Istituto nazionale di statistica è un ente di ricerca pubblico. È presente nel Paese dal 1926 ed è il principale produttore di statistica ufficiale a supporto dei cittadini e dei decisori pubblici.

L'organizzazione dell'Istituto è stata ridisegnata con il Decreto del Presidente della Repubblica n. 166 del 7 settembre 2010. Gli organi di governo sono il Presidente – che è nominato con decreto del Presidente della Repubblica e ha la rappresentanza legale dell'Istituto – e il Consiglio, che ha funzioni di indirizzo, programmazione e controllo.

Il Comstat esercita le funzioni direttive dell'Istat nei confronti degli uffici di statistica del Sistan, il Sistema statistico nazionale, mentre il Collegio dei revisori accerta la regolare tenuta della contabilità. L'Istituto svolge un ruolo di indirizzo, coordinamento, promozione e assistenza tecnica alle attività statistiche degli enti e degli uffici che fanno parte del Sistan.

L'Istat fa parte del Sistema Statistico Europeo e collabora con gli altri soggetti del sistema statistico internazionale.

Fino al 1998 la titolarità delle rilevazioni civili e penali era dell'Istat, successivamente è passata alla Direzione generale di statistica del Ministero della Giustizia. Sono rimaste in ogni caso di pertinenza dell'Istituto Nazionale di Statistica alcune rilevazioni afferenti al settore penale.

Di seguito vengono riportate le diverse analisi condotte dall'ISTAT (informazioni tratte da <<https://www.istat.it/it/informazioni-e-servizi/per-i-giornalisti/istat-in-breve>>).

Delitti per i quali l'Autorità giudiziaria ha iniziato l'azione penale

L'Istituto Nazionale di Statistica a partire dal 2000 effettua con cadenza annuale la *Rilevazione dei delitti e delle persone per i quali l'Autorità giudiziaria ha iniziato l'azione penale*. In tale rilevazione il principale fenomeno osservato è pertanto la criminalità e le unità di analisi sono: le persone denunciate per delitto, i procedimenti giudiziari penali, le contravvenzioni denunciate, i delitti denunciati, la parte offesa e le persone denunciate per contravvenzione.

La rilevazione è totale e deriva da fonti amministrative; in particolare viene acquisito il Registro SICP² gestito dal Ministero della Giustizia e utilizzato dagli uffici giudiziari per la registrazione dei procedimenti e della loro movimentazione; le unità di rilevazione sono le Procure della Repubblica presso i Tribunali.

Vengono effettuati dei controlli e delle correzioni dei dati attraverso la localizzazione deterministica degli errori e dei valori anomali e l'uso di varie tecniche di imputazione; si effettuano inoltre controlli di congruenza con precedenti dati della stessa indagine, controlli di congruenza tra indicatori e/o dati della stessa edizione dell'indagine relativi a diverse unità di analisi e controlli di congruenza con dati di altre indagini o altre fonti (informazioni disponibili al seguente link <<https://www.istat.it/it/archivio/giustizia+penale>>).

Minorenni denunciati per delitto

L'indagine rileva annualmente la criminalità minorile e le principali caratteristiche sociali e demografiche dei minori denunciati per aver commesso un reato. Le unità di analisi sono le contravvenzioni denunciate, i Minorenni denunciati per delitto, i delitti denunciati e i Minorenni denunciati per contravvenzione. I dati coprono l'intero territorio nazionale italiano e la periodicità delle stime è annuale, i dati sono disponibili a partire dal 2000 e sono estratti dal registro Re.Ge.; le unità di rilevazione sono le Procure della Repubblica presso il Tribunale per i Minorenni.

Nel trattamento dei dati viene prevista una correzione manuale degli errori effettuando un ritorno presso le unità di rilevazione, il controllo e correzione dei dati viene effettuato attraverso la localizzazione deterministica degli errori e dei valori anomali e l'applicazione di metodi di imputazione. Vengono inoltre effettuati controlli di congruenza con precedenti dati della stessa indagine e con dati di altre indagini o altre fonti (informazioni tratte dal seguente sito <<https://www.istat.it/it/archivio/234686>>).

Sicurezza delle donne

L'Istat da lungo tempo è impegnato nella misurazione del fenomeno della violenza di genere contro le donne. La prima indagine interamente

² Il Registro SICP ha sostituito il Re.Ge. La maggior parte degli uffici sono migrati da Re.Ge. a SICP nel 2014. Il distretto di Firenze è migrato da Re.Ge. a SICP a partire dal 2011, il distretto di Ancona nel 2013, insieme ad alcune sedi di altri distretti. In alcune sedi la migrazione è avvenuta nel 2015.

ed esplicitamente dedicata alla violenza sulle donne – denominata Indagine sulla sicurezza delle donne – è stata condotta dall'Istat nel 2006, con il contributo finanziario del Ministero per le pari opportunità, grazie all'attiva collaborazione progettuale dei centri antiviolenza e anche con il supporto di alcune donne vittime di violenza.

L'indagine Sicurezza delle donne nasce con l'intento di far luce sul fenomeno della violenza contro le donne in Italia in tutte le sue diverse forme, in termini di prevalenza ed incidenza, ed in particolare la violenza domestica.

Attraverso le informazioni derivanti da tale indagine l'Istat può fornire dati importanti ai fini delle politiche di prevenzione e contrasto della violenza di genere, documentando la diffusione della violenza fisica sessuale e psicologica, la tipologia di autori, la gravità della violenza, il luogo, le conseguenze.

La seconda edizione dell'indagine sulla violenza contro le donne è stata effettuata nel 2015. Più in particolare, l'indagine si propone di rilevare e descrivere:

- l'estensione e le caratteristiche del fenomeno della violenza extra-familiare e della violenza domestica e quindi il numero, la dinamica e le peculiarità dei diversi episodi di violenza;
- il periodo in cui si è verificata la violenza, ad esempio nel corso della vita o nei 12 mesi;
- le caratteristiche delle vittime, la loro reazione all'episodio di violenza e le conseguenze fisiche, psicologiche ed economiche delle violenze che hanno subito;
- le caratteristiche degli autori delle violenze, con particolare attenzione agli autori delle violenze in famiglia;
- l'incidenza del sommerso, a partire da quanti degli eventi subiti sono stati effettivamente denunciati alle forze dell'ordine e i motivi della non denuncia;
- la dinamica dell'evento e la storia della relazione di coppia nei casi in cui la violenza è agita in famiglia o, comunque, da un partner della donna;
- la violenza subita prima dei 16 anni.

Dai dati è possibile individuare e stimare i possibili fattori di rischio e quelli protettivi a livello individuale e sociale e i costi sociali della violenza, riconducibili direttamente e indirettamente alla donna e gli eventuali figli, al maltrattante e alla società, misurati attraverso alcune ricadute negative come l'impossibilità della vittima di condurre le normali attività quotidiane,

di lavorare, l'utilizzo dei servizi sociali e sanitari, o i costi direttamente sostenuti per far fronte ai danni conseguenti alla violenza (spese per cure mediche o psicologiche, per danni materiali o di tipo legale), nonché alcune informazioni attinenti i possibili costi legati all'iter giudiziario.

L'indagine campionaria dell'Istat sulla sicurezza delle donne fornisce anche la stima delle donne che hanno subito atti persecutori (stalking: una condotta reiterativa caratterizzata da molestie e/o minacce dirette a una persona che producono effetti coartanti sulla libertà psichica della vittima e un'indesiderata intrusione nella sua sfera individuale).

Nella seconda indagine per la prima volta è stato applicato un disegno campionario con tecnica di indagine mista, che consente le stime sulla violenza per alcune cittadinanze straniere. Per ulteriori dettagli si veda la nota metodologica (informazioni tratte da <<https://www.istat.it/it/archivio/237805>>).

Servizi offerti dai Centri antiviolenza alle donne vittime

L'Istat ha svolto per la prima volta l'indagine sui servizi offerti dai Centri antiviolenza alle donne vittime, in collaborazione con il Dipartimento per le Pari opportunità (Dpo) presso la Presidenza del Consiglio, le regioni e il Consiglio nazionale delle ricerche (Cnr – Irrps). L'indagine è stata effettuata nei mesi di giugno – luglio 2018 e sono stati intervistati 281 centri antiviolenza rispondenti ai requisiti dell'Intesa del 2014. Tra questi 253 hanno completato il questionario. I centri antiviolenza forniscono in prevalenza servizi di ascolto e accoglienza, supporto legale, orientamento e accompagnamento ad altri servizi, supporto psicologico, aiuto nel percorso di allontanamento dal partner violento, orientamento lavorativo, sostegno all'autonomia. In alcuni casi il servizio è fornito direttamente dal centro, in altri, dal centro in collaborazione con i servizi sul territorio, in altri ancora, il centro assolve la funzione di indirizzamento (informazioni disponibili al seguente link <[https://temi.camera.it/leg18/post/OCD15_14623/sintesi-risultanze-indagini-istat-centri-antiviolenza-e-case-rifugio-anno-2018.html#:~:text=I%20pi%C3%B9%20frequenti%20sono%20quelli,lavorativo%20\(80%2C5%25\)](https://temi.camera.it/leg18/post/OCD15_14623/sintesi-risultanze-indagini-istat-centri-antiviolenza-e-case-rifugio-anno-2018.html#:~:text=I%20pi%C3%B9%20frequenti%20sono%20quelli,lavorativo%20(80%2C5%25)>)>).

Incidenti stradali con lesione a persone

La rilevazione riguarda gli incidenti stradali verificatisi nell'arco di un anno solare sull'intero territorio nazionale, verbalizzati da un'autorità di polizia. L'incidente stradale viene definito come “quell'evento in cui è coinvolto almeno un veicolo sulla rete stradale, verificatosi nelle vie o piazze aperte alla circolazione, che comporti lesioni a persone (morti entro

30 giorni e/o feriti)” - (Convenzione di Vienna del 1968, UNECE, ITF, Eurostat). La rilevazione ha come riferimento normativo la Decisione del Consiglio europeo n. 704 del 30 novembre 1993 relativa alla creazione di una banca di dati comunitaria sugli incidenti stradali.

L'informazione statistica sull'incidentalità è raccolta dall'Istat mediante una rilevazione totale a cadenza mensile, nel caso di specifici accordi e convenzioni avviene con cadenza trimestrale. L'unità di rilevazione è il singolo incidente stradale. La rilevazione è riferita al momento in cui l'incidente si è verificato. Le variabili rilevate riguardano i principali aspetti del fenomeno: data e località dell'incidente, organo di rilevazione, localizzazione dell'incidente, tipo di strada, segnaletica, fondo stradale, condizioni meteorologiche, natura dell'incidente, tipo di veicoli coinvolti, circostanze dell'incidente, conseguenze dell'incidente alle persone. Per la raccolta dei dati e il monitoraggio delle informazioni sull'incidentalità stradale, l'Istat ha sottoscritto degli accordi con una pluralità di enti tra cui ACI, Ministero delle Infrastrutture e dei Trasporti, Ministero dell'Interno – Servizio di Polizia Stradale, Ministero della Difesa - Carabinieri, Polizia Provinciale e Locale, Regioni e Province autonome. In conformità a quanto disposto dal *Codice dell'amministrazione digitale* (art. 47 del d.lgs. n. 82 del 2005), i dati devono essere inviati all'Istat in modalità informatizzata. A partire dal 2019, l'Istat mette a disposizione delle Polizie Locali un nuovo sistema di acquisizione dati Istat *Gino*, disponibile all'indirizzo <<https://gino.istat.it/incidenti>>. I comandi di Polizia locale possono accedere al sistema, dove è possibile inserire online le informazioni relative a ciascun sinistro rilevato, attraverso la compilazione guidata di un questionario web. Utilizzando questo sistema, le Polizie Locali e gli Organi intermedi di rilevazione hanno la possibilità di utilizzare anche la funzione di *upload* per caricare un file di formato *.csv*, secondo un tracciato record reperibile nella sezione “Documenti e istruzioni” disponibile all'indirizzo citato. La presente rilevazione è inserita nel Programma statistico nazionale e quindi l'Istat è tenuta a svolgerla per legge. La normativa vigente prevede la facoltà di rispondere o meno ad alcuni quesiti di natura “sensibile” indicati sulla circolare e sull'informativa all'interessato (informazioni disponibili al seguente link <<https://www.istat.it/it/archivio/4609>>).

Titoli protestati

L'Istat effettua a partire dal 2000 e mensilmente una rilevazione totale di tipo amministrativo che copre l'intero territorio nazionale sui titoli di credito protestati (cambiali, tratte e assegni). La tecnica di acquisizione dei dati consiste in un archivio informatizzato di dati pre-elaborati su dati

statistici rilevati con l'atto amministrativo da enti intermedi.

In dettaglio vengono acquisiti informazioni dal Registro dei protesti (REPR) alimentato da tutte le Camere di Commercio d'Italia. Il controllo e la correzione dei dati si basa sulla localizzazione deterministica di errori e valori anomali e imputazione mediante regole deterministiche; viene inoltre effettuato un controllo di congruenza dei dati con precedenti dati della stessa indagine (informazioni disponibili al seguente link <<https://www.istat.it/it/archivio/234704>>).

Condannati per delitto con sentenza irrevocabile

Il diritto penale sanziona comportamenti illeciti in base alla gravità, connotandoli come delitti o come contravvenzioni, i primi sono più gravi.

Il numero delle condanne per delitto rappresenta un importante aspetto del complesso fenomeno della criminalità. Il focus informativo è fortemente legato all'attività delle Autorità Giudiziarie, rappresentando l'output finale del loro lavoro. Il potenziale informativo è quindi meno legato all'effettiva incidenza della criminalità rispetto ad altre fonti di misurazione della stessa, quali ad esempio le denunce di reato alle forze dell'ordine.

La rilevazione riguarda la totalità dei condannati con sentenza irrevocabile per delitto commesso più grave, viene offerto un focus sui condannati di età inferiore ai 18 anni, ed infine il quadro complessivo di tutti i delitti che sono stati oggetto del complesso delle condanne.

L'unità di rilevazione è il Registro delle Condanne con sentenza definitiva del Casellario Centrale Giudiziale alimentato dai dati dei Casellari Giudiziali Locali. Le unità di analisi sono le condanne per delitti con riferimento all'iscrizione della condanna con sentenza irrevocabile nel Casellario Giudiziale. Le principali variabili osservate sono la tipologia del delitto, l'Autorità che ha emesso la sentenza e grado della sentenza, luogo e momento del commesso delitto, principali caratteristiche dell'autore (genere, età, luogo di nascita), principali misure di sicurezza, tipo e durata della pena inflitta. La periodicità dell'indagine era annuale. La rilevazione, iniziata nel 1980, ha assunto la configurazione attuale a partire dal 1996. Tuttavia a partire dal 2007 vi è stata un'interruzione della serie storica a causa di modifiche della base dati del Casellario e a causa del diverso trattamento metodologico e informatico del dato (informazioni tratte dai seguenti link: <<https://www4.istat.it/it/archivio/52159>>; <https://www4.istat.it/it/files/2012/02/nota_descrittiva.pdf?title=Condanne+con+sentenza+irrevocabile+-+01%2Ffeb%2F2012+-+Nota+descrittiva.pdf>).

Separazioni personali dei coniugi

L'Istat conduce la rilevazione delle separazioni personali dei coniugi e la rilevazione degli scioglimenti e cessazione degli effetti civili del matrimonio. I dati rilevati consentono di analizzare le diverse componenti del fenomeno: procedimento e ricorso all'assistenza legale, matrimonio all'origine, caratteristiche dei coniugi ed eventuali figli coinvolti, provvedimenti economici quali il contributo per il mantenimento e l'assegnazione della casa coniugale. Queste informazioni permettono di far luce sulle modifiche strutturali e sulle tendenze recenti relativamente allo scioglimento dei matrimoni formalizzato a seguito di un iter giudiziario o extra-giudiziale.

L'Istat rileva le separazioni e i divorzi condotti presso le cancellerie civili dei tribunali della Repubblica, raccogliendo i dati relativi a ogni singolo procedimento concluso dal punto di vista giudiziario nell'anno di riferimento.

Per le separazioni e i divorzi effettuati presso i tribunali la rilevazione è organizzata con due modalità distinte di raccolta dei dati:

- in ottemperanza agli obblighi di legge previsti in materia di “dema-terializzazione delle comunicazioni tra pubbliche amministrazioni” la modalità principale è la raccolta dei dati via web, tramite la piattaforma Istat-Gestione Indagini Online (Gino), attiva dal mese di aprile 2018;
- in caso di impedimento oggettivo nell'utilizzo di questa modalità è prevista, in alternativa, la compilazione e l'invio dei dati con la classica modalità cartacea.

Per le separazioni e i divorzi effettuati presso i comuni è prevista la compilazione di modelli individuali elettronici. La compilazione, acquisizione e trasmissione dei modelli avviene in questo caso esclusivamente per via telematica a cura degli ufficiali di stato civile, attraverso un'utenza personalizzata. Sia le circolari, sia le istruzioni sono disponibili nella home page delle due piattaforme di acquisizione insieme ad una breve descrizione delle indagini.

Per le separazioni e i divorzi effettuati presso i tribunali è prevista la compilazione di modelli individuali che vengono trasmessi all'Istat entro il giorno 10 del mese successivo al trimestre cui si riferiscono i dati.

Per le separazioni e i divorzi effettuati presso i comuni la rilevazione ha una periodicità mensile ed i modelli devono essere trasmessi entro la fine del mese successivo alla conclusione del procedimento che porta alla separazione/divorzio dei coniugi.

Le suddette rilevazioni sono inserite nel Programma statistico nazionale

2017-2019 approvato con DPR 31 gennaio 2018, e nel Programma statistico nazionale 2017-2019 – Aggiornamento 2018-2019 e vi è obbligo di risposta (informazioni tratte dal seguente sito <<https://www.istat.it/it/archivio/8758>>).

Sicurezza dei cittadini

L'indagine Sicurezza dei cittadini è un'indagine campionaria periodica, che ha lo scopo di rilevare quanto le persone si sentono sicure nel proprio ambiente di vita e quanto siano diffusi alcuni fatti come, ad esempio, il furto del portafoglio o dell'automobile.

La rilevazione fa parte del sistema integrato di Indagini Multiscopo sulle famiglie. È realizzata tramite interviste telefoniche e, per le persone prive di un recapito telefonico, con interviste faccia a faccia.

La rilevazione prevede alcuni quesiti sensibili per i quali non c'è obbligo di risposta e ai quali, quindi, la persona intervistata può rifiutarsi di rispondere.

La rilevazione rientra tra quelle comprese nel Programma statistico nazionale, che raccoglie l'insieme delle rilevazioni statistiche necessarie al Paese.

Vengono intervistate le persone dai 14 anni in su, ogni persona viene estratta con criterio di scelta casuale dalle Liste anagrafiche comunali (LAC), il campione è costituito da circa 50.000 persone distribuite in tutto il territorio nazionale. Le informazioni vengono raccolte tramite un questionario elettronico e con tecnica di rilevazione *Computer Assisted* (informazioni tratte dal sito <<https://www.istat.it/it/archivio/164581>>).

4. Il Ministero della Giustizia e la Direzione Generale di Statistica (DGStat)

L'art. 110 della Costituzione affida al Ministro della Giustizia l'organizzazione ed il funzionamento dei servizi relativi alla Giustizia.

Il Ministero dunque si occupa dell'organizzazione giudiziaria, del personale e dei servizi occorrenti all'esercizio dell'attività giudiziaria, ed ha funzioni amministrative strettamente connesse alle funzioni giurisdizionali, sia nel campo civile che penale. Cura, inoltre, la gestione degli archivi notarili, la vigilanza sugli ordini e collegi professionali, l'amministrazione del casellario, la cooperazione internazionale in materia civile e penale, l'istruttoria delle domande di grazia da proporre al Presidente della Repubblica.

Nel settore penitenziario, il Ministero della giustizia svolge compiti di

attuazione della politica dell'ordine e della sicurezza negli istituti e servizi penitenziari e nel trattamento dei detenuti, nonché di amministrazione del personale penitenziario e di gestione dei supporti tecnici necessari. Il Ministero ha inoltre competenza in materia di minori e di gestione del personale e dei beni relativi al settore della giustizia minorile e dell'esecuzione penale esterna.

Il Ministero della Giustizia è il soggetto che si occupa a livello nazionale di statistica giudiziaria. Come ben dettagliato nel sito, i compiti del Ministero della Giustizia, oltre a quelli previsti dal decreto legislativo 6 settembre 1989, n. 322 (Norme sul Sistema Statistico Nazionale e sulla riorganizzazione dell'Istat), consistono in: realizzazione e gestione di banche dati di statistica giudiziaria; redazione del programma statistico nazionale attraverso l'individuazione di criteri unici e omogenei sul territorio nazionale; assicurare a livello centrale, per i rapporti con l'ISTAT, il coordinamento con gli uffici del Ministero, gli uffici amministrativi decentrati e gli uffici giudiziari; rapporti con organismi europei e internazionali di settore.

Nel 2001 è stata istituita all'interno del Ministero della Giustizia la Direzione Generale di Statistica - attualmente Direzione Generale di statistica e analisi organizzativa (DGStat) - che si occupa dell'analisi statistica del sistema giudiziario. In particolare, ai sensi del decreto legislativo 322/1989 la DGStat è ufficio del Sistema Statistico Nazionale (SISTAN) ed è titolare dal gennaio 1998 delle rilevazioni per il settore civile, dal gennaio 1999 di quelle relative al settore penale e dal gennaio 2003 di quelle del settore amministrativo contabile.

Per le rilevazioni civili e penali la titolarità è passata dall'Istat alla DGStat, mentre le rilevazioni del settore amministrativo-contabile sono di nuova istituzione e costituiscono, insieme alle precedenti, il presupposto essenziale per interventi di correzione e di miglioramento dell'intera organizzazione.

La DGStat rappresenta uno strumento di trasparenza per i cittadini, in grado di fornire, rendendole disponibili all'interno e all'esterno, informazioni aggiornate sui dati e sui flussi fondamentali della domanda di giustizia nel nostro Paese e sulla capacità di risposta del sistema.

È, inoltre, l'unità organizzativa di riferimento per la statistica giudiziaria nei confronti di numerosi enti, istituzioni e organizzazioni nazionali e internazionali, tra i quali il Parlamento, il Governo, il Consiglio Superiore della Magistratura, il Consiglio d'Europa e la sua Commissione per l'efficienza dei sistemi giudiziari europei (Cepej), la Banca Mondiale, le università e gli organi d'informazione.

La DGStat si occupa delle statistiche relative ai flussi giudiziari

(sopravvenienze, definizioni e pendenze³), all'arretrato e alla durata media degli affari del settore civile e penale, i cui dati provengono da fonte amministrativa.

Le elaborazioni vengono messe a disposizione dei cittadini e degli utenti professionali (come uffici giudiziari, magistrati, avvocati) sul sito del Ministero e sulla pagina dedicata alle statistiche (informazioni tratte dal sito del Ministero Giustizia).

Vengono di seguito riportati alcuni dettagli sulla tipologia di informazioni disponibili e sul processo di produzione dei dati nei settori civile, penale e nella mediazione civile e commerciale.

Area civile – Monitoraggi trimestrali, mensili, serie storica

I dati statistici in ambito civile provengono da fonte amministrativa ovvero dai registri informatizzati utilizzati dagli uffici giudiziari per la gestione dei procedimenti.

Per il Tribunale e la Corte d'appello la fonte dei dati ministeriali è il DWGC (*Datawarehouse della Giustizia Civile*) che raccoglie tutte le informazioni presenti nei registri SICID (Sistema Informativo della Cognizione Civile Distrettuale) e SIECIC (Sistema Informativo delle Esecuzioni Civili Individuali e Concorsuali) in uso negli uffici giudiziari; sono pertanto disponibili informazioni relative al singolo procedimento.

Per il Giudice di Pace e il Tribunale per i minorenni il processo di acquisizione prevede l'invio dei dati aggregati da parte degli uffici giudiziari alla DGStat con cadenza trimestrale attraverso appositi modelli di rilevazione. Nei casi di inadempienza, al fine di garantire la completezza e la consistenza della serie storica, vengono stimati i dati degli uffici che non hanno compilato i modelli con opportune tecniche di interpolazione statistica basate sul loro trend storico.

Per la Cassazione i dati aggregati sono comunicati dall'Ufficio statistico, ivi distaccato, a cui compete l'attività di estrazione e la successiva attività di elaborazione.

Nell'area pubblica del sito del Ministero della Giustizia sono pubblicati i monitoraggi periodici, le analisi relative a specifici istituti quali ad esempio la mediazione civile e commerciale o le procedure di composizione della crisi da sovra-indebitamento meglio descritte di seguito, nonché gli studi, le analisi e le ricerche della DGStat su alcuni argomenti di interesse.

È pertanto disponibile un monitoraggio distrettuale con cadenza

³ Le misure statistiche sono definite nel Capitolo 3.

trimestrale – relativo a Tribunali e Corti d'Appello - con le informazioni sulle iscrizioni (cioè i nuovi fascicoli aperti), sulle definizioni (fascicoli esitati con sentenza o con altre modalità di definizione), sulle pendenze (cioè il numero di procedimenti ancora aperti alla fine del periodo), nonché alcuni indicatori come l'indice di ricambio (*clearance rate*) meglio descritto successivamente e la variazione delle pendenze.

Il monitoraggio fornisce trimestralmente anche informazioni sull'andamento dell'arretrato civile, attraverso la cosiddetta 'stratigrafia' delle pendenze (divise per anno di iscrizione). Occorre precisare che per il calcolo dell'arretrato rilevano i soli procedimenti che, alla data di riferimento, non sono stati risolti entro i termini di ragionevole durata previsti dalla legge (cosiddetti procedimenti "a rischio Pinto"⁴). I termini di ragionevole durata variano a seconda del grado di appello e sono i seguenti:

- 3 anni per i procedimenti in primo grado;
- 2 anni per i procedimenti in appello;
- 1 anno per i procedimenti in Cassazione.

Nei monitoraggi trimestrali civili vengono esclusi dall'analisi statistica i procedimenti del giudice tutelare, l'accertamento tecnico preventivo in materia previdenziale (ATP) e i procedimenti di 'ricevimento e verbalizzazione di dichiarazione giurata' in quanto:

- la definizione dei procedimenti del giudice tutelare non dipende dall'attività del giudice, ma dalla cessazione della misura in essere;
- l'Accertamento Tecnico Preventivo (ATP) in materia di previdenza costituisce un'attività di filtro preventivo che il giudice delega a specifiche commissioni tecniche, prima che un eventuale contenzioso possa essere avviato innanzi al Tribunale;
- l'attività di ricevimento e verbalizzazione di dichiarazione giurata è di carattere prettamente amministrativo.

Vien pubblicato anche un monitoraggio mensile con l'informazione relativa a iscrizioni e definizioni.

Inoltre, è pubblicata una serie storica con i flussi dei procedimenti classificati per macromateria, materia e un ulteriore dettaglio. A differenza dei monitoraggi trimestrali e mensili, nella serie storica viene fornita

⁴ La cosiddetta legge Pinto, ovverosia la [legge numero 89 del 24 marzo 2001](#), è un provvedimento che riconosce a coloro che hanno dovuto affrontare un processo di durata irragionevole la possibilità di richiedere un'equa riparazione per il danno patrimoniale o non patrimoniale subito. Si tratta, chiaramente, di uno strumento processuale volto a combattere il fenomeno, assai diffuso in Italia, della lunghezza eccessiva dei processi.

l'informazione relativa ai:

- sopravvenuti: come somma di procedimenti nuovi iscritti e dei sopravvenuti derivanti a seguito di eventi di riassunzione, separazione fascicolo, mutamento rito;
- definiti: come somma di procedimenti definiti con sentenza definitiva e procedimenti definiti senza sentenza; in particolare, i definiti con sentenza definitiva conteggiano le sentenze in base alla data di pubblicazione, comprendendo i dispositivi letti in udienza e le sentenze a verbale; i definiti senza sentenza includono tutti gli altri eventi di definizione alternativi alla sentenza (ad esempio riunioni, improcedibilità, cancellazione, decreti di accoglimento o rigetto), compresi i definiti per mutamento/trasformazione rito (ad esempio trasformazione di separazione consensuale in giudiziale e viceversa) e i sospesi ed interrotti;
- pendenti a fine periodo;
- durata media effettiva.

Con cadenza annuale vengono pubblicati i dati dei flussi dei procedimenti civili dei Giudici di Pace dettagliati per macromateria e materia.

Mediazioni civili e commerciali

Ai sensi del d.lgs. n. 28 del 2010 e successive modificazioni, la mediazione è condizione di procedibilità per le seguenti materie: condominio, diritti reali, divisione, successioni ereditarie, patti di famiglia, locazione, comodato, affitto di aziende, risarcimento del danno derivante da responsabilità medica e sanitaria e da diffamazione con il mezzo della stampa o con altro mezzo di pubblicità, contratti assicurativi, bancari e finanziari.

La DGStat effettua il monitoraggio statistico dei procedimenti di mediazione civile trattati presso gli organismi abilitati. La rilevazione sull'andamento delle iscrizioni ha preso avvio con l'entrata in vigore del suddetto Decreto Legislativo e copre tutte le tipologie di mediazione (obbligatoria in quanto condizione di procedibilità, volontaria e demandata dal giudice). La procedura di acquisizione dei dati prevede la compilazione on-line (ovvero l'invio dei fogli elettronici) di due modelli: 1. modello trimestrale di rilevazione dei flussi di mediazione (maschera d'acquisizione denominata "Riepilogo") che rileva i flussi dei procedimenti trattati da ciascun organismo nel periodo di riferimento (pendenti iniziali, iscritti, definiti, pendenti finali) con indicazione della materia del procedimento (diritti reali, divisione etc.); 2. scheda per singolo procedimento di mediazione definito

(maschera d'acquisizione denominata "Scheda mediazione") che raccoglie le informazioni di sintesi relative a ciascun procedimento concluso: esito dei procedimenti, durata, personalità giuridica delle parti coinvolte, materia, incontro preliminare di programmazione e numero complessivo incontri, procedimenti demandati dal giudice anche per le materie non obbligatorie, indennità corrisposte agli Organismi di mediazione, etc.

I dati sono pubblicati a cadenza trimestrale e sono elaborati effettuando apposite stime per gli uffici non rispondenti.

Organismi di composizione della crisi da sovraindebitamento

La rilevazione è stata introdotta nel 2017 e raccoglie i flussi sui procedimenti di "composizione della crisi da sovraindebitamento" (accordo, piano del consumatore, liquidazione del patrimonio) e le istanze di esdebitazione trattate presso gli organismi abilitati. I dati sono trasmessi con cadenza annuale dagli Organismi di composizione della crisi presenti nel Registro istituito dal comma 2° dell'art. 15 della L.3/2012 e tenuto presso il Dipartimento per gli affari di giustizia del Ministero della Giustizia.

Area penale – I flussi dei procedimenti

Anche per l'area penale, i dati provengono principalmente da fonti amministrative, ovvero dai registri informatizzati per la gestione dei procedimenti penali e dagli estrattori appositamente predisposti per la produzione dei flussi statistici. A differenza del settore civile, dove è stato sviluppato un sistema di analisi statistica (il datawarehouse della giustizia civile sopra descritto), nell'area penale il processo di acquisizione dei dati prevede l'invio da parte degli uffici giudiziari alla DGStat con cadenza trimestrale dei flussi statistici aggregati. Anche in questo caso, al fine di garantire la completezza e la consistenza della serie storica, vengono stimati i dati degli uffici inadempienti.

Per la Cassazione i dati sono comunicati dall'Ufficio statistico alla DGStat.

Nell'area pubblica del sito del Ministero della Giustizia (<<https://webstat.giustizia.it>>) sono pubblicati i report con i flussi dei procedimenti penali degli uffici giudicanti e requirenti; le principali modalità di definizione dei procedimenti penali; la durata effettiva dei procedimenti penali.

Le misure statistiche sui flussi cambiano in base agli uffici di riferimento. Ad esempio, i sopravvenuti sono ottenuti:

- in Procura della Repubblica come somma dei procedimenti iscritti

nel registro Noti nel periodo di riferimento in base alla “data di iscrizione”, dei procedimenti restituiti dal GIP o dal giudice del dibattimento all’ufficio PM nel periodo in base alla “data restituzione atti” e dei procedimenti derivanti da eventuali “stralci” effettuati nel periodo di riferimento in base alla “data di iscrizione”;

- nella sezione GIP/GUP del Tribunale come somma dei procedimenti che pervengono al GIP con una richiesta definitiva da parte del PM in base alla “data di pervenimento della richiesta definitiva” al GIP e dei procedimenti restituiti dal dibattimento ma già considerati esauriti nello stesso periodo o in periodi precedenti in base alla “data restituzione atti”;
- in Tribunale Ordinario – sez. dibattimento, in Corte di assise, al Giudice di Pace e in Corte di Appello come somma dei procedimenti iscritti nel registro Noti nel periodo di riferimento in base alla “data di iscrizione” e dei procedimenti derivanti da eventuali “stralci” effettuati nel periodo di riferimento e conteggiati attraverso la “data di iscrizione”.

I definiti sono ottenuti:

- in Procura della Repubblica come somma dei procedimenti per i quali è stata emessa dal PM una richiesta definitiva, a prescindere dall’esito della stessa, in base alla data della richiesta definitiva del PM, ad eccezione delle richieste di citazione diretta a giudizio in cui viene considerata la data di trasmissione atti al dibattimento;
- nella sezione GIP/GUP del Tribunale come somma dei procedimenti definiti nel periodo di riferimento con decreto, ordinanza o sentenza in base alla data di deposito della sentenza se il procedimento è definito con sentenza; alla data trasmissione atti al dibattimento se il procedimento è definito con decreto che dispone il giudizio, alla data di esecutività del decreto penale di condanna se il procedimento è definito con decreto penale; alla data del provvedimento emesso per tutte le altre tipologie di provvedimento (decreti di archiviazione, ordinanze di restituzione atti, riunioni);
- in Tribunale Ordinario – sez. dibattimento, in Corte di assise, al Giudice di Pace e in Corte di Appello come somma dei procedimenti definiti nel periodo di riferimento con decreto, ordinanza o sentenza in base alla data di deposito della sentenza se il procedimento è definito con sentenza e alla data del provvedimento emesso per tutte le altre tipologie di provvedimento (ordinanze di trasmissione atti, riunioni).

I pendenti finali sono ottenuti conteggiando tutti i procedimenti iscritti nel registro per i quali ancora non è stato emesso un provvedimento definitivo alla data di osservazione.

Per la Sorveglianza vengono raccolti semestralmente i dati relativi all'attività giurisdizionale dei 29 Tribunali e dei 58 Uffici di Sorveglianza. I dati sono estratti dal registro informatizzato SIUS e le grandezze analizzate - per anno, distretto e sede - sono le misure concesse (provvedimenti di accoglimento di istanza di misure alternative alla detenzione carceraria, o di differimento sospensione della pena o di espulsione dal territorio nazionale), le misure dichiarate inefficaci (declaratorie di inefficacia della esecuzione delle misure alternative o di sospensione), le misure revocate, le misure cessate per sopravvenienza di nuovi titoli di condanna, quelle cessate per altri motivi e le misure in esecuzione.

Le intercettazioni

L'intercettazione è un mezzo di ricerca della prova previsto e disciplinato dall'art. 266 e seguenti del codice di procedura penale. La DGStat ha avviato un monitoraggio statistico periodico a partire dal 2003 presso gli uffici giudiziariquirenti: Procure Generali presso le Corti di Appello, Procure della Repubblica presso i Tribunali per i minorenni e Procure della Repubblica presso i Tribunali ordinari. I dati vengono comunicati dagli Uffici giudiziari alla DGStat su base trimestrale. Le variabili osservate sono i bersagli distinti per tipologia di intercettazione (telefonica, ambientale e di altro tipo), tipologia di reato (ordinario, mafia, terrorismo) e ufficio di provenienza, l'importo liquidato dagli uffici giudiziari al netto dell'IVA, la durata media delle intercettazioni sulla base dei decreti di autorizzazione e di proroga emessi.

Il monitoraggio PNRR

Il Piano Nazionale di Ripresa e Resilienza (PNRR) prevede nel settore giudiziario degli obiettivi di riduzione della durata dei procedimenti civili e penali e, limitatamente agli uffici di merito e al settore civile, di riduzione dell'arretrato (cfr. paragrafo 5 del capitolo VI). Il Ministero ha, pertanto, dedicato una sessione del sito alla pubblicazione del monitoraggio periodico dei risultati raggiunti dagli uffici. In particolare, vengono messe a disposizione del cittadino e degli uffici giudiziari le informazioni sull'andamento degli indicatori relativi ai Tribunali, alle Corti d'Appello e alla Corte di Cassazione.

5. Il Consiglio Superiore della Magistratura e l'Ufficio statistico

Il Consiglio Superiore della Magistratura (CSM) è un organo di rilievo costituzionale che si occupa dell'amministrazione della giurisdizione e di garantire l'autonomia e l'indipendenza della magistratura. Tra gli organi consiliari permanenti previsti dal Regolamento del CSM vi è l'Ufficio Statistico con il compito di elaborare i dati inerenti agli uffici giudiziari, ai magistrati e alle attività del Consiglio, assicurarne l'aggiornamento e garantirne l'uniformità, così fornendo i supporti di conoscenza alle Commissioni e alle altre articolazioni consiliari.

Un'apposita delibera del Consiglio ha disciplinato il funzionamento e la dotazione dell'Ufficio statistico, determinandone altresì il coordinamento dell'attività con le funzioni della Struttura Tecnica per l'Organizzazione (STO).

La Struttura Tecnica dell'Organizzazione, prevista dall'art. 18 del Regolamento interno del Csm, è un organo consiliare permanente composto da dieci magistrati nominati dal Consiglio, previo interpello, tra coloro che abbiano maturato esperienza in tema di organizzazione e informatizzazione degli uffici, di analisi dei flussi, di carichi di lavoro e di pendenze dei procedimenti e dei processi. L'incarico ha durata triennale e può essere rinnovato, a richiesta, per non più di una volta e per non più di un anno. La STO è presieduta dal Presidente *pro tempore* della Settima commissione, è coordinata e diretta dalla commissione stessa ed ha diversi compiti tra cui quelli di acquisire ed analizzare informazioni sui carichi di lavoro, sui flussi e sulle pendenze dei procedimenti e dei processi sia a livello nazionale sia distrettuale, al fine di verificare l'efficienza e l'efficacia dei progetti di organizzazione degli uffici giudiziari e di effettuare adeguate comparazioni tra gli stessi; e di fornire supporto al Consiglio sia per la verifica dei risultati operativi ottenuti attraverso le attività di indirizzo e di regolamentazione, sia per la definizione periodica del fabbisogno informativo e formativo in questo settore, sia per l'interscambio di dati con il Ministero della Giustizia e le sue articolazioni. La STO può svolgere la propria attività in coordinamento con l'Ufficio statistico del Consiglio.

L'Ufficio Statistico collabora con la Struttura tecnica per l'organizzazione nella predisposizione delle rilevazioni di interesse, partecipa alle sedute della stessa per i profili di competenza e fornisce il risultato delle analisi dei dati. Le articolazioni del Consiglio interloquiscono direttamente con l'Ufficio Statistico per l'affidamento di incarichi e progetti di ricerca. In merito alle richieste provenienti da altri Enti o da privati, il Comitato di Presidenza

autorizza l'Ufficio Statistico alle elaborazioni necessarie e all'eventuale invio dei risultati.

I principali compiti dell'Ufficio Statistico concernono:

- analisi dei flussi di lavoro delle Commissioni e della tempestività delle procedure di evasione delle pratiche di interesse del Consiglio;
- supporto alla individuazione delle sedi da pubblicare nei bandi per la copertura delle vacanze nell'organico di magistrati degli uffici giudiziari nonché dei posti da assegnare ai Magistrati Ordinari in Tirocinio;
- analisi dei flussi di lavoro degli uffici requirenti e giudicanti di supporto alle pratiche consiliari;
- monitoraggio della performance degli uffici giudiziari rispetto a obiettivi prefissati;
- rilevazione dei dati utili alla redazione del format in attuazione della Risoluzione sui programmi di gestione ex art. 37 del D. Lgs. n. 98/2011; analisi dei format inviati dagli Uffici Giudiziari e dai Consigli Giudiziari;
- rilevazioni su temi di interesse, ad es. ufficio del processo, violenza di genere, questionario ai RID, misure di prevenzione;
- valutazione sull'impatto di alcune riforme (modifica delle piante organiche, modifica delle competenze territoriali, ecc.);
- progettazione, sviluppo e gestione del *datawarehouse* che integrerà i flussi dei procedimenti in materia civile e penale degli uffici giudiziari e dei singoli magistrati con le pratiche in carico al CSM.

Per lo svolgimento di tali attività il CSM ha stipulato un protocollo di intesa con il Ministero della Giustizia sulla base del quale il Ministero pone a disposizione del CSM i dati acquisiti nel sistema statistico del dominio giustizia. Inoltre, al fine di rendere più agevole e completa la consultazione del sistema statistico del dominio giustizia da parte del CSM per gli scopi istituzionali propri e degli uffici giudiziari, è prevista la realizzazione, a cura del Consiglio, di un datamart che permetterà autonome elaborazioni sui dati statistici relativi ai magistrati e al sistema tabellare.

Il datawarehouse civile e penale del Csm

Nell'ambito del progetto di reingegnerizzazione del sistema informativo del CSM è prevista la realizzazione di un sistema informatizzato centrale di raccolta dei dati (datawarehouse del CSM) al fine di consentire le

elaborazioni statistiche di supporto alle attività del Consiglio.

Il *datawarehouse* (dwh) del Consiglio sarà alimentato dal datawarehouse ministeriale civile, dai dati delle statistiche comparate inviate dagli uffici giudiziari e dalle banche dati anagrafiche del Consiglio; esso rappresenterà la principale fonte per ogni valutazione statistica consiliare e dovrà essere utilizzato in tutte le procedure consiliari che richiedano una elaborazione numerica dei flussi di lavoro. Con tale strumento sarà possibile, inoltre, fornire elaborazioni e dati statistici agli uffici giudiziari e ai magistrati. Saranno disponibili report sui flussi dei procedimenti civili e penali, per ufficio, per sezione e per singolo magistrato.

Il nuovo dwh si presenta come un sistema servente rispetto alle funzioni già esercitate dal CSM e dal circuito complessivo dell'autogoverno. La novità introdotta dal nuovo sistema riguarda la raccolta delle informazioni dagli uffici giudiziari, raccolta che già il CSM operava in varie occasioni e momenti e che con il dwh viene *automatizzata*: l'alimentazione del nuovo sistema statistico del CSM sarà costituita pertanto da flussi automatici di dati provenienti dagli uffici giudiziari senza alcun ulteriore onere di raccolta o di elaborazione dei dati a carico degli uffici stessi. Ciò rappresenterà anche un servizio reso agli uffici, ai Consigli Giudiziari, nonché alle Commissioni Flussi costituite presso i Consigli, e ai magistrati, in quanto riduce gli adempimenti a cui gli stessi sono tenuti nell'ambito di varie procedure consiliari (si pensi alla raccolta di informazioni per i progetti organizzativi ovvero per il deposito di domande di tramutamento o per concorsi per direttivi).

Capitolo II

Tutela della riservatezza e titolarità dei dati statistici giudiziari

L'utilizzo dei dati statistici comporta il rispetto delle regole dirette a garantire la riservatezza statistica e la protezione dei dati personali. In questo capitolo verranno brevemente richiamate le norme esistenti in materia di tutela della privacy e le norme riguardanti la titolarità dei dati statistici giudiziari.

1. Il trattamento dei dati personali per scopi statistici

Il codice in materia di protezione dei dati personali - Decreto legislativo 30 giugno 2003, n. 196 - contiene le norme nazionali relative alla tutela dei dati personali, nonché le definizioni dei termini specifici adottati in materia. Si riportano in particolare le seguenti definizioni, utili ai fini della collocazione del trattamento dei dati utilizzati in ambito giudiziario per finalità statistiche:

TRATTAMENTO: qualunque operazione o complesso di operazioni, effettuati anche senza l'ausilio di strumenti elettronici, concernenti la raccolta, la registrazione, l'organizzazione, la conservazione, la consultazione, l'elaborazione, la modificazione, la selezione, l'estrazione, il raffronto, l'utilizzo, l'interconnessione, il blocco, la comunicazione, la diffusione, la cancellazione e la distruzione di dati, anche se non registrati in una banca di dati;

DATO PERSONALE: qualunque informazione relativa a persona fisica, identificata o identificabile, anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi compreso un numero di identificazione personale;

DATO SENSIBILE: qualunque informazione che rivela l'origine razziale od etnica, le convinzioni religiose, filosofiche, le opinioni politiche, l'appartenenza sindacale, relativa alla salute o alla vita sessuale;

DATI GIUDIZIARI: i dati personali idonei a rivelare provvedimenti di cui

all'articolo 3, comma 1, lettere da a) a o) e da r) a u), del d.P.R. 14 novembre 2002, n. 313, in materia di casellario giudiziale, di anagrafe delle sanzioni amministrative dipendenti da reato e dei relativi carichi pendenti, o la qualità di imputato o di indagato ai sensi degli articoli 60 e 61 del codice di procedura penale;

DATO ANONIMO (che si contrappone al dato personale): il dato che in origine, o a seguito di trattamento, non può essere associato ad un interessato identificato o identificabile;

TITOLARE: la persona fisica, la persona giuridica, la pubblica amministrazione e qualsiasi altro ente, associazione od organismo cui competono, anche unitamente ad altro titolare, le decisioni in ordine alle finalità, alle modalità del trattamento di dati personali e agli strumenti utilizzati, ivi compreso il profilo della sicurezza;

RESPONSABILE: la persona fisica, la persona giuridica, la pubblica amministrazione e qualsiasi altro ente, associazione od organismo preposti dal titolare al trattamento di dati personali.

Le regole generali relative al trattamento dei dati per finalità statistiche prevedono che i dati personali debbano essere trattati per scopi specifici, legittimi ed espliciti che devono essere resi noti agli interessati attraverso apposita informativa; devono anche essere pertinenti, completi e non eccedenti rispetto agli scopi della raccolta.

I dati statistici trattati dalla DGStat e dall'Ufficio Statistico del CSM per l'analisi dei flussi dei procedimenti civili e penali sono anonimi per quanto concerne l'identificazione dell'indagato, dell'imputato, della parte offesa e delle parti civili; non è quindi possibile identificare in alcun modo – direttamente o indirettamente - l'interessato. I dati trattati dalla DGStat sono anonimi anche per quanto riguarda l'identificazione del magistrato.

Naturalmente occorre fare riferimento al Regolamento generale sulla protezione dei dati (GDPR, dall'inglese General Data Protection Regulation) sul trattamento dei dati personali.

2. La titolarità dei dati

Il decreto n. 102 del 24 maggio 2012, stabilisce la tipologia e le modalità di estrazione, raccolta e trasmissione dei dati statistici dell'Amministrazione della giustizia.

In base all'articolo 3 i sistemi informatici del dominio giustizia

costituiscono l'infrastruttura tecnologica deputata alla raccolta, gestione, archiviazione, elaborazione e trasmissione dei dati statistici. Il sistema statistico del Ministero della giustizia assicura l'unitarietà e l'efficienza della gestione dei dati statistici. Il responsabile del sistema statistico è titolare dei dati statistici. La gestione tecnica del sistema è attribuita al responsabile per i sistemi informativi automatizzati.

In base all'art. 4 i dati e le informazioni trattati nell'ambito delle attività giurisdizionali ed amministrative del Ministero della giustizia costituiscono dati statistici a norma dell'art. 6-bis, co. 3, del d. lgs. 6 settembre 1989, n. 322 e sono utilizzati anche per le misurazioni e le valutazioni previste dall'art. 2 del d. lgs. 27 ottobre 2009, n. 150. Ciascun ufficio giudiziario o amministrativo assicura la completezza e il tempestivo aggiornamento dei dati. Il dirigente ovvero, in caso di sua mancanza, il responsabile dell'ufficio, sono responsabili della qualità dei dati e ne verificano periodicamente la correttezza e l'aggiornamento, assumendo tutte le iniziative necessarie. Il responsabile del sistema statistico verifica la qualità dei dati raccolti e, ove ravvisi anomalie, invita l'ufficio ad intraprendere ogni iniziativa necessaria a ripristinare la fedeltà, la completezza e l'aggiornamento dei dati, assegnando un termine. Gli uffici giudiziari ed amministrativi hanno l'obbligo di fornire i dati richiesti.

In base all'art. 5 la raccolta e l'estrazione dei dati è effettuata attraverso i sistemi informativi automatizzati del Ministero della giustizia che gestiscono le attività giurisdizionali ed amministrative. Il responsabile per i sistemi informativi automatizzati, d'intesa con il responsabile del sistema statistico, stabilisce le modalità operative di estrazione e raccolta dei dati statistici assicurando che i dati personali delle parti private siano resi anonimi, mediante procedure automatizzate, contestualmente alla loro raccolta. La raccolta può anche essere effettuata mediante apposite procedure informatiche o manuali stabilite dal responsabile per i sistemi informativi automatizzati, d'intesa con il responsabile del sistema statistico.

È sempre consentita la raccolta ed estrazione dei dati e delle informazioni concernenti:

- a) la classe di età delle persone fisiche;
- b) la nazionalità delle persone fisiche;
- c) la residenza, il domicilio e la dimora delle persone fisiche;
- d) la natura, tipologia, sede, stabilimento e domicilio delle persone giuridiche;
- e) la natura e la tipologia della controversia e del procedimento;
- f) la natura del reato e dell'illecito;

- g) la tipologia del reato e dell'illecito;
- h) la rubrica del reato e dell'illecito;
- i) il luogo e il periodo di consumazione del reato e dell'illecito;
- j) i dati identificativi del magistrato assegnatario del procedimento e di colui che ha emesso l'atto e il provvedimento, anche non definitorio;
- k) i dati identificativi del personale che ha trattato il procedimento, effettuando registrazioni ed annotazioni sul sistema informatico;
- l) la tipologia dell'atto e del provvedimento, anche non definitorio, adottato;
- m) la modalità di definizione del procedimento, i dati amministrativi e contabili.

I dati e le informazioni di cui alle lettere c) ed h) possono essere acquisiti solo successivamente all'emissione del provvedimento di archiviazione o all'esercizio dell'azione penale, a norma degli artt. 405, 408 e 411 del codice di procedura penale.

I dati statistici sono oggetto di elaborazione ed aggregazione per classi omogenee, definite dal responsabile del sistema statistico.

In base all'art. 7 hanno diritto di consultare le elaborazioni statistiche disponibili:

- a) i magistrati ordinari ed onorari, limitatamente ai procedimenti di loro competenza o nei quali hanno svolto le funzioni o emesso atti e provvedimenti;
- b) il dirigente amministrativo dell'ufficio giudiziario, ovvero, in caso di sua mancanza, il responsabile dell'ufficio, limitatamente ai dati relativi alle attività svolte dal personale amministrativo dell'ufficio;
- c) i magistrati con funzioni direttive e semidirettive, per i procedimenti di cui alla lettera a) ed i dati di cui alla lettera b), limitatamente ai procedimenti di competenza dell'ufficio o della sezione o del gruppo di lavoro che dirigono o coordinano;
- d) il Presidente della Corte di appello per gli uffici giudicanti del distretto;
- e) il Procuratore generale presso la corte di appello, per i procedimenti di cui alla lettera a), per gli uffici requirenti del distretto;
- f) il Procuratore nazionale antimafia, per i procedimenti di cui alla lettera a), per le direzioni distrettuali antimafia di tutti i distretti;
- g) il Presidente della Corte di cassazione per i procedimenti di cui alla lettera a), per gli uffici giudicanti di tutti i distretti;

- h) il Procuratore generale presso la Corte di cassazione, per i procedimenti di cui alla lettera a), per gli uffici requirenti di tutti i distretti;
- i) l'Ispettorato generale per i procedimenti di cui alla lettera a) ed i dati di cui alla lettera b), limitatamente ai procedimenti di competenza del magistrato, dell'ufficio o della sezione o del gruppo di lavoro oggetto di ispezione o di accertamento a fini disciplinari;
- j) il Procuratore generale presso la Corte di cassazione e la direzione generale dei magistrati, per i procedimenti di cui alla lettera a), limitatamente ai procedimenti di competenza del magistrato, dell'ufficio o della sezione o del gruppo di lavoro oggetto di ispezione o di accertamento a fini disciplinari;
- k) il Consiglio giudiziario per i procedimenti di cui alla lettera a), relativamente ai magistrati e agli uffici giudicanti e requirenti del distretto;
- l) il Consiglio direttivo della Corte di cassazione, per i procedimenti di cui alla lettera a), relativamente ai magistrati e agli uffici giudicanti e requirenti della medesima corte;
- m) il Consiglio superiore della magistratura per i procedimenti di cui alla lettera a) relativamente ai magistrati e a tutti gli uffici giudiziari.

Il responsabile del sistema statistico ha accesso a tutti i dati statistici, ad eccezione di quelli ai quali la legge vieta l'accesso per ragioni di particolare riservatezza. La consultazione è effettuata, di regola, senza accesso diretto e con modalità tali da non pregiudicare la funzionalità del sistema, tenuto conto delle risorse tecniche, organizzative e finanziarie disponibili. Nei casi in cui l'accesso diretto è autorizzato dal responsabile del sistema statistico, il responsabile per i sistemi informativi automatizzati ne stabilisce le modalità tecniche.

3. L'accesso ai dati – Dati pubblici e dati aperti

Nel capitolo 1 sono stati riportati i dati che i diversi enti che si occupano di statistica giudiziaria mettono a disposizione dei cittadini. I dati pubblicati on line possono essere definiti 'dati aperti'. Sono, infatti, dati aperti (open data) quelli che oltre a essere disponibili on line, sono forniti in un formato utilizzabile e sono liberamente consultabili e utilizzabili da tutti in forma gratuita con i relativi metadati (cfr. documentazione AGID).

Ci sono ulteriori dati pubblici che non sono disponibili on line ma possono essere richiesti agli organi competenti. Normalmente l'accesso a tali dati, sempre in modalità anonima, viene fornito quando vi è un interesse di ricerca.

La pubblica amministrazione, infatti, produce, raccoglie ed elabora tante informazioni, alcune delle quali vengono messe a disposizione della comunità nel rispetto dei principi di comunicazione, trasparenza e riuso, altre sono prodotte per finalità interne; di queste ultime, alcune possono essere opportunamente preparate e diffuse ai ricercatori che ne fanno richiesta, altre sono soggette a vincoli di tutela della privacy e non possono pertanto essere diffuse.

Ci sono anche i dati cosiddetti a conoscibilità limitata, come quelli coperti dal segreto di stato o le opere d'ingegno coperte dal diritto d'autore.

I dati aperti sono quindi un sottoinsieme dei dati pubblici.

Il SISTAN, attraverso i diversi enti che ne fanno parte, mette a disposizione dei cittadini tanti dati in formato aperto e ha organizzato diversi convegni volti ad approfondire i profili tecnici e giuridici di utilizzo degli open data anche con riferimento alla tutela della riservatezza.

Capitolo III

Misure statistiche e indicatori utilizzati in ambito giudiziario

La conoscenza delle fonti dei dati è necessaria per avviare un percorso conoscitivo nell'ambito delle statistiche giudiziarie, ma è altrettanto importante tradurre il dato in informazione attraverso la conoscenza specifica delle misure statistiche giudiziarie adottate e l'utilizzo di adeguate misure di sintesi e specifici indicatori.

La stessa valutazione della performance degli uffici giudiziari, in termini di efficienza rispetto alle risorse disponibili e di complessità delle pratiche trattate, passa sia attraverso la rappresentazione statistica dei dati, come può esserlo ad esempio la stratigrafia delle pendenze, sia attraverso l'utilizzo di indicatori, come ad esempio la durata, l'indice di ricambio, l'indice di smaltimento, la produttività pro-capite, il turn over.

Ciascun indicatore mette in evidenza uno degli aspetti necessari alla valutazione dell'andamento dei flussi ma potrebbe essere utile costruire un indicatore composito ottenuto come combinazione di più indici eventualmente pesati, ovvero ricorrere a modelli matematici di ottimizzazione lineare che permettono di fornire un indicatore di output di efficienza ottenuto da più indicatori di input senza dover associare ad essi dei pesi.

1. Le misure statistiche

Gli indicatori di performance quantitativi usualmente adottati in ambito giudiziario si basano principalmente sulle diverse misure utilizzate per rappresentare i flussi giudiziari. Le misure possono rappresentare dati di stato se riferite ad un determinato istante o dati di flusso se intervenute in un determinato arco temporale.

Le principali misure statistiche sono:

- le pendenze all'inizio di un periodo considerato (al tempo t): fascicoli ancora aperti per i quali, quindi, non è stato registrato un evento di definizione antecedente al periodo iniziale di osservazione (t);

- i procedimenti iscritti (tra il tempo t e il tempo $t+1$): i nuovi procedimenti aperti nel periodo considerato;
- le sopravvenienze (tra il tempo t e il tempo $t+1$): nel settore civile comprendono sia i procedimenti iscritti nel periodo oggetto di osservazione sia i fascicoli nuovamente aperti a seguito di riassunzione di procedimenti precedentemente sospesi, interrotti, o fascicoli per i quali è stato effettuato un mutamento rito e che quindi vengono considerati sopravvenuti con il nuovo rito processuale; nel settore penale comprendono i procedimenti iscritti nel periodo oggetto di osservazione¹, i procedimenti derivanti da eventuali stralci, i procedimenti precedentemente definiti restituiti dal GIP o dal dibattimento;
- le definizioni con sentenza (tra il tempo t e il tempo $t+1$): procedimenti per i quali è stata pubblicata o depositata la sentenza nel periodo osservato;
- gli altrimenti definiti (tra il tempo t e il tempo $t+1$): i procedimenti per i quali è stato emesso un provvedimento (o una richiesta) definitorio (decreto, ordinanza, inammissibilità, estinzione, conciliazione, sospensione, interruzione, riunione, ecc.) diverso dalla sentenza o i procedimenti civili per i quali viene effettuato un cambiamento di rito nel periodo osservato (ad esempio le trasformazioni di separazioni e divorzi consensuali in giudiziari e viceversa);
- le definizioni (tra il tempo t e il tempo $t+1$): comprendono sia i procedimenti definiti con sentenza sia i procedimenti definiti altrimenti nel periodo oggetto di osservazione;
- le pendenze finali (al tempo $t+1$): fascicoli ancora aperti alla fine del periodo considerato; le pendenze finali al tempo $t+1$ possono essere ottenute applicando la seguente formula matematica: Pendenze alla fine del periodo_(t+1) = Pendenze iniziali_(t) + Sopravvenienze_(t, t+1) – Definizioni_(t – t+1).

Normalmente si preferisce estrarre dai sistemi informativi le pendenze finali e calcolare le pendenze iniziali partendo dalla formula sopra riportata, da cui quindi risulta:

$$\text{Pendenze iniziali}_{(t)} = \text{Pendenze alla fine del periodo}_{(t+1)} + \text{Definizioni}_{(t, t+1)} - \text{Sopravvenienze}_{(t, t+1)}.$$

Le pendenze sono misure di stato, mentre le definizioni e le sopravvenienze

¹ Nell'ufficio del Gip/Gup gli iscritti sono i fascicoli *pervenuti* dalla Procura della Repubblica con richiesta definitoria.

sono misure di flusso.

Le misure statistiche sono basate, in alcuni casi, su criteri diversi di calcolo a seconda che siano riferite all'attività dell'ufficio giudiziario o a quella del magistrato.

Ad esempio, nel settore civile il procedimento si considera definito per l'ufficio quanto viene pubblicata la sentenza, mentre viene considerato concluso per il magistrato con il deposito della sentenza.

Nell'ufficio del PM il procedimento viene considerato definito per l'ufficio con decreto di citazione a giudizio solo dopo l'emissione del decreto medesimo, a seguito di indicazione da parte del Tribunale della data dell'udienza dibattimentale, e quindi con conseguente trasmissione del fascicolo al tribunale (la data definizione coincide quindi con la data trasmissione degli atti al dibattimento); diversamente per rilevare l'attività del magistrato, un procedimento è definito con citazione diretta a giudizio al momento della richiesta di citazione diretta.

Per l'ufficio Gip/Gup e per il Tribunale dibattimentale il procedimento si considera definito con sentenza con il deposito della motivazione della sentenza in cancelleria, mentre viene considerato concluso per il magistrato con l'emissione della sentenza.

2. Gli indicatori statistici

Gli indicatori (o indici) sono utilizzati per descrivere un fenomeno in modo sintetico e sono solitamente ottenuti con semplici operazioni di differenza o rapporto tra misure statistiche. Gli indicatori possono essere assoluti o relativi; i primi sono espressi nella stessa unità di misura del fenomeno oggetto di osservazione (pensiamo ad esempio alla differenza tra i procedimenti sopravvenuti nell'anno corrente rispetto a quelli sopravvenuti nell'anno precedente), mentre i secondi non dipendono dall'unità di misura del fenomeno e si ottengono rapportando tra loro due misure assolute o rapportando una misura statistica al suo valore massimo (pensiamo al rapporto tra i procedimenti sopravvenuti nell'anno corrente rispetto a quelli sopravvenuti nell'anno precedente). Tra gli indicatori relativi rientrano anche gli indici normalizzati che possono assumere i valori compresi in un determinato intervallo finito compreso tra 0 e 1 o tra -1 e +1.

Gli indicatori basati sulle differenze di due misure possono essere assoluti $I = \text{sop}_t - \text{sop}_{t-1}$ (come nella differenza tra i procedimenti sopravvenuti

nell'anno corrente rispetto a quelli sopravvenuti nell'anno precedente) o relativi $IR = (sop_t - sop_{t-1}) / sop_{t-1}$ (se la differenza assoluta viene rapportata ai sopravvenuti nell'anno precedente). Possono assumere valori positivi o negativi.

Gli indicatori ottenuti come rapporto (detti rapporti statistici) mettono a confronto due misure tra loro collegate e sono normalmente classificati come:

- Rapporti di composizione, quando si rapporta una parte del fenomeno al totale (procedimenti pendenti ultratriennali su procedimenti pendenti totali);
- Rapporti di derivazione, quando si rapporta la modalità di un fenomeno per quella corrispondente di un altro che, sul piano logico o temporale, ne costituisce l'antecedente o il parametro di riferimento (magistrati cessati su organico medio);
- Rapporto di coesistenza, ottenuto rapportando tra loro la frequenza di una modalità rispetto alla frequenza corrispondente di un'altra modalità in un fenomeno in cui la frequenza è rapportata per classi (magistrati vincitori di concorso di sesso maschile su magistrati vincitori di concorso di sesso femminile);
- Rapporto di densità ottenuto confrontando un fenomeno con la dimensione spaziale o temporale dello stesso (sopravvenuti su bacino di utenza).

I tassi sono rapporti ottenuti tra un fenomeno di flusso (definiti in un determinato periodo) e un fenomeno di stato (pendenti a fine periodo).

Infine, i numeri indici – utilizzati principalmente in ambito economico – sono dei rapporti statistici che misurano la variazione di un fenomeno nel tempo o nello spazio. La grandezza posta al denominatore viene chiamata 'base' e può essere fissa (rapporto sempre i dati rispetto a quello osservato nell'anno x) o mobile (rapporto i dati rispetto a quelli dell'anno precedente).

3. I principali indicatori di performance quantitativi e qualitativi in giustizia

La durata dei procedimenti: durata effettiva e disposition time (o durata prognostica)

Esistono fondamentalmente due indicatori rappresentativi della durata dei procedimenti giudiziari (sia di una causa civile che di un processo penale o amministrativo): la durata effettiva e la durata media prevedibile.

La *durata effettiva* misura il tempo che intercorre dalla data di iscrizione di una causa fino alla data in cui viene emesso o pubblicato il provvedimento definitivo. Il completamento di una causa o di un processo può essere considerato il momento di lettura del dispositivo da parte del giudice (laddove previsto dalla procedura) oppure la successiva data di deposito della minuta della sentenza o la data di pubblicazione da parte della cancelleria. Tutti i termini hanno una significatività informativa dal momento che può essere di interesse misurare il tempo di emissione della decisione, o il ciclo di lavoro totale fino al completamento di tutte le attività inerenti al procedimento, compresa quindi l'attività di deposito della motivazione e di pubblicazione del provvedimento da parte della cancelleria.

Durata effettiva in giorni = (Data definizione – Data iscrizione).

Attraverso il Datawarehouse del Ministero è possibile misurare la durata effettiva di tutti i singoli procedimenti civili trattati. Ciò permette di esaminare nel dettaglio le durate per materia, ufficio, e teoricamente anche per specifiche fasi processuali. Tuttavia, non per tutti i sistemi informativi in uso è possibile calcolare la durata effettiva di tutti i procedimenti definiti.

Pertanto, per stimare la durata vengono utilizzate due formule prese in prestito dalla teoria della logistica.

La prima formula - utilizzata anche in ambito internazionale (cfr. documenti Cepej) - rappresenta una stima della durata dei procedimenti pendenti ad una certa data supponendo costante nel tempo il numero di definizioni rilevato nell'ultimo anno e viene denominata *Disposition time* (DT) o *durata prognostica*. Essa misura il tempo medio prevedibile di definizione dei procedimenti mettendo a confronto tutti i procedimenti ancora aperti a una certa data (pendenze) con quelli definiti nell'ultimo anno fino a quella data (definiti):

Disposition time in anni = Pendenti finali_(t+1) / Definiti_(t, t+1).

Disposition time in giorni = Pendenti finali_(t+1) / Definiti_(t, t+1) * 365.

La durata effettiva e il disposition time sono ugualmente importanti e la lettura congiunta dei due indicatori può aiutare a comprendere l'andamento dei flussi di un ufficio: una durata effettiva elevata e un disposition time più basso non è indice di una cattiva performance ma può essere indicatore del fatto che l'ufficio si è concentrato maggiormente sullo smaltimento di fascicoli più risalenti.

Un ulteriore indice di durata media dei procedimenti è quello che si basa

sull'utilizzo della formula ISTAT, data dal rapporto tra la pendenza media in un dato anno e la semisomma dei procedimenti sopravvenuti ed esauriti.

Tale formula è la trasposizione della cosiddetta "formula di magazzino" ed esprime il tempo che intercorre tra la data di iscrizione di un procedimento e la sua definizione.

Durata con formula di magazzino in anni: $(\text{Pendenti iniziali}_{(t)} + \text{Pendenti finali}_{(t+1)}) / 2) / ((\text{Sopravvenuti}_{(t, t+1)} + \text{Definiti}_{(t, t+1)}) / 2)$.

Moltiplicando per 365 si ottiene il valore in giorni.

Pendenze e Arretrato

I provvedimenti che sono stati iscritti in un ufficio giudiziario vengono considerati "pendenti" fin quando non viene emesso o pubblicato un provvedimento decisorio (sentenza, decreto o altro). I procedimenti pendenti possono indicare pertanto il carico di lavoro cui è sottoposto un giudice, un tribunale o un sistema giudiziario in un dato momento.

Le pendenze fisiologiche (o giacenze) comprendono quei procedimenti avviati da poco tempo e che presumibilmente saranno smaltiti in un tempo ragionevole; le pendenze patologiche invece comprendono i procedimenti "a rischio Pinto" e cioè quei procedimenti che alla data di riferimento non sono stati risolti entro i termini previsti dalla legge e per i quali la parte può chiedere allo Stato un risarcimento per l'irragionevole durata del processo. I termini considerati ragionevoli dalla legge sono fissati in tre anni per i procedimenti in primo grado, due anni per i procedimenti in appello, un anno per i procedimenti in Cassazione, sei anni per una procedura fallimentare.

Pertanto, l'arretrato di un Tribunale al 31/12/2021 è dato dai procedimenti pendenti al 31/12/2021 e iscritti prima dell'1/01/2019 (o prima dell'1/01/2016 per i fallimenti); l'arretrato di una Corte d'Appello al 30/06/2021 è dato dai procedimenti pendenti al 30/06/2021 e iscritti prima dell'1/07/2019.

Lo smaltimento dell'arretrato è uno degli obiettivi che il dirigente di un ufficio giudiziario giudicante deve prevedere nel programma di gestione civile (di cui si parlerà più avanti). Il rapporto tra il numero effettivo di procedimenti ultrabiennali o ultratriennali definiti in un anno e l'obiettivo di smaltimento indicato nel programma di gestione può fornire un ulteriore indicatore per valutare la *capacità di pianificazione* dello smaltimento di un ufficio giudiziario.

Stratigrafia delle pendenze

La stratigrafia delle pendenze è una fotografia dei procedimenti pendenti ad una determinata data, dettagliata per anno di iscrizione dei procedimenti. Quindi, estrapolati tutti i fascicoli per i quali ad una determinata data non è stato emesso o pubblicato un provvedimento decisorio (sentenza, decreto o altro), gli stessi vengono raggruppati e presentati distintamente in base all'anno di iscrizione del procedimento.

L'indice di ricambio (Clearance rate)

L'indice di ricambio è calcolato come il rapporto tra le definizioni e le sopravvenienze e ha lo scopo di misurare l'efficienza di un ufficio o di un sistema giudiziario in termini di capacità di mantenere in equilibrio i flussi attraverso un numero di definizioni almeno pari alle sopravvenienze, in un dato lasso di tempo.

Indice di ricambio $_{(t, t+1)} = \text{Definiti}_{(t, t+1)} / \text{Sopravvenuti}_{(t, t+1)}$

Tale indicatore può essere espresso anche in percentuale moltiplicandolo per 100.

L'indice di ricambio di un ufficio assume valore inferiore all'unità se nell'intervallo di tempo considerato l'ufficio non riesce a smaltire (definire) tanti procedimenti quanti ne vengono iscritti: ciò implica la formazione di arretrato. Al contrario un indice di ricambio maggiore o uguale ad 1 indica una elevata efficienza dell'ufficio che riesce a definire un numero uguale o maggiore di procedimenti rispetto a quelli che sopravvivono sempre nello stesso intervallo di tempo.

Tale indice, così come altri indicatori di performance, è collegato all'organico e alla sua copertura, alla durata dei procedimenti, alla complessità e anzianità degli stessi nonché alla capacità organizzativa dell'ufficio e del sistema giudiziario nel suo complesso.

L'indice di smaltimento

L'indice di smaltimento esprime il numero di procedimenti definiti nel periodo in un ufficio giudiziario nel corso dell'anno per ogni procedimento in carico nell'ufficio; il carico è dato dalla somma dei pendenti all'inizio del periodo e dei sopravvenuti nell'anno considerato.

Tale indicatore, a differenza dell'indice di ricambio che può essere calcolato per qualsiasi periodo di tempo (mese, anno, più anni), va calcolato su un anno, sia esso solare o giudiziario².

² Per anno giudiziario si intende il periodo che va dall'1 luglio di un anno al 30 giugno dell'anno successivo.

Indice di smaltimento_(t, t+1) = Definiti_(t, t+1) / (Pendenti iniziali_(t) + Sopravvenuti_(t, t+1)).

Tale indicatore può essere espresso anche in percentuale moltiplicandolo per 100.

Come si può intendere dal nome dell'indice stesso esso esprime la capacità di smaltimento di un ufficio; più si avvicina all'unità, maggiore è la capacità di smaltimento. Si tratta, comunque, di un indicatore più difficile da interpretare non essendoci un valore che è possibile prendere come soglia oggettiva di smaltimento ottimale (come invece nel caso dell'indice di ricambio in cui un valore ≥ 1 è positivo).

La valutazione sulla capacità di smaltimento degli uffici infatti va effettuata congiuntamente all'osservazione dell'indice di variazione delle pendenze; un valore basso dello smaltimento non significa improduttività dell'ufficio se corrisponde a una positiva variazione percentuale delle pendenze; è, però, indicatore di una elevata quota di pendenze rispetto al numero di definizioni. Inoltre, tale indicatore va osservato nel tempo: l'ufficio deve porsi come obiettivo l'aumento nel tempo del valore di tale indicatore, sintomo di una maggiore capacità di definizione rispetto al carico di lavoro.

Appare evidente, per costruzione dell'indice, che se le pendenze iniziali sono molto basse o pressoché nulle l'indice di smaltimento converge all'indice di ricambio.

Indicatori di performance qualitativa

La qualità della risposta del sistema giudiziario è più difficile da valutare, non essendo spesso disponibili le informazioni necessarie. Uno degli indicatori qualitativi di misurazione dell'efficienza di un sistema giudiziario è dato dal tasso di impugnazione in appello e in Cassazione. Un elevato tasso di impugnazione misura indirettamente l'instabilità e l'incertezza della decisione, oltre a comportare un maggior numero di iscrizioni in appello e in Cassazione con il conseguenziale aumento della durata dei procedimenti.

Per calcolare il tasso di impugnazione occorre conoscere il numero di ricorsi in un determinato periodo di tempo e rapportarlo al numero di decisioni complessive, tenendo conto dei termini previsti dalla legge per la presentazione dell'impugnazione. La Corte di Cassazione pubblica sul sito, nella sezione statistiche, gli indici territoriali di ricorso civile e penale per Cassazione, calcolati secondo la formula: (Procedimenti iscritti/popolazione residente) x100.000. Tali indicatori permettono di

conoscere mediamente quanti procedimenti sono giunti in Corte da una determinata area, ogni 100.000 abitanti residenti, consentendo quindi di depurare la sopravvenienza dall'ammontare della popolazione che l'ha originata e rendendo confrontabili dati relativi a territori differenti.

Per valutare la complessità dell'attività svolta da un ufficio giudiziario si possono distinguere i flussi per materia (tipologia di reato nel penale, codice oggetto nel civile) eventualmente attribuendo dei pesi diversi ai procedimenti in funzione, oltre che della fattispecie, del rito processuale, o del numero di parti o di imputati.

La dotazione di personale e il turn over

La *dotazione di personale amministrativo* è importante in quanto il funzionamento di un ufficio giudiziario dipende anche dal corretto e puntuale svolgimento di attività amministrative e di supporto al lavoro del magistrato. La congruità del personale amministrativo sul territorio può essere valutata rapportandola *al carico di lavoro dell'ufficio* e al *personale di magistratura in organico e in servizio*.

Un altro indicatore importante è il *tasso di scopertura dei giudici*, che indica la percentuale di posti scoperti rispetto al numero di magistrati previsto in organico per ciascuna sede.

L'*indice di carico pro-capite* è dato dal rapporto tra il numero di pendenze e sopravvenienze e le unità di personale in servizio. L'*indice di esaurimento pro-capite* mette a confronto i procedimenti definiti con le risorse umane disponibili, misurando quindi la produttività pro-capite.

Il *turnover* viene normalmente definito come il tasso di rotazione del personale ovvero il flusso di persone in entrata (turnover positivo) e in uscita (turnover negativo) dall'organico di una organizzazione.

Diversi sono gli indicatori che possono essere utilizzati per stimare il tasso di rotazione del personale. I principali sono:

- tasso di turnover complessivo = $((n. \text{ entrati nel periodo} + n. \text{ usciti nel periodo}) / \text{organico ad inizio periodo}) * 100$;
- tasso di turnover negativo = $(n. \text{ usciti nel periodo} / \text{organico ad inizio periodo}) * 100$, eventualmente distinguendo il turnover volontario (usciti per fuori ruolo o trasferimento ad altra sede e/o ufficio) da quello involontario (usciti per cessazione);
- tasso di turnover positivo = $(n. \text{ entrati nel periodo} / \text{organico ad inizio periodo}) * 100$;

- tasso di compensazione = (n. entrati nel periodo / n. usciti nel periodo).

L'Ufficio statistico del CSM calcola periodicamente il tasso di *turnover quinquennale* del personale togato di magistratura negli uffici giudiziari e lo mette a disposizione sul sito del CSM nella sezione riservata alle Statistiche.

Come utilizzare gli indicatori di performance

La valutazione della performance degli uffici giudiziari può essere effettuata confrontando i diversi indicatori tra loro. Tuttavia, in questo modo è spesso difficile effettuare una valutazione unidimensionale assoluta e comparativa delle performance dei vari uffici.

L'ufficio statistico del CSM ha utilizzato un indicatore di efficienza che prende in considerazione una pluralità di indicatori, garantendo allo stesso tempo una riduzione dimensionale della misura e una adeguata interpretabilità del fenomeno misurato.

L'indicatore proposto si basa su *Modelli di Data Envelopment Analysis (DEA)* che consistono in tecniche di programmazione lineare che non formulano nessuna ipotesi sulla relazione matematica esistente tra performance e risorse in dotazione in quanto essa non è semplice da definire nel caso di uffici giudiziari e in generale di amministrazioni pubbliche. Tali modelli DEA consentono genericamente di misurare le prestazioni (produttività) delle unità produttive (uffici giudiziari) sulla base di multipli *input* e *output* e di valutare l'efficienza delle stesse unità mettendole a confronto.

Gli indici di efficienza derivanti dall'applicazione di tali modelli variano tra 0 ed 1, estremi compresi; valori più vicini all'unità indicano maggiore efficienza, valori più vicini a zero indicano minore efficienza.

I modelli DEA possono essere *input-oriented* o *output-oriented*; nel primo caso le unità efficienti sono quelle che utilizzano un numero minore di risorse per ottenere un dato output, nel secondo caso invece le unità efficienti sono quelle che con un dato numero di risorse a disposizione (input) ottengono il massimo output possibile.

Il CSM ha utilizzato il modello DEA di tipo *output-oriented* in quanto la dotazione di risorse disponibili nei vari uffici giudiziari è fissa o comunque può essere modificata solo attraverso provvedimenti decisi o approvati da enti esterni. Sono infatti il Ministero della Giustizia e il Csm che determinano l'eventuale variazione delle piante organiche degli uffici e le coperture degli organici; non può il singolo ufficio intervenire su tali variabili.

Un esempio di misure di input può essere: il tasso di copertura del

personale togato e il numero di procedimenti iscritti sull'organico di magistratura; un esempio di variabili di output possono essere: l'indice di ricambio, il numero di procedimenti definiti sull'organico, etc.

La misura di efficienza di un dato Tribunale civile ottenuta dal modello DEA orientato all'output indica quanto produce l'ufficio sulla base delle sue possibilità: per esempio un valore della misura dell'efficienza pari a 0,8 indica che l'ufficio produce all'80% delle sue possibilità e sulla base della stessa dotazione di risorse potrebbe ottenere un output maggiore del 20% rispetto a quello da lui ottenuto.

Capitolo IV

Statistica per la giustizia: come descrivere e sintetizzare i dati

La statistica descrittiva comprende un insieme di tecniche usate per descrivere le caratteristiche dei dati osservati, classificarle, rappresentarle graficamente e sintetizzarle attraverso l'utilizzo di indicatori di posizione e di variabilità.

In questo capitolo si forniscono le principali definizioni statistiche e i principali strumenti utilizzabili per descrivere e sintetizzare i dati giudiziari.

1. *Unità statistiche e caratteri*

In statistica la realtà che si intende indagare è definita con il termine *fenomeno*, intendendo tutto ciò che è suscettibile di osservazione diretta o indiretta; ad esempio “il reddito totale di un certo Paese” oppure “il numero di udienze” tenute da ciascun giudice (fenomeno individuale) o dall'intero ufficio (fenomeno collettivo).

Le *unità statistiche* sono le unità (individui generici, giudici, etc.) oggetto di studio sulle quali vengono rilevate le caratteristiche legate al fenomeno. Ad esempio, le unità statistiche possono essere i giudici del Tribunale d'Italia su cui si rileva il numero di udienze tenute nell'anno.

L'insieme delle unità statistiche costituisce la popolazione di interesse chiamata anche *collettivo statistico* oppure *universo statistico*; essa può essere “finita” se costituita da un numero finito di unità oppure “infinita” se comprende un numero innumerabile o non precisabile di unità. Nel nostro esempio di prima la popolazione di interesse è data dall'insieme dei giudici dei Tribunali di Italia.

Ogni unità statistica è portatrice di più caratteristiche, chiamate *caratteri* o *attributi*; il modo in cui un carattere si manifesta in una unità statistica è detto *modalità*.

I caratteri possono essere raggruppati in due grandi gruppi: Caratteri qualitativi e Caratteri quantitativi. I caratteri qualitativi si dividono a loro volta nei seguenti sottogruppi:

- Caratteri qualitativi sconnessi: caratteri per le cui modalità non è possibile stabilire un ordinamento di alcun tipo (ad eccezione di quello alfabetico), si può solo dire se sono diverse o uguali (genere, luogo di nascita, macromaterie attribuite a una sezione civile, uffici di Tribunale, etc.).
- Caratteri qualitativi ordinati: caratteri per le cui modalità è possibile stabilire un ordinamento come ad esempio il titolo di studio, il grado di soddisfazione, la funzione ricoperta all'interno di un ufficio giudiziario (i.e. magistrato in tirocinio, giudice, presidente di sezione, presidente), etc.

I caratteri quantitativi si dividono in:

- Caratteri discreti: caratteri che assumono un numero finito o infinito ma numerabile di modalità o valori tra cui è possibile effettuare operazioni di differenza/somma o di prodotto/rapporto. Essi sono essenzialmente numeri che risultano da un conteggio: numero di telefonate in un giorno, *numero di definizioni per ufficio giudiziario*, etc.
- Caratteri continui: caratteri che assumono un numero infinito di valori (numeri reali) all'interno di un determinato intervallo; ad esempio altezza, peso, tempo di attesa, *durata di un procedimento espressa in tempo esatto* (giorni, minuti e secondi).

2. Distribuzione statistica e frequenza

Le informazioni concernenti un determinato fenomeno possono essere raccolte attraverso due tipi di rilevazioni: totale o campionaria; nel primo caso la rilevazione è estesa a tutte le unità della popolazione oggetto di studio, nel secondo caso soltanto ad una parte di essa.

- ✓ Nel caso di rilevazione totale si usa la statistica descrittiva con lo scopo di operare una sintesi di tutte le manifestazioni individuali che caratterizzano il fenomeno considerato;
- ✓ Nel caso di rilevazione campionaria invece la statistica inferenziale permette appunto di inferire cioè di generalizzare i risultati ottenuti sul campione a tutte le unità della popolazione da cui il campione proviene.

Una distribuzione statistica consiste nella distribuzione delle caratteristiche rilevate sulle unità della popolazione. Ad esempio, se X è

un carattere qualitativo o quantitativo su cui è possibile investigare, esso assumerà in corrispondenza delle unità statistiche $i=1, \dots, N$ le modalità o valori $x_1, \dots, x_N$.

Assumiamo che le unità statistiche siano le sedi di Tribunale e il carattere qualitativo sia la dimensione: la Tab. 21 in appendice riporta la distribuzione statistica della dimensione degli uffici per le sedi di Tribunale e Procura della Repubblica a dicembre 2022.

Accade spesso che alcune modalità possano ricorrere più di una volta. Se si indica con n_i il numero di volte secondo cui la modalità x_i appare nel collettivo considerato, la distribuzione si può riscrivere sapendo che ad ogni x_i corrisponde un numero n_i , tale numero è detto *frequenza assoluta* e l'insieme delle coppie (x_i, n_i) $i=1, \dots, i, \dots, N$ è detta distribuzione di frequenze secondo il carattere X (in questo caso la dimensione). Si ottiene pertanto la seguente Tab. 1:

Tab. 1. Distribuzione di frequenze assolute del collettivo dei Tribunali secondo la dimensione

Dimensione (x_i)	Numero Tribunali (n_i)
Metropolitani	3
Grandi	3
Medio grandi	20
Medio piccoli	66
Piccoli	48

Per un carattere qualitativo o quantitativo discreto, si definisce frequenza relativa corrispondente alla modalità x_i il rapporto n_i/N cioè il rapporto tra la frequenza assoluta corrispondente alla modalità x_i e il totale delle unità statistiche rilevate N dato dalla somma di tutte le n_i , $i=1, \dots, m$; m =numero totale di modalità (6 nella Tab. 1): $N = \sum_{i=1}^m n_i$.

La frequenza relativa si indica con $f_i = n_i/N$ e la somma delle f_i deve essere pari a 1. Pertanto, in formule si ha $\sum_{i=1}^m \frac{n_i}{N} = \sum_{i=1}^m f_i = 1$. Moltiplicando le frequenze relative per 100 si ottengono le frequenze relative percentuali; in questo caso la somma dovrà essere pari a 100, $\sum_{i=1}^m f_i \% = 100$.

La tabella seguente mostra la distribuzione delle frequenze relative percentuali inerenti alla tabella precedente.

Tab. 2. Distribuzione di frequenze relative percentuali dei Tribunali secondo la dimensione di organico.

Dimensione	Numero relativo percentuale di sedi ($f_i\%$)
Metropolitani	2,14%
Grandi	2,14%
Medio grandi	14,28%
Medio piccoli	47,14%
Piccoli	34,28%
Totale	100 %

La Tab. 2 sovrastante mostra che il 47,14% delle sedi di Tribunale è di dimensione medio piccola.

3. Le rappresentazioni grafiche di variabili qualitative e quantitative discrete

Al fine di descrivere e sintetizzare in modo semplice e immediato l'informazione relativa a un dato fenomeno possono essere utilizzati degli strumenti grafici che si differenziano tra loro a seconda della tipologia di variabile/i che sintetizza/no il fenomeno considerato.

Nel caso di variabili qualitative (sconnesse o ordinate) i grafici più utilizzati per una rappresentazione visiva sono i grafici a barre (o a rettangoli distanziati) e i grafici circolari anche detti a torta.

Nel caso di grafici a barre la frequenza di ogni modalità è rappresentata dall'area di un rettangolo di base unitaria; i rettangoli sono tutti staccati tra loro e alla base di ciascuno di essi viene riportata l'etichetta della modalità a cui si riferisce la frequenza osservata. L'ordine delle barre può essere casuale nel caso di variabili qualitative connesse, può invece avere una direzione nel caso di variabili qualitative ordinate.

Di seguito viene rappresentato il grafico a barre relativo alla Tab. 1.

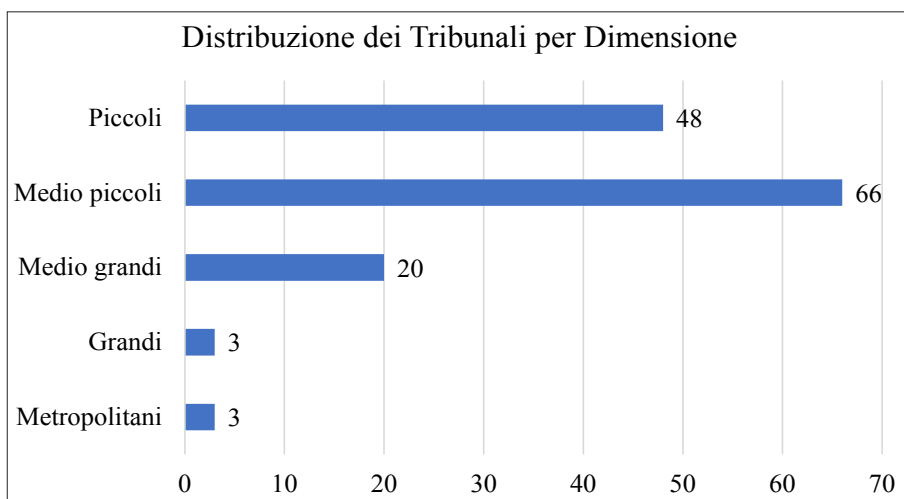


Grafico 1

Nel caso di grafici a torta, la frequenza totale N è rappresentata dalla superficie di un cerchio. La frequenza della i -esima modalità n_i è invece rappresentata dall'area di un settore circolare. Per determinare le ampiezze degli angoli al centro, per ciascuna delle m modalità osservate, è sufficiente moltiplicare le frequenze di ciascuna modalità per la costante $360^\circ/N$. Viene di seguito riportato il grafico a torta relativo ai dati in valore assoluto e percentuale della Tab. 1 e della Tab. 2.

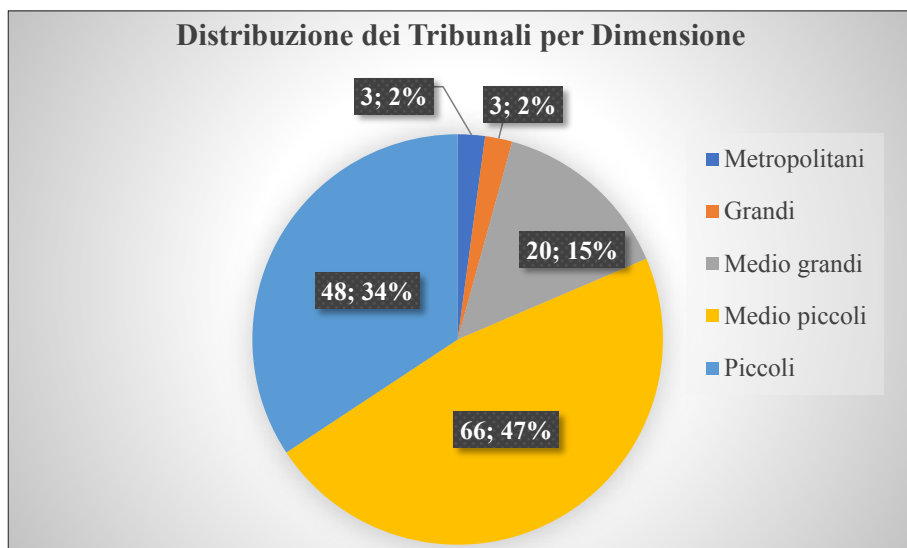
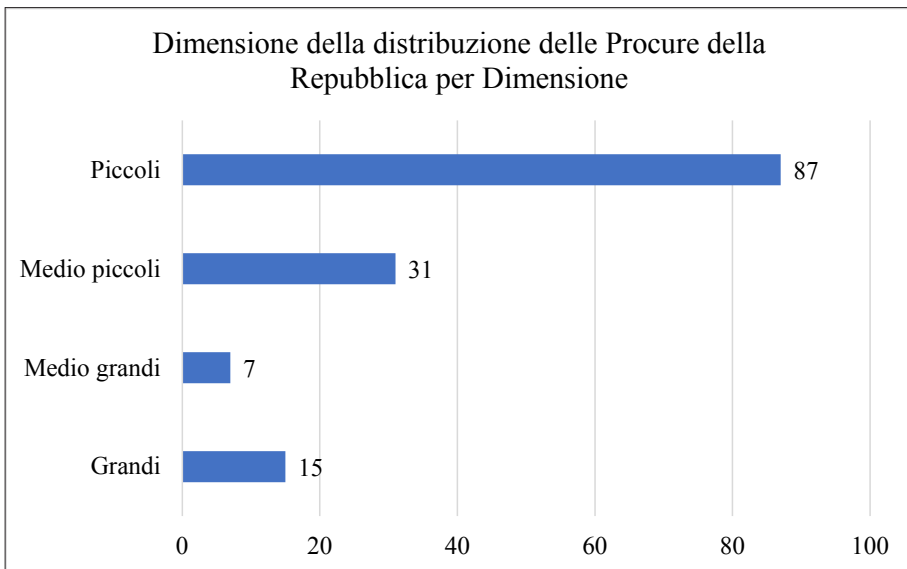
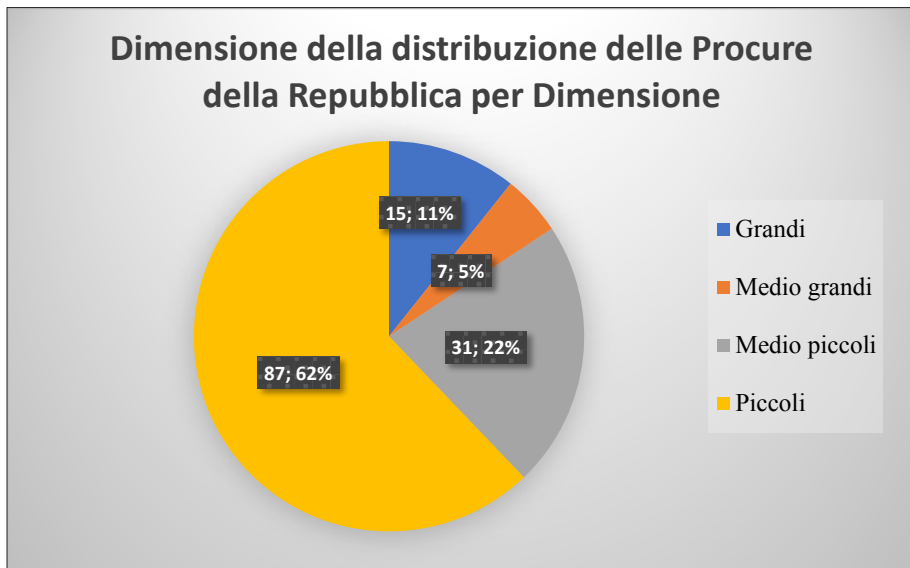


Grafico 2

Di seguito si riportano i grafici a barre e a torta relativi alla distribuzione assoluta e percentuale delle Procure della Repubblica per Dimensione.

*Grafico 3**Grafico 4*

Nel caso di variabili quantitative discrete, come ad esempio il numero di udienze civili tenute in un semestre da un campione di magistrati addetti al settore civile di alcuni Tribunali italiani, la rappresentazione grafica più opportuna è la rappresentazione per punti.

Consideriamo ad esempio la tabella seguente che riporta il numero fittizio di magistrati addetti al civile di alcuni Tribunali per numero di udienze espletate (anch'esse fittizie) e il Grafico 5 che ne mostra la rappresentazione grafica.

Tab. 3. Distribuzione del numero di magistrati addetti al civile per numero di udienze civili

Numero di udienze civili (x_i)	Numero di magistrati addetti al civile (n_i)
5	8
7	6
10	5
20	6
25	3
30	2

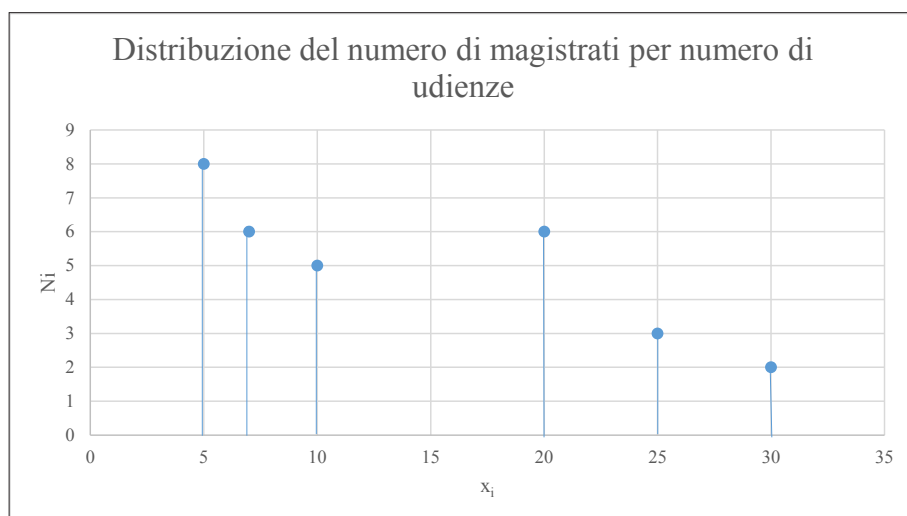


Grafico 5

4. Le rappresentazioni grafiche di variabili quantitative continue

Nel caso di un carattere quantitativo continuo invece i grafici maggiormente utilizzati sono: l'istogramma di frequenze, il boxplot. Nel caso di due variabili quantitative continue è molto utilizzato il grafico a dispersione che può mettere in evidenza l'eventuale direzione della relazione esistente tra le due variabili. Descriviamo di seguito come si costruiscono i grafici suddetti.

L'*istogramma* permette di rappresentare graficamente la forma di una distribuzione di frequenza in classi, di una variabile quantitativa continua X .

Un esempio di distribuzione quantitativa continua nel settore della giustizia è quello rappresentato nella tabella seguente in cui le unità i sono le Corti di Appello e la variabile continua è data dalle pendenze finali pro-capite (X). La tabella sottostante mostra la distribuzione unitaria dei pendenti finali pro-capite (per magistrato in organico) per ufficio di Corte di Appello; le Corti di Appello sono 29 considerando anche le sedi distaccate di Bolzano, Sassari e Taranto, pertanto l'indice i che le identifica va da 1 a 29. I pendenti finali costituiscono la variabile aleatoria X che assume valori $x_1, \dots, x_{29}$ (es. $x_1=320$ in corrispondenza di $i=1$ =Ancona).

Tab. 4. Pendenze finali nelle Corti di Appello al 31.12.2021

Corte di Appello	Pendenti finali pro-capite (x_i)
Ancona	320
Bari	346
Bologna	402
Bolzano	61
Brescia	181
Cagliari	167
Caltanissetta	165
Campobasso	157
Catania	444
Catanzaro	343
Firenze	356
Genova	238
L'Aquila	262
Lecce	366

Corte di Appello	Pendenti finali pro-capite (x_i)
Messina	208
Milano	112
Napoli	538
Palermo	318
Perugia	213
Potenza	303
Reggio Calabria	351
Roma	512
Salerno	189
Sassari	273
Taranto	289
Torino	196
Trento	42
Trieste	188
Venezia	316

Nostre elaborazioni su Dati Ministero della Giustizia

A partire dai dati riportati in Tab. 4 è possibile ricavare una distribuzione di frequenze secondo il carattere quantitativo continuo “Pendenti finali pro-capite” (X) espresso in intervalli, come mostrato dalla tabella seguente.

Tab. 5. Distribuzione del numero di Corti di Appello per classi¹ di pendenze finali pro-capite

Pendenti finali pro-capite(x_i)	Numero di Corti di Appello (n_i)
[0-100)	2
[100-200)	8
[200-300)	6
[300-400)	9
[400-500)	2
[500-600)	2

Nostre elaborazioni su Dati Ministero della Giustizia

Nel caso di caratteri quantitativi continui raggruppati in classi, è possibile

¹ In una classe di valori le parentesi tonde indicano l'esclusione dei valori estremi, le parentesi quadre ne indicano l'inclusione.

utilizzare l'Istogramma con basi uguali o diverse a seconda dell'omogeneità o meno dell'ampiezza delle classi.

Sull'asse delle ascisse dell'istogramma si dispongono in ordine crescente le classi osservate della variabile, ciascuna con la sua ampiezza a_i (nell'esempio della Tab. 5 pari a 100) che costituisce la base sopra la quale si costruisce un rettangolo, la cui altezza d_i è la densità di frequenza della classe stessa $d_i = n_i/a_i$. La densità di frequenza della classe i esprime quanta parte della frequenza della classe compete in media ad ogni segmento di ampiezza unitaria in essa contenuto.

Se si rappresentano le frequenze assolute, l'area totale dell'istogramma rappresenta l'intero ammontare di osservazioni N ; se si rappresentano le frequenze relative f_i l'area totale dell'istogramma sarà uguale a 1. L'area totale dell'istogramma è ripartita tra i singoli rettangoli proporzionalmente al numero di casi entro ciascuna classe; l'area di ciascun rettangolo dell'istogramma è pari alla frequenza assoluta o relativa della classe corrispondente e può essere pertanto espressa come prodotto tra base per altezza di ciascun rettangolino $n_i = d_i * a_i$.

Le classi sono intervalli contigui, disgiunti e ordinati e i rettangoli dell'istogramma devono essere pertanto uniti.

Nell'esempio di Tab. 5 le classi sono gruppi di pendenze finali pro-capite e l'ampiezza degli intervalli è uniforme e pari a 100; in questo caso l'istogramma può essere rappresentato direttamente con rettangoli aventi per base l'intervallo di valori della classe e come altezza la frequenza corrispondente in quanto quest'ultima è proporzionale alla densità in modo uguale per ogni classe (Grafico 6).

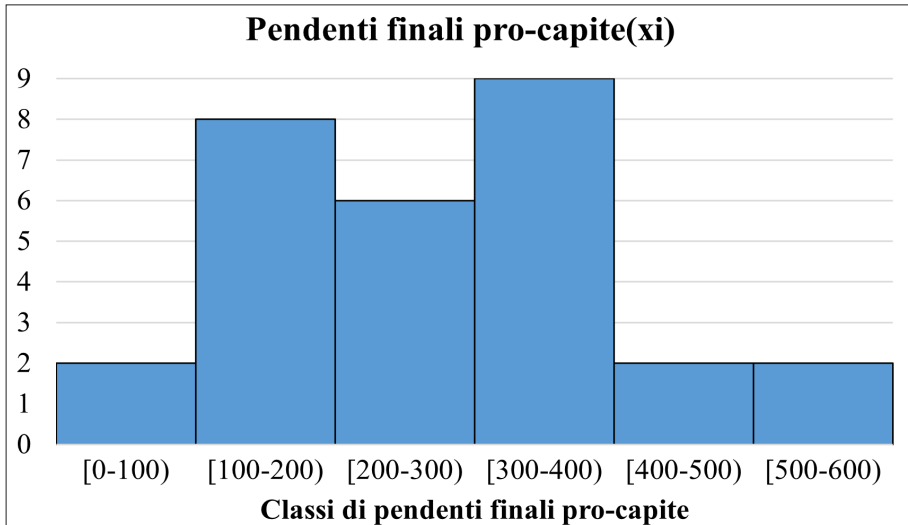


Grafico 6 - Nostre elaborazioni su dati Ministero della Giustizia

Tab. 6. Distribuzione del numero di Corti di Appello per classi di pendenze finali pro-capite

Pendenti finali pro-capite(x_i)	Ampiezza della classe a_i	Numero di Corti di Appello (n_i)	Densità (d_i)= n_i/a_i
[0-50)	50	1	0,020
[50-150)	100	2	0,020
[150-180)	30	3	0,100
[180-300)	120	10	0,083
[300-350)	50	6	0,120
[350-390)	40	3	0,075
[390-550)	160	4	0,025

Nostre elaborazioni su Dati Ministero della Giustizia

Nella Tab. 6 le classi relative alle pendenze finali pro-capite nelle Corti d'Appello hanno una ampiezza diversa; in questo caso l'altezza dell'istogramma è pari alla densità.

Il boxplot o grafico a baffi invece permette di rappresentare indici di posizione relativi a un certo carattere nonché i valori anomali detti anche outlier, cioè valori che si comportano in modo diverso rispetto al resto della distribuzione. Il boxplot verrà spiegato in modo più dettagliato nel seguito

del presente contributo, dopo aver descritto gli indicatori statistici necessari alla sua costruzione.

Il *grafico a dispersione* si può usare come strumento di controllo di qualità, di previsione, di analisi di marketing o per altri problemi di statistica e di statistica giudiziaria.

Il grafico a dispersione mostra la relazione eventualmente esistente tra due variabili quantitative misurate sulle stesse unità statistiche; la tabella seguente riporta un esempio di due misure rappresentate dai flussi giudiziari e in particolare dalle iscrizioni per organico (iscrizioni pro-capite, X) e dalle pendenze finali pro-capite (Y) in corrispondenza delle stesse unità statistiche in questo caso rappresentate dagli uffici giudicanti di secondo grado.

Tab. 7. Definizioni e sopravvenienze pro-capite nelle Corti di Appello

Corti di Appello	Pendenti finali pro-capite 2021 (y_i)	Iscritti pro-capite 2021 (x_i)
Ancona	320	173
Bari	346	154
Bologna	402	185
Bolzano	61	56
Brescia	181	106
Cagliari	167	100
Caltanissetta	165	103
Campobasso	157	119
Catania	444	173
Catanzaro	343	170
Firenze	356	141
Genova	238	115
L'Aquila	262	158
Lecce	366	145
Messina	208	178
Milano	112	108
Napoli	538	141
Palermo	318	173
Perugia	213	155
Potenza	303	139

Corti di Appello	Pendenti finali pro-capite 2021 (y_i)	Iscritti pro-capite 2021 (x_i)
Reggio Calabria	351	104
Roma	512	147
Salerno	189	118
Sassari	273	138
Taranto	289	148
Torino	196	120
Trento	42	55
Trieste	188	116
Venezia	316	131

Nostre elaborazioni su Dati Ministero della Giustizia

Nel grafico a dispersione ogni unità è rappresentata da un punto di coordinate (x, y) , dove i valori delle ascisse rappresentano i valori x di una variabile X , che in questo caso corrisponde alla variabile continua “iscritti pro-capite”, e quelli delle ordinate rappresentano i valori dell’altra variabile y , che nel nostro esempio corrisponde alla variabile continua “pendenti finali pro-capite”.

Il grafico a dispersione perciò visualizza in uno spazio cartesiano² il modo in cui agiscono due o più variabili: l’insieme di punti nel suo complesso mostra il comportamento del fenomeno studiato. Il grafico serve a mostrare le possibili relazioni esistenti tra variabili.

² Uno spazio cartesiano è un sistema di riferimento formato da rette ortogonali che si intersecano in un punto chiamato origine; se lo spazio cartesiano è costituito da due rette che si intersecano (due dimensioni) si chiama piano cartesiano e si divide per convenzione in 4 regioni denominate quadranti e identificate con numeri romani in senso antiorario partendo dalla regione (quadrante) con ascissa e ordinata positiva. Nel Grafico 8 è rappresentato solo il primo quadrante.

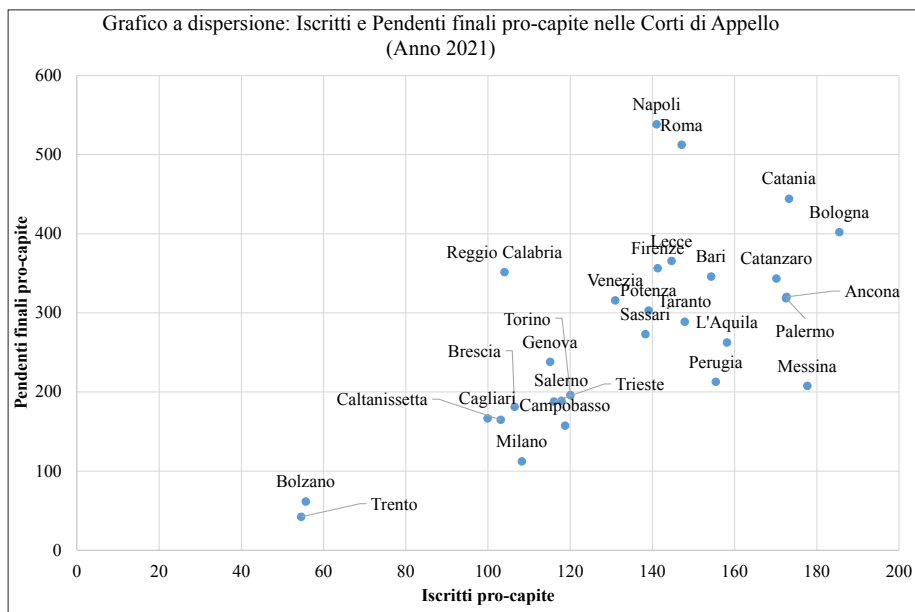


Grafico 7 - Nostre elaborazioni su dati del Ministero della Giustizia

La relazione tra due variabili si può verificare, infatti, anche solo guardando il relativo grafico a dispersione; in particolare se l'insieme di punti (sciame) va verso l'alto la relazione tra le due variabili sarà positiva o meglio i due caratteri saranno positivamente correlati: all'aumentare dell'uno aumenta l'altro. Nell'esempio raffigurato si può facilmente notare che all'aumentare delle iscrizioni pro-capite aumentano le pendenze finali e viceversa.

Se lo sciame dei punti è diretto verso il basso la correlazione tra le due variabili è negativa cioè all'aumentare dell'una diminuisce l'altra, se invece la distribuzione dell'insieme dei punti sul grafico cartesiano non segue un criterio non esiste una relazione tra i due caratteri considerati.

Il grafico seguente mostra perfettamente queste tre situazioni:

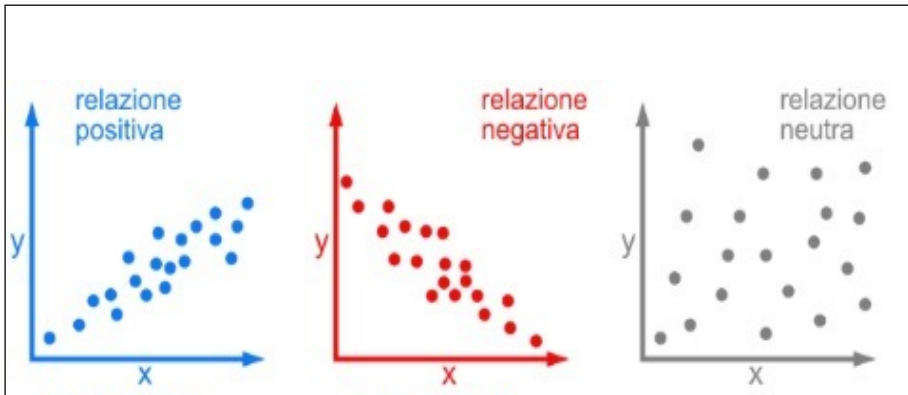


Grafico 8

5. Indici di posizione

Quando si analizzano variabili quantitative è utile accompagnare le rappresentazioni grafiche con specifici indici e misure adeguati a sintetizzare il fenomeno.

Gli indici di posizione sintetizzano la posizione di una distribuzione di frequenza mediante un valore reale rappresentativo della globalità del fenomeno e tale da riassumere gli aspetti ritenuti più importanti. Gli indici di posizione più utilizzati sono la Moda, la Media e la Mediana.

La Moda

La moda di una distribuzione di frequenza X è la modalità più frequente. Nell'esempio di cui alla Tab. 1 sulla distribuzione di frequenze assolute del collettivo dei Tribunali secondo la dimensione, la caratteristica che ha la più elevata frequenza è data dai Tribunali 'medio piccoli' (66 su 140).

La Media

La media mira ad esprimere la posizione del centro ideale sul quale tendono a gravitare le modalità della distribuzione; è chiamato infatti anche indice di tendenza centrale.

Data una distribuzione di frequenze $\{(x_1, n_1), \dots, (x_m, n_m)\}$ secondo un carattere quantitativo X (in cui per ogni modalità/valore del carattere

è indicata la sua frequenza, l'ultimo valore è l'm-esimo), la sua media aritmetica è data da

$$\bar{X} = \frac{1}{N}(x_1n_1 + \dots + x_in_i + \dots + x_mn_m) = \frac{1}{N}\sum_{i=1}^m x_in_i = \sum_{i=1}^m x f_i \quad [1]$$

in cui $f_i = n_i/N$.

Nel caso di distribuzione unitaria (come quella presentata in Tab. 4, in cui N rappresenta il numero totale di osservazioni) la media aritmetica ovviamente è pari a

$$\bar{X} = \frac{1}{N}(x_1 + \dots + x_i + \dots + x_N) = \frac{1}{N}\sum_{i=1}^N x_i, \quad [2]$$

pertanto, vale la condizione per cui $\bar{X}N = \sum_{i=1}^N x_i$.

Sulla base, quindi, dei dati in Tab. 4 relativi al carattere quantitativo X= "pendenze pro-capite" per le Corti di Appello, la media in base alla formula [2] è pari a $\bar{X} = (320+346+\dots+316) / 29 = 271$ dove 29 è il numero di Corti di Appello.

Se si considera un campione di magistrati di numerosità N=15 con numero di udienze mensili per magistrato pari rispettivamente a: 1, 2, 4, 7, 7, 7, 10, 10, 10, 10, 10, 13, 14, 14, 15. La media totale delle udienze è data semplicemente dalla somma delle udienze mensili diviso il numero totale di giudici: $(1+2+3+\dots+15) / 15 = 134/15 = 8,93$.

La stessa distribuzione di udienze mensili per magistrato può essere formalizzata secondo il formato riportato nella tabella seguente.

Tab. 8.

Num. Udienze mensili (x_i)	Num. Magistrati (n_i)	Num. Tot. Udienze ()
1	1	1
2	1	2
4	1	4
7	3	21
10	5	50
13	1	13
14	2	28
15	1	15
Totale	15	134

Per calcolare la media in questo caso è necessario prima calcolare il numero di udienze mensili totali moltiplicando il numero di udienze mensili per numero di magistrati corrispondenti e dividendo la somma di tali prodotti, pari a 134, per il numero totale di magistrati pari a 15, ottenendo 8,93 udienze mensili in media totali.

La media calcolata in questo modo è chiamata media ponderata, dove i pesi sono pari alle frequenze assolute (o relative) delle unità a cui corrispondono i singoli valori della variabile (vedi [1]).

Quando i numeri di una serie non sono indipendenti l'uno dall'altro oppure si caratterizzano per grandi fluttuazioni (ad esempio tassi di crescita elevati seguiti da tassi di crescita bassi nel periodo successivo) può essere calcolata la *Media geometrica* (Mg), molto in uso soprattutto nel mondo finanziario.

$$Mg = \bar{X}_g = (\prod_{i=1}^N x_i)^{1/N}$$

Nell'esempio sopra la media geometrica è pari a 3,24.

Un altro tipo di media che dà uguale peso a ciascun punto di dati è la Media armonica (Ma) utilizzata spesso per trovare relazioni moltiplicative o divisorie tra le frazioni senza preoccuparsi dei denominatori comuni (durata media di viaggio data una durata di diversi viaggi). La Media armonica è definita come il reciproco della media aritmetica dei reciproci:

$$Ma = \bar{X}_a = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}}$$

Nell'esempio in Tabella 8 la media armonica è pari a 5,06.

La Mediana, percentili e quartili

Consideriamo la distribuzione di un carattere quantitativo come ad esempio quello relativo alle pendenze pro-capite per ufficio della Tab. 4. Ordiniamo i valori del carattere quantitativo in ordine crescente o decrescente; la mediana è il valore (delle pendenze) che bipartisce il collettivo (gli uffici di Corte di Appello) ossia che lascia alla sua destra e alla sua sinistra un numero uguale di unità e cioè il 50%. Se il numero di unità è dispari la mediana è la modalità che occupa la posizione centrale $(N+1)/2$, mentre se N è pari per convenzione si assume come mediana la semisomma delle modalità che occupano i due posti centrali $(N/2 + N/2+1)/2$.

Nell'esempio, poiché le Corti sono in numero dispari e pari a 29, dopo aver ordinato le Corti in base ai valori delle pendenze pro-capite dal più piccolo al più grande (vedi tabella seguente), è facile vedere che la modalità mediana è la Corte di Appello corrispondente alla 15° posizione $(29+1)/2$. Sassari è pertanto la Corte di Appello mediana e il valore mediano delle pendenze è 273.

Tab. 9

Corte di Appello	Pendenti pro-capite (x.)
Trento	42
Bolzano	61
Milano	112
Campobasso	157
Caltanissetta	165
Cagliari	167
Brescia	181
Trieste	188
Salerno	189
Torino	196
Messina	208
Perugia	213
Genova	238
L'Aquila	262
Sassari	273
Taranto	289
Potenza	303
Venezia	316
Palermo	318
Ancona	320
Catanzaro	343
Bari	346
Reggio Calabria	351
Firenze	356
Lecce	366
Bologna	402
Catania	444
Roma	512
Napoli	538

In questo caso la distribuzione si può definire abbastanza simmetrica in quanto la mediana non dista molto dal valor medio che è pari a 271; in generale la mediana è un indice di posizione molto più robusto della media in quanto a differenza di essa non risente dell'influenza dei valori anomali.

La mediana corrisponde al secondo quartile o al 50° percentile.

I percentili dividono la distribuzione in 100 parti, ad esempio il 1° percentile lascia alla sua sinistra un centesimo degli elementi della distribuzione ossia l'1%; il 10° percentile lascia alla sua sinistra il 10% degli elementi della distribuzione, etc. I quartili dividono la distribuzione in 4 parti: il 1° quartile lascia pertanto alla sua sinistra il 25% della distribuzione e corrisponde pertanto al 25° percentile, il 3° quartile corrisponde al 75° percentile. Di seguito viene riportato il calcolo dei quartili di una distribuzione unitaria (in cui cioè ogni ufficio rappresenta una unità) facendo riferimento all'esempio sopra riportato.

Il primo quartile (Q1) è il valore delle pendenze tale che il 25% (1/4) della distribuzione assume valori minori o uguali ad esso. Pertanto, dopo aver ordinato la distribuzione, il primo quartile è circa pari al valore che corrisponde alla posizione $N \cdot 0,25$. Nell'esempio proposto N è pari a 29 e pertanto la posizione del primo quartile è pari a 7,25; essendo esso superiore a 7 si considera pertanto la posizione 8. Il valore delle pendenze che lascia dietro di sé il 25% della distribuzione è 188 e corrisponde alla Corte di Appello di Trieste.

Il terzo quartile (Q₃) è il valore delle pendenze tale che il 75% (3/4) della distribuzione assume valori minori o uguali ad esso. Pertanto, dopo aver ordinato la distribuzione, il terzo quartile è circa pari al valore che corrisponde alla posizione $N \cdot 0,75$. Nell'esempio proposto N è pari a 29 e pertanto la posizione del primo quartile è pari a 21,75, essendo superiore alla 21° si considera pertanto la posizione 22. Il valore delle pendenze che lascia dietro di sé il 75% della distribuzione è 346 e corrisponde alla Corte di Appello di Bari.

Nella tabella sovrastante il valore mediano è segnato in rosso mentre quello dei quartili in blu.

Avendo a disposizione i quartili si può calcolare il *range interquartile* che quantifica l'estensione del 50% della distribuzione del carattere che si trova attorno alla mediana: il 50% delle unità statistiche che presentano una modalità prossima a quella centrale.

Più ampio è il range interquartile, maggiore è la *dispersione*, cioè la distanza, la diversità, delle unità statistiche rispetto alla mediana.

Il range interquartile è espresso nella stessa unità di misura del carattere e non è influenzato dall'eventuale presenza di valori estremi o anomali (cioè distanti, diversi dagli altri) assunti dal carattere nel collettivo in esame.

Nell'esempio considerato il range interquartile è pari a $346-188=158$ pendenze pro-capite.

Un altro modo per calcolare i quartili e perciò anche la mediana è quello che si basa sulle frequenze relative cumulate³ (F_i). Consideriamo sempre la tabella delle pendenze pro-capite per le Corti di Appello in cui le Corti sono le unità ordinate secondo l'ordinamento crescente dei pendenti finali. Poiché stiamo considerando una distribuzione unitaria (in cui, come già detto, ogni ufficio rappresenta una unità), le frequenze assolute n_i associate a ogni Corte sono pari a 1; a partire da n_i si possono calcolare le frequenze cumulate N_i sommando ogni n_i corrispondente a un certo valore delle pendenze ai valori precedenti delle pendenze. A questo punto è possibile calcolare le frequenze cumulate relative F_i rapportando le frequenze cumulate N_i in corrispondenza di ogni unità al totale delle Frequenze cumulate che corrisponde a 29, il totale delle Corti di Appello. È possibile poi trovare il valore mediano guardando alla colonna F_i ; poiché la mediana corrisponde al valore che divide la distribuzione in due parti uguali e che quindi lascia il 50% della distribuzione sopra e il 50% sotto, andiamo a guardare quando F_i è pari al 50% o il primo valore di F_i più grande del 50% (perché significa che contiene il 50%). Tale valore F_i nell'esempio considerato è pari a 52% e il valore delle pendenze pro-capite che corrisponde a tale valore è pari a 273 e rappresenta il valore mediano. La stessa cosa può essere fatta per il primo e il terzo quartile guardando al valore delle pendenze che corrisponde al 25% (o al primo valore di F_i più grande di 25%) e al 75% (o al primo valore di F_i più grande di 75%) rispettivamente: il valore delle pendenze pro-capite corrispondente al primo quartile è pari a 188 mentre quello relativo al terzo quartile è 346.

La tabella seguente riporta le pendenze civili ordinate, la distribuzione di frequenze, le frequenze cumulate assolute e relative; i valori dei quartili sono in grassetto.

³ La frequenza cumulata (assoluta), associata ad una modalità o a una classe di modalità, è pari alla somma della sua frequenza assoluta e di quelle delle modalità che la precedono.

Tab. 9a

Corte di Appello	Pendenti pro-capite (x_i)	n_i	Frequenze cumulate (N_i)	F_i
Trento	42	1	1	3%
Bolzano	61	1	2	7%
Milano	112	1	3	10%
Campobasso	157	1	4	14%
Caltanissetta	165	1	5	17%
Cagliari	167	1	6	21%
Brescia	181	1	7	24%
Trieste	188	1	8	28%
Salerno	189	1	9	31%
Torino	196	1	10	34%
Messina	208	1	11	38%
Perugia	213	1	12	41%
Genova	238	1	13	45%
L'Aquila	262	1	14	48%
Sassari	273	1	15	52%
Taranto	289	1	16	55%
Potenza	303	1	17	59%
Venezia	316	1	18	62%
Palermo	318	1	19	66%
Ancona	320	1	20	69%
Catanzaro	343	1	21	72%
Bari	346	1	22	76%
Reggio Calabria	351	1	23	79%
Firenze	356	1	24	83%
Lecce	366	1	25	86%
Bologna	402	1	26	90%
Catania	444	1	27	93%
Roma	512	1	28	97%
Napoli	538	1	29	100%

Il Boxplot

Il boxplot è il migliore metodo di rappresentazione di una distribuzione statistica attraverso i suoi indici di posizione.

Esso consiste in una “scatola” la cui linea interna rappresenta la mediana della distribuzione, mentre le linee estreme della scatola rappresentano il primo e il terzo quartile all’interno delle quali si trova il 50% delle osservazioni. All’aumentare della distanza tra terzo e primo quartile aumenta la dispersione delle osservazioni intorno alla mediana; tra le due linee fuori dalla scatola sono collocati la maggior parte degli uffici (circa l’80%). I punti esterni alle linee rappresentano i valori anomali della distribuzione.

Il grafico seguente è esplicativo di come è costruito un boxplot.

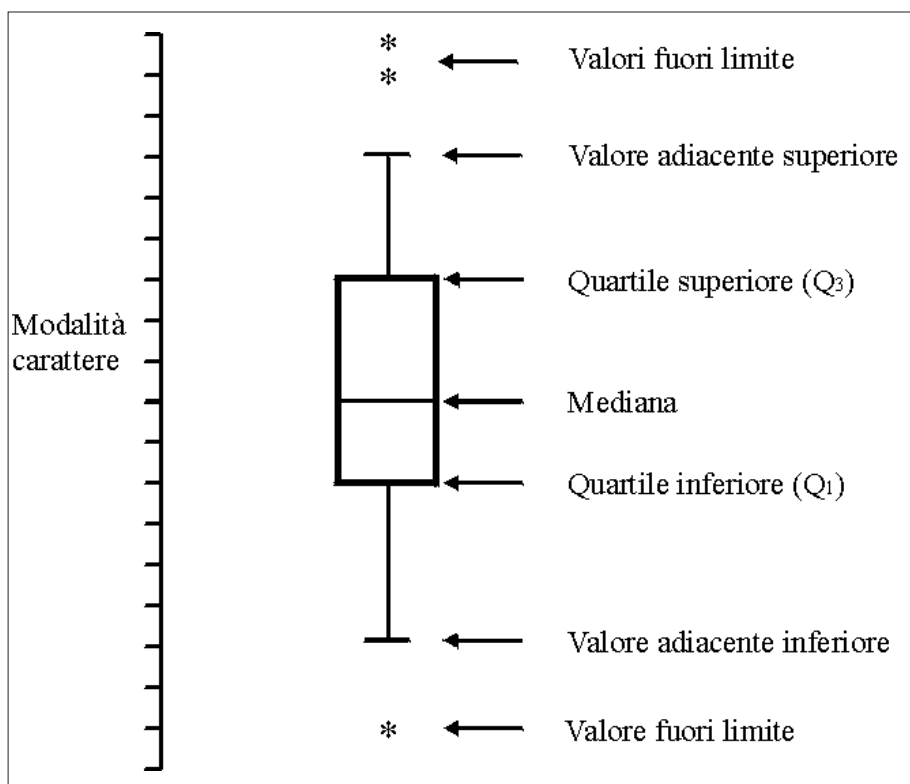


Grafico 9

Di seguito viene rappresentato il boxplot relativo alle pendenze pro-capite delle Corti di Appello.

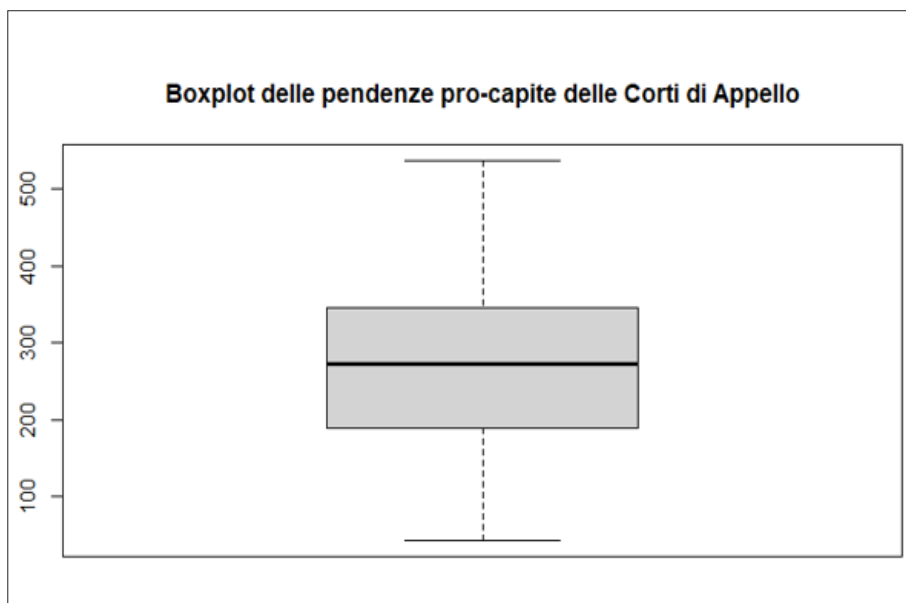


Grafico 9a. -Nostre elaborazioni su dati del Ministero della Giustizia

6. Il concetto di variabilità e gli indici relativi

I caratteri rilevati nelle unità statistiche hanno la peculiarità di presentare valori o attributi diversi del tutto o solo in parte. L'attitudine dei caratteri ad assumere modalità differenti si chiama variabilità nel caso di variabili quantitative, mentre mutabilità nel caso di variabili qualitative.

La variabilità assume significati specifici a seconda del contesto in cui si manifesta, e quindi può rappresentare una misura di imprecisione, di incertezza, di concentrazione.

Gli indici di variabilità sono nulli se tutti i termini della distribuzione sono uguali tra loro e cioè se la popolazione è omogenea rispetto alla variabile considerata che pertanto risulta concentrata in una sola modalità.

Al contrario tali indici assumono valori maggiori all'aumentare della distanza tra i valori assunti dalla variabile; ad esempio due unità con valori

di una variabile pari a 12 e 10 sono molto vicine tra loro; al contrario due ulteriori unità con valori pari a 12 e 100 hanno una maggiore distanza tra loro e pertanto maggiore è la variabilità.

La costruzione di un indice di variabilità riflette i diversi obiettivi di misurazione; ci sono indici che misurano

- ✓ la variabilità del carattere mediante una sintesi delle misure degli scarti in valore assoluto tra ogni termine della distribuzione e una media M scelta come riferimento ($|x_i - M|$).
La media M può coincidere con la media della distribuzione della variabile ($\bar{X}$) oppure può coincidere con un altro parametro preso come riferimento,
- ✓ la variabilità del carattere mediante una sintesi delle misure degli scarti in valore assoluto tra tutti i termini della distribuzione tra loro ($|x_i - x_j|$).
- ✓ la variabilità del carattere considerando la diversità tra le modalità di due particolari unità della distribuzione o tra due quantili⁴.

La varianza

Uno degli indici di variabilità più comunemente usati è la *varianza*, che appartiene al primo gruppo delle tipologie di indici sopra menzionati.

La varianza è la Media dei quadrati degli scarti dei valori di una variabile dalla sua media; tradotto in formule equivale a:

$$\text{Var}(X) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2$$

Il numeratore della varianza è detto *devianza*, mentre la radice quadrata della varianza è detta *deviazione standard* (o scarto quadratico medio); anch'essi sono indici di variabilità.

La varianza varia tra 0 a $+\infty$: assume valore nullo quando tutti i valori sono uguali alla media e pertanto la variabile X assume un solo valore su tutte le unità del collettivo; più i valori si distanziano dalla media e perciò differiscono da essa più il valore della varianza cresce. Nell'esempio della Tab. 4 relativo alle pendenze pro-capite, la varianza è pari a 14.672; ciò indica che effettivamente esiste variabilità ma non avendo la varianza un limite superiore non è sempre possibile quantificare opportunamente l'ampiezza di tale variabilità. Tale indicatore può, però, essere utilizzato per

⁴ Il quantile di ordine α (con α compreso tra 0 ed 1) è un valore che divide la popolazione di riferimento in due parti: una quota di valori più piccoli pari ad α ed una quota pari ad $1-\alpha$ più grandi.

confrontare tra loro distribuzioni omogenee con riferimento all'unità di misura (ad esempio per paragonare la variabilità nelle pendenze pro-capite di un distretto rispetto a quella di un altro).

Un indice di variabilità che possa essere utilizzato anche per confrontare fenomeni che hanno unità di misura differenti e che quantifichi anche il livello di variabilità è il coefficiente di variazione.

Il coefficiente di variazione

Il coefficiente di variazione (Cv) è dato dal rapporto tra la deviazione standard e la media della variabile presa in considerazione; è un indice di dispersione che permette di confrontare misure di fenomeni riferite a unità di misura diverse. Il coefficiente di variazione può essere considerato anche un indice della precisione della misura e può essere utilizzato anche per confrontare la variabilità di diverse distribuzioni.

In generale se il Cv è superiore al 100% significa che la variabile considerata ha una variabilità superiore alla media e pertanto i valori assunti sono molto diversi tra loro e rispetto al valore medio. Tuttavia, la variabilità può essere consistente anche se il Cv è inferiore al 100% e il suo livello si definisce solo in confronto al Cv di un'altra o di altre variabili.

Le tabelle seguenti riportano la media, la deviazione standard e il coefficiente di variazione delle sopravvenienze degli uffici di Tribunale e di Corte di Appello per macro-materia civile.

Se volessimo confrontare il numero di sopravvenienze negli uffici giudiziari d'Italia per ciascuna materia/settore, dovremmo attenderci una elevata variabilità rispetto alla media nazionale in virtù del diverso ammontare delle iscrizioni in funzione della dimensione territoriale di ciascun ufficio. Potrebbe essere interessante anche verificare se tale variabilità sia diversa tra materie differenti.

La Tab. 10 e la Tab. 11 mostrano che sia i Tribunali che le Corti di Appello presentano nel numero di sopravvenienze del Lavoro e Previdenza e dei Procedimenti speciali sommari, coefficienti di variazione più grandi rispetto alle altre macro materie e circa pari o superiori al 100% (variabilità maggiore della media delle sopravvenienze nella stessa macro materia del Lavoro e Previdenza e in quella dei Procedimenti speciali). In particolare, per i Tribunali il Cv è pari a 144% per Lavoro e Previdenza e al 148% per i Procedimenti speciali (Tab. 10); per quanto concerne invece le Corti di Appello il Cv per il Lavoro e Previdenza è pari a 101% mentre per i Procedimenti speciali è pari a 99% (Tab. 11).

Tab. 10. Media, Deviazione standard e coefficiente di variazione delle sopravvenienze nei Tribunali per macromateria civile.

Macromaterie	Media	Deviazione standard	Cv
Altre procedure	17	17	99%
Civile ordinario	23	32	138%
Esecuzioni civili	18	23	132%
Fallimentare e altre procedure concorsuali	2	3	140%
Lavoro e previdenza	15	21	144%
Procedimenti speciali sommari	21	31	148%
Volontaria giurisdizione	10	10	101%
Totale	106	130	123%

Nostre elaborazioni su fonte Tableau civile (Flussi 2014-2019)

Tab. 11. Media, Deviazione standard e coefficiente di variazione delle sopravvenienze nelle Corti di Appello per macromateria civile.

Macromaterie	Media	Deviazione standard	Cv
Civile ordinario	12.513	11.432	91%
Lavoro e previdenza	6.913	6.967	101%
Procedimenti speciali sommari	121	119	99%
Volontaria giurisdizione	5.039	4.187	83%
Totale complessivo	24.585	21.647	88%

Nostre elaborazioni su fonte Tableau civile (Flussi 2014-2019)

7. Il rapporto di concentrazione

Si è già accennato alla concentrazione, affermando che un indicatore di variabilità consente di verificare quanto i valori di una distribuzione siano o meno concentrati attorno alla media; infatti, i due concetti – variabilità e concentrazione – sono strettamente connessi tra loro. La concentrazione fa specifico riferimento a caratteri quantitativi trasferibili cioè a caratteri le cui modalità possono essere trasferite tra le unità della stessa popolazione. Un esempio classico di carattere trasferibile è il reddito.

L'obiettivo della *concentrazione* è quello quindi di studiare il modo in cui l'ammontare totale di un carattere quantitativo trasferibile si ripartisce tra le unità che costituiscono la popolazione oggetto di studio.

In giustizia un carattere che potrebbe considerarsi trasferibile (ad esempio attraverso politiche di geografia giudiziaria) è il flusso di sopravvenienze negli uffici di Tribunale; le unità N sono quindi i Tribunali e il carattere X sono le sopravvenienze, e rappresenta la distribuzione dei procedimenti sopravvenuti rilevati in una popolazione P costituita dagli $N=140$ Tribunali.

Se i Tribunali avessero tutti le stesse sopravvenienze per cui $x_1 = \dots = x_i = \dots = x_N$, dove x_i rappresenta il numero di fascicoli sopravvenuti nel Tribunale i , oppure se ogni unità (ufficio) possedesse $1/N$ dell'ammontare di sopravvenienze globali, si avrebbe *concentrazione nulla* o *equi distribuzione*: la variabilità del carattere nella popolazione oggetto di studio in questo caso sarebbe nulla.

Se invece tutti i fascicoli sopravvenissero in un solo Tribunale e gli altri 139 uffici avessero sopravvenienze nulle si parlerebbe di *massima concentrazione* e in tal caso la distribuzione si caratterizzerebbe per *variabilità massima*. Analogo esempio potrebbe applicarsi nell'ipotesi di analizzare le assegnazioni degli affari ai diversi magistrati di un ufficio.

Nei casi intermedi a quelli estremi menzionati (equi distribuzione e massima concentrazione) si parla di carattere concentrato: un carattere è tanto più concentrato quanto maggiore è la sperequazione, cioè la mancanza di uniformità, del carattere stesso tra le unità della popolazione che nell'esempio corrisponde al collettivo dei Tribunali.

Per capire bene la costruzione dell'indice di concentrazione è necessario considerare una distribuzione, ad esempio quella delle sopravvenienze nei vari Tribunali e ordinare tale distribuzione in ordine crescente (o non decrescente) rispetto alle sopravvenienze.

Si può indicare con q_i la porzione di carattere totale posseduto dalle prime i unità che presentano un valore del carattere X non superiore a x_i

$$q_i = \frac{x_1 + x_2 + \dots + x_i}{x_1 + x_2 + \dots + x_i + \dots + x_N} = \frac{X(i)}{X(N)}$$

e si può indicare con p_i la frazione di unità rispetto a N costituenti la popolazione considerata e alle quali compete l'ammontare parziale

$$X(i): p_i = \frac{i}{N} \quad (i=1, \dots, N).$$

Nel caso di equi distribuzione, ossia quando l'ammontare globale è distribuito in parti uguali tra le N unità, per esempio tutti i Tribunali hanno lo stesso numero di fascicoli sopravvenuti, allora $p_i = q_i$ poiché in tale situazione una data frazione di unità possiede la medesima frazione dell'ammontare

globale (per esempio $\frac{1}{4}$ dei Tribunali possiede $\frac{1}{4}$ delle sopravvenienze).

Nel caso invece di massima concentrazione ovvero quando l'ammontare globale è posseduto da una sola unità si ha che $p_i = \frac{i}{N}$ per $(i=1, \dots, N-1)$ e $q_i = 0$ e $p_N = q_N = 1$ in generale vale sempre la condizione $p_i \geq q_i$.

Il *Rapporto di Concentrazione di Gini* è un indice che misura il grado di concentrazione ed è dato dalla seguente formula

$$R(X) = \frac{\sum_{i=1}^{N-1} (p_i - q_i)}{\sum_{i=1}^{N-1} p_i}$$

dove il denominatore rappresenta il massimo del numeratore

$$\max \sum_{i=1}^{N-1} (p_i - q_i) = (\sum_{i=1}^{N-1} p_i)$$

che si realizza quando tutte le $q_i = 0$ ($i=1, \dots, N-1$).

Il rapporto di concentrazione di Gini $R(X)$ varia tra 0 ed 1; è pari a 0 nel caso di equi distribuzione mentre assume valore pari a 1 nel caso di massima concentrazione.

La tabella seguente riporta le misure necessarie per il calcolo del rapporto di concentrazione delle iscrizioni pro-capite nelle Corti di Appello italiane.

Guardando soprattutto alle p_i e alle q_i si può calcolare che il 50% delle Corti di Appello possiede il 40% circa delle iscrizioni pro-capite complessive (non si è molto lontani dall'equi distribuzione).

Tab. 12. Attributi per il calcolo del Rapporto di Concentrazione delle iscrizioni pro-capite nelle Corti di Appello italiane

Corti di Appello	Iscritti pro-Capite 2021	Cumulata degli iscritti	q_i	i	p_i	$p_i - q_i$
			0		0	
Trento	55	55	0,01	1	0,03	0,02
Bolzano	56	110	0,03	2	0,07	0,04
Cagliari	100	210	0,05	3	0,10	0,05
Caltanissetta	103	313	0,08	4	0,14	0,06
Reggio Calabria	104	417	0,11	5	0,17	0,06
Brescia	106	524	0,14	6	0,21	0,07

Corti di Appello	Iscritti pro-Capite 2021	Cumulata degli iscritti	q_i	i	p_i	$p_i - q_i$
Milano	108	632	0,16	7	0,24	0,08
Genova	115	747	0,19	8	0,28	0,08
Trieste	116	863	0,22	9	0,31	0,09
Salerno	118	981	0,25	10	0,34	0,09
Campobasso	119	1.100	0,28	11	0,38	0,10
Torino	120	1.220	0,32	12	0,41	0,10
Venezia	131	1.351	0,35	13	0,45	0,10
Sassari	138	1.489	0,38	14	0,48	0,10
Potenza	139	1.628	0,42	15	0,52	0,10
Napoli	141	1.769	0,46	16	0,55	0,09
Firenze	141	1.910	0,49	17	0,59	0,09
Lecce	145	2.055	0,53	18	0,62	0,09
Roma	147	2.202	0,57	19	0,66	0,09
Taranto	148	2.350	0,61	20	0,69	0,08
Bari	154	2.504	0,65	21	0,72	0,08
Perugia	155	2.660	0,69	22	0,76	0,07
L'Aquila	158	2.818	0,73	23	0,79	0,06
Catanzaro	170	2.988	0,77	24	0,83	0,06
Palermo	173	3.160	0,82	25	0,86	0,05
Ancona	173	3.333	0,86	26	0,90	0,04
Catania	173	3.506	0,91	27	0,93	0,02
Messina	178	3.684	0,95	28	0,97	0,01
Bologna	185	3.869	1	29	1	0

Nostre elaborazioni dati del Ministero della Giustizia

Sommando l'ultima colonna ($p_i - q_i$) fino a Messina (N-1) e dividendola per la somma della penultima colonna (p_i) ottengo $R(X)$ che è pari a circa 0,13 indicando presenza di una bassa concentrazione delle iscrizioni pro-capite negli uffici di secondo grado e pertanto una distribuzione abbastanza

omogenea della stessa variabile nei vari uffici. Il rapporto di concentrazione dà una misura dell'Area di concentrazione che si esplicita di seguito.

I valori riportati utili per il calcolo di $R(X)$, e presenti come esempio anche nella Tab. 12, possono essere rappresentati graficamente con la cosiddetta curva di Lorenz (o curva di concentrazione), sviluppata da Max O. Lorenz nel 1905 come strumento grafico utile per l'analisi della distribuzione del reddito: sull'asse delle ascisse si rappresentano le p_i mentre sull'asse delle ordinate le q_i cioè le quantità relative cumulate.

Il segmento passante per i punti di coordinate (0,0) e (1,1) è il segmento di equi distribuzione (arancione), la curva al di sotto del segmento è la curva di concentrazione (azzurra) e l'area tra il segmento e la curva individua l'*Area di concentrazione*: più l'area è grande e maggiormente il carattere si distribuisce nelle mani di poche o di una sola unità, più l'area è piccola e più la curva si avvicina alla retta di equi distribuzione.

Il grafico seguente mostra la curva di Lorenz delle iscrizioni pro-capite nelle Corti di Appello italiane. Si può notare che la curva di concentrazione in questo caso è abbastanza vicina alla retta di equi distribuzione, infatti il rapporto di concentrazione di Gini è basso e pari a 0,13, mostrando una certa omogeneità nella distribuzione delle iscrizioni pro-capite nelle Corti di Appello (sono pochi gli uffici in cui il carattere è concentrato maggiormente e comunque non in modo troppo differente rispetto agli altri).

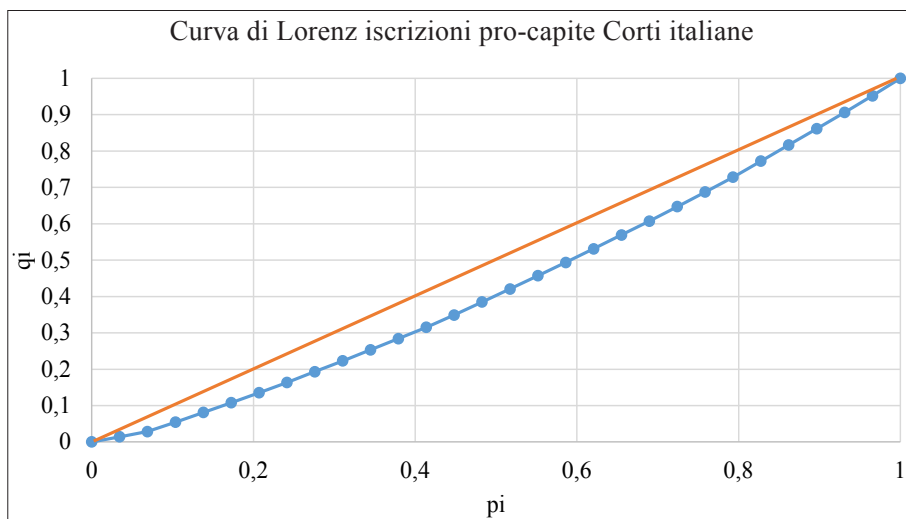


Grafico 10 - Nostre elaborazioni su dati Ministero della Giustizia

Diverso è il discorso se si considerano le iscrizioni totali e non pro-capite relative al 2021 nelle Corti di Appello; il rapporto di concentrazione di Gini è in questo caso pari a 0,43 circa, mostrando una modesta concentrazione delle iscrizioni in un gruppo di uffici rispetto agli altri (vedi Grafico 11, il 50% della popolazione possiede circa il 20% del reddito). La differenza rispetto all'analisi precedente effettuata sulle iscrizioni pro-capite è sicuramente dovuta alla diversa dimensione degli uffici di secondo grado.

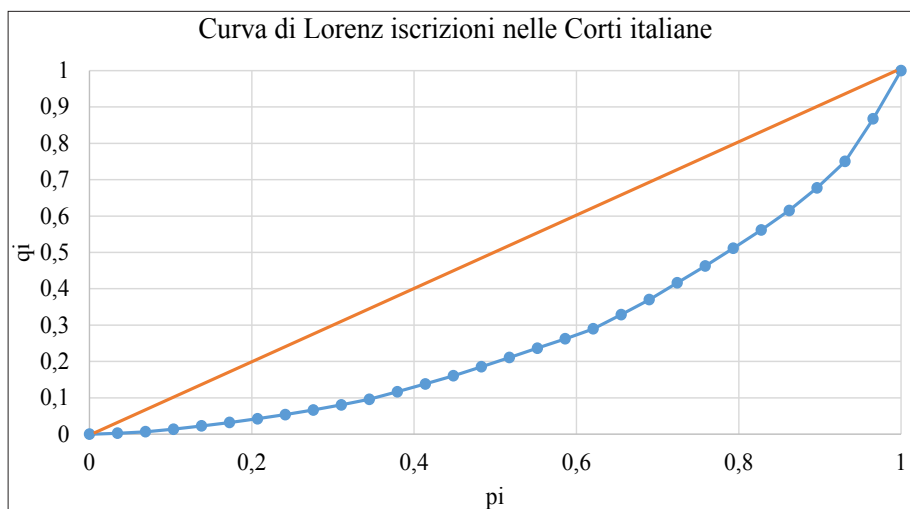


Grafico 11

8. La Mutabilità

Nel caso di caratteri qualitativi, sconnessi o ordinati, non si parla di variabilità ma di mutabilità. La mutabilità è la tendenza dei caratteri qualitativi, sconnessi o ordinati ad assumere differenti modalità; se un carattere si manifesta sempre con la stessa modalità sulla popolazione oggetto di studio allora la mutabilità è nulla e la popolazione è detta *omogenea* rispetto a quel carattere.

Si consideri una distribuzione in cui il carattere qualitativo X - ad esempio X = dimensione dei Tribunali - ha cinque (m) distinte modalità x_1, x_2, x_3, x_4 e x_5 - ovvero “metropolitano”, “grande”, “medio grande”, “medio piccolo”, “piccolo”; il totale delle frequenze, pari al numero di tribunali, è 140 e tale totale può essere ripartito tra le cinque modalità (m) in diversi modi,

a seconda dei quali vi sarà un diverso grado di mutabilità. Nello specifico:

- se il carattere assume la stessa modalità su ciascuna delle 140 unità, come accennato sopra, si ha massima *omogeneità* e quindi mutabilità zero in quanto le modalità non mutano in corrispondenza delle unità.
- se, invece, si hanno frequenze equidistribuite per ciascuna modalità dimensionale (ciascuna modalità si presenta $140:5=28$ volte) allora si ha massima mutabilità che coincide con la condizione di massima *eterogeneità* tra le unità ed *equi distribuzione* delle modalità osservate.

Gli indici di entropia più noti e utilizzati per la misurazione della mutabilità sono l'indice di entropia introdotto da Hartley e generalizzato da Shannon e l'indice di Gini.

L'indice di entropia di Shannon è il seguente:

$$H = - \sum_{i=1}^m f_i \log f_i$$

Esso varia tra 0 e $\log(m)$: 0 nel caso di massima omogeneità e $\log(m)$ nel caso di massima eterogeneità. L'indice di entropia normalizzato è il seguente e consiste nel dividere H per il suo massimo in modo tale che l'indice vari tra 0 ed 1 (omogeneità e eterogeneità rispettivamente).

$$H^* = \frac{-\sum_{i=1}^m f_i \log f_i}{\log m}$$

L'indice di mutabilità di Gini è invece definito dalla seguente espressione e varia tra 0 e $(m-1)/m$

$$G = 1 - \sum_{i=1}^m f_i^2$$

Normalizzando l'indice di Gini si ottiene il seguente indice che assume valore nullo in assenza di mutabilità e valore unitario nel caso di massima mutabilità ed eterogeneità e perciò di equi distribuzione delle unità tra le m modalità totali.

$$G^* = 1 - \sum_{i=1}^m f_i^2 \quad (m / m-1)$$

La tabella seguente mostra le misure necessarie per il calcolo degli indici di entropia di Shannon e di Gini per quanto concerne la dimensione dei Tribunali italiani.

Tab. 13. Misure per il calcolo dell'entropia della distribuzione dei Tribunali per dimensione

Dimensione	Numero relativo di sedi (f_i)	$f_i \cdot \log f_i$	f_i^2
Metropolitani	0,02	-0,08	0,01
Grandi	0,02	-0,08	0,01
Medio grandi	0,14	-0,28	0,08
Medio piccoli	0,47	-0,35	0,13
Piccoli	0,34	-0,37	0,13

L'indice di entropia di Shannon H può essere calcolato sommando la penultima colonna ($f_i \cdot \log f_i$) e cambiandogli di segno, così come la formula richiede; $H = -1,15$, mentre H^* cioè l'indice di Shannon normalizzato si ottiene in questo caso dividendo H per il log (5) in quanto $m=5$ rappresenta il numero totale di modalità, $H^*=0,72$.

L'indice di entropia di Gini si ottiene sottraendo da 1 la somma dell'ultima colonna (f_i^2) $G=0,65$, mentre l'indice di Gini normalizzato si ottiene in questo caso moltiplicando G per $5/4$ ed è pari pertanto a $G^*=0,81$.

Guardando ad entrambi gli indici di entropia normalizzati ($H^*=0,72$ e $G^*=0,81$) si può notare la presenza di una eterogeneità abbastanza elevata con modalità molto mutevoli.

9. Distribuzioni doppie di frequenze

Se le frequenze assolute vengono associate non a singole modalità di una variabile ma a coppie di modalità (x_i, y_j) $i=1, \dots, m$ e $j=1, \dots, t$ (dove $m=t$ oppure $m \neq t$) appartenenti a due variabili diverse X e Y , si hanno distribuzioni doppie di frequenze. Le distribuzioni doppie possono essere rappresentate attraverso tabelle di frequenza a doppia entrata in cui le frequenze di cella n_{ij} rappresentano appunto il numero di unità su cui si manifestano simultaneamente le modalità ij appartenenti alle variabili X e Y .

Viene riportato di seguito un esempio relativo alla dimensione dei Tribunali per classi di sopravvenienze degli stessi individuate sulla base dei quartili. In questo caso l'indice di riga i (il numero di modalità della dimensione) va da 1 a 5, mentre j che indica le modalità della variabile di colonna Y va da 1 a 4, pertanto ad esempio $n_{43}=34$ rappresentando la frequenza di Tribunali su cui si manifestano le modalità "medio piccoli" e

sopravvenienze della classe “70.674-112.422” in concomitanza. Vengono rappresentati i totali di riga e di colonna, dette anche marginali di riga e colonna; si può notare che l’ammontare di colonna è pari a 35 per tutte le modalità delle classi di sopravvenienze perché come anticipato prima le classi sono state create sulla base dei quartili che dividono la variabile in 4 parti uguali. La somma dei totali marginali di riga così come la somma dei totali marginali di colonna è pari al totale delle unità. Le frequenze marginali di riga e di colonna vengono indicate come n_i e n_j per $i=1, \dots, s$ e $j=1, \dots, t$; nell’esempio specifico $s=5$ e $t=4$ e $n_1=3$, $n_2=3$, $n_3=18$, $n_4=65$, $n_5=51$, invece $n_{.1}=n_{.2}=n_{.3}=n_{.4}=35$.

Tab. 14. Tabella doppia di frequenze per dimensione e classi di sopravvenienze

Dimensione	Classi di sopravvenienze				Totale
	≤ 44.018	44.018-70.674	70.674-112.422	≥ 112.422	
Metropolitani				3	3
Grandi				3	3
Medio grandi			1	17	18
Medio piccoli	1	18	34	12	65
Piccoli	34	17			51
Totale	35	35	35	35	140

A partire dalla Tab. 14 è possibile derivare un’altra tabella (Tab. 15) con le frequenze relative semplicemente rapportando il contenuto di ogni cella al totale delle frequenze e perciò dei Tribunali. Si ottiene perciò la tabella doppia seguente delle frequenze relative in cui questa volta $f_{43} = n_{43}/N$ è pari a $0,24=34/140$. Le frequenze marginali relative di riga e di colonna si indicano rispettivamente con f_i e f_j nell’esempio $i=1, \dots, 5$ e $j=1, \dots, 4$. Come si può notare dalla Tab. 15, $f_{1.}=0,02$, $f_{2.}=0,13$, $f_{3.}=0,46$, $f_{4.}=0,02$, $f_{5.}=0,36$, mentre $f_{.1}=f_{.2}=f_{.3}=f_{.4}=0,25$.

La somma delle marginali di riga così come la somma delle marginali di colonna nella Tab. 15 è pari a 1.

Tab. 15. Tabella doppia di frequenze relative per dimensione e classi di sopravvenienze

Dimensione	Classi di sopravvenienze				Totale
	≤ 44.018	44.018- 70.674	70.674- 112.422	≥ 112.422	
Metropolitani	0,00	0,00	0,00	0,02	0,02
Grandi	0,00	0,00	0,00	0,02	0,02
Medio grandi	0,00	0,00	0,01	0,12	0,13
Medio piccoli	0,01	0,13	0,24	0,09	0,46
Piccoli	0,24	0,12	0,00	0,00	0,36
Totale	0,25	0,25	0,25	0,25	1,00

Capitolo V

Statistica per la giustizia: come analizzare i dati

La trasformazione dei dati in informazione è possibile attraverso l'analisi dettagliata del fenomeno e delle sue variabili, l'individuazione delle eventuali relazioni esistenti, la verifica delle tendenze in atto, la stima dei parametri di interesse sull'intera popolazione di riferimento, l'applicazione di metodi di elaborazione e modellazione dei dati che richiedono conoscenze statistiche specifiche.

In questo capitolo si forniscono i principali strumenti impiegabili per analizzare, comprendere, valutare, interpretare i dati giudiziari e individuare le possibili variabili che ne influenzano l'andamento.

1. *La connessione statistica*

Il carattere Y si dice indipendente dal carattere X se il fatto che X presenti la modalità x_1 o la x_2 o la x_s non modifica la distribuzione di Y ossia le distribuzioni di Y condizionate alle modalità x_s di X .

In particolare, la Y è indipendente dalla X se la frequenza assoluta n_{ij} è uguale a Nn_{ij} per ogni i e ogni j , dove

$$n_{ij}^* = \frac{n_i \cdot n_j}{N}$$

e cioè quando la frequenza osservata n_{ij} è uguale alla frequenza teorica n_{ij}^* per ogni i e ogni j .

Ovviamente si può anche dire che Y è indipendente dalla X se la frequenza relativa osservata f_{ij} è uguale alla frequenza relativa teorica f_{ij}^* per ogni i e ogni j dove

$$f_{ij}^* = f_i \cdot f_j$$

La *connessione* è una deviazione da una situazione di indipendenza statistica; perché due caratteri siano connessi è sufficiente verificare che almeno una frequenza assoluta effettiva della distribuzione doppia non

soddisfi l'uguaglianza $n_{ij} = n_{ij}^*$.

Il X^2 è un indice che permette di verificare la connessione tra due caratteri X e Y ed è definito come

$$X^2 = \sum_i^m \sum_j^t \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*}$$

e misura la distanza tra la tabella effettiva e la tabella teorica. Più il valore del X^2 è grande, maggiore è la distanza da una situazione di indipendenza; il valore dell'indice tuttavia risulta essere più elevato all'aumentare di N e tale inconveniente può essere ovviato dividendolo per il totale delle unità. Si ha pertanto l'indice $\varphi^2 = \frac{X^2}{N}$ che a sua volta può essere normalizzato in modo da poter variare tra 0 e 1 e indicare così rispettivamente connessione nulla e connessione massima. Si ha pertanto l'indice di Cramer

$$V = \frac{\varphi^2}{\max \varphi^2} = \frac{\varphi^2}{\min(m-1, t-1)} = \frac{X^2}{N \min(m-1, t-1)}.$$

Consideriamo l'esempio in Tab. 15 e calcoliamo la tabella delle frequenze teoriche N_{ij}^* secondo quanto detto precedentemente.

Tab. 16

Dimensione	Classi di sopravvenienze				Totale
	≤ 44.018	44.018- 70.674	70.674- 112.422	≥ 112.422	
Metropolitani	0,75	0,75	0,75	0,75	3
Grandi	0,75	0,75	0,75	0,75 (=3*35/140)	3
Medio grandi	4,5	4,5	4,5	4,5	18
Medio piccoli	16,25	16,25	16,25	16,25	65
Piccoli	12,75	12,75	12,75	12,75	51
Totale	35	35	35	35	140

Successivamente calcoliamo anche la tabella delle quantità $\frac{(N_{ij} - N_{ij}^*)^2}{N_{ij}^*}$

utilizzando la Tab. 14 per i valori N_{ij} e la Tab. 16 per i valori N_{ij}^* .

Tab.17

Dimensione	Classi di sopravvenienze				Totale
	≤ 44.018	44.018- 70.674	70.674- 112.422	≥ 112.422	
Metropolitani	0,75	0,75	0,75	6,75	0
Grandi	0,75	0,75	0,75	6,75	0
Medio grandi	4,5	4,5	2,72	34,72	0
Medio piccoli	14,31	0,19	19,39	1,11	0
Piccoli	35,42	1,42	12,75	12,75	0
Totale	0	0	0	0	0

A questo punto l'indice X^2 è pari alla somma delle celle della tabella sovrastante e cioè 162, il $\varphi^2 = 162/140 = 1,15$ e la V di Cramer è pari a $V = \frac{\varphi^2}{3} = 0,38$ dove 3 è il minimo tra il numero delle modalità di righe -1 e delle modalità di colonna -1. Il valore della V indica la presenza di una connessione tra la variabile X= dimensione e la variabile Y= classi di sopravvenienze anche se questa è bassa.

2. La Concordanza

Nel caso di variabili categoriche può essere misurato il grado di concordanza solo se le variabili sono dotate dello stesso numero di modalità e se ogni modalità di X è affine a una e solo una modalità di Y e viceversa: ad esempio si vuole valutare la coincidenza di opinioni di due giudici. Un indice di concordanza molto utilizzato è il *K di Cohen* (<https://ssmtstatistica.files.wordpress.com/2015/12/kappacoen.pdf>) che fornisce appunto in questo caso una misura del grado di accordo tra due giudici; per capire meglio la costruzione dell'indice si consideri la tabella seguente.

Tab. 18. Tabella di contingenza sui giudizi dei giudici A e B su tre Categorie

		Giudice A			Totale
		Categoria 1	Categoria 2	Categoria 3	
Giudice B	Categoria 1	88	14	18	120
	Categoria 2	10	40	10	60
	Categoria 3	2	6	12	20
Totale		100	60	40	200

Nel calcolo del K di Cohen si considerano soltanto le frequenze poste sulla diagonale principale, quelle cioè che riguardano l'incrocio tra le categorie affini. Se le frequenze osservate sono maggiori rispetto ai valori teorici allora si ha una concordanza altrimenti si ha una discordanza. In dettaglio l'indice K di Cohen è calcolato come $K = \frac{\theta - \hat{\theta}}{1 - \hat{\theta}}$, dove $\theta = \sum_{i=1}^I f_{ii}$, e $\hat{\theta} = \sum_{i=1}^I f_{i.} f_{.i}$ e le f_{ii} sono le frequenze relative delle celle diagonali e le $f_{i.} f_{.i}$ è il prodotto delle marginali delle celle diagonali (i varia da 1 a $I=3$ nell'esempio).

L'indice di Cohen varia tra $-\hat{\theta}/(1 - \hat{\theta})$ e 1: assume valore unitario nel caso di perfetta concordanza, assume valore pari a $-\hat{\theta}/(1 - \hat{\theta})$ nel caso di perfetta discordanza e assume valore pari a zero nel caso di completa indipendenza tra i giudizi dei giudici. Nell'esempio in Tab. 18, il minimo che l'indice K può assumere è -0,69.

Tab. 19. Tabella di contingenza delle frequenze relative sui giudizi dei giudici A e B su tre Categorie

		Giudice A			
		Categoria 1	Categoria 2	Categoria 3	Totale
Giudice B	Categoria 1	0,44	0,07	0,09	0,6
	Categoria 2	0,05	0,2	0,05	0,3
	Categoria 3	0,01	0,03	0,06	0,1
	Totale	0,5	0,3	0,2	1

Tab. 20. Tabella di contingenza delle frequenze relative teoriche sui giudizi dei giudici A e B su tre Categorie

		Giudice A			Totale
		Categoria 1	Categoria 2	Categoria 3	
Giudice B	Categoria 1	0,3	0,18	0,12	0,6
	Categoria 2	0,15	0,09	0,06	0,3
	Categoria 3	0,05	0,03	0,02	0,1
	Totale	0,5	0,3	0,2	1

A partire dalle tabelle precedenti è possibile calcolare rispettivamente θ e $\hat{\theta}$ come somme dei valori di diagonale delle Tabelle 19 e 20, pertanto il valore dell'indice K è dato da $K = \frac{\theta - \hat{\theta}}{1 - \hat{\theta}} = \frac{0,7 - 0,41}{1 - 0,41} = 0,492$.

Si può notare come il valore di K indichi che i due giudici siano moderatamente concordanti nelle valutazioni.

3. La relazione tra variabili

La relazione tra variabili può essere misurata attraverso la correlazione, un indice simmetrico che verifica se esiste o meno la tendenza di due caratteri a variare insieme; la correlazione non verifica l'esistenza di un rapporto di causalità ovvero di una relazione tra due variabili in cui una è un fattore determinante al manifestarsi dell'altra variabile (es. il fumo come fattore determinante nella comparsa del cancro al polmone).

A volte può sembrare che alcune variabili sembrino correlate ma in realtà non vi è una vera relazione sottesa ma piuttosto la presenza di una terza variabile non osservata che può far sembrare la relazione esistente tra le due più forte o più debole di quanto in realtà non sia. In questo ultimo caso si parla di correlazione spuria; un esempio è la correlazione che sembra forte tra il numero di matrimoni e il numero di rondini in cielo che non è dovuta al fatto che effettivamente il numero di matrimoni vari al variare del numero di rondini o viceversa ma al fatto che in certi paesi le rondini compaiono durante le loro migrazioni in primavera ed autunno che sono pure i periodi preferiti dalle coppie di convogliare a nozze. La variabile in comune alle rondini e ai matrimoni è la stagione primaverile la quale produce una correlazione appunto spuria.

Se la correlazione è un indice simmetrico, la regressione è un modello "asimmetrico" in cui una variabile può "influenzare" l'altra; la logica sarà illustrata nel seguito.

Come già descritto in precedenza, il grafico a dispersione è molto usato anche al fine di rappresentare l'esistenza di una relazione lineare tra due variabili quantitative il cui legame di dipendenza l'una dall'altra può essere tradotto in forma funzionale. In questo caso una delle due variabili, generalmente quella con antecedenza logica, viene considerata come variabile *indipendente* perché si suppone sia *esplicativa* del comportamento dell'altra (ad esempio il fumo nella comparsa del cancro al polmone). A volte è possibile che entrambi i caratteri siano dipendenti l'uno dall'altro e allora è necessario ricercare la legge di dipendenza di Y da X e di X da Y.

La legge di dipendenza di Y da X più semplice è quella espressa dal polinomio di grado 1: la curva che rappresenta la relazione funzionale tra X e Y è una retta.

La retta di regressione lineare è esplicitata dalla seguente equazione $Y = a + bX + e$, dove a è l'intercetta della retta e cioè il valore di Y quando il valore di X è pari a zero, b è il coefficiente angolare della retta che ne indica la pendenza ed e è il termine di errore che indica ciò che il modello non riesce a cogliere.

Data una nuvola di punti, rappresentata nel diagramma a dispersione, è possibile stimare la retta di regressione o meglio i parametri della retta che minimizza gli scarti tra i punti e la retta stessa adattandosi così quanto più possibile ai dati. L'indice di bontà di adattamento ai dati è l'indice di determinazione o R^2 che può variare tra 0 ed 1: 0 nel caso di assenza di dipendenza lineare, 1 nel caso di perfetta dipendenza. In quest'ultimo caso l'errore è pari a 0 e ciò significa che non c'è differenza tra le y osservate e quelle stimate dal modello e la retta passa perfettamente per i punti.

Nel grafico seguente viene rappresentata la retta di regressione lineare e la sua equazione con le stime dei parametri tra le due variabili "pendenze pro-capite" e "iscritti pro-capite" delle Corti di Appello. Il coefficiente angolare stimato è pari a 2,47 circa e indica che all'aumentare di una unità delle iscrizioni aumentano mediamente di 2,47 le pendenze finali. Inoltre, il coefficiente di determinazione R^2 è pari a 0,46 circa, indicando un modesto adattamento del modello ai dati osservati e pertanto una modesta dipendenza lineare tra le due variabili: l'aumento delle pendenze finali pro-capite dipende linearmente e mediamente dall'aumento delle iscrizioni pro-capite. Ovviamente questo è solo un esempio che non tiene conto delle definizioni che giocano un ruolo fondamentale nella determinazione delle pendenze finali e delle restanti variabili che possono influenzare la variazione delle pendenze e/o delle iscrizioni.

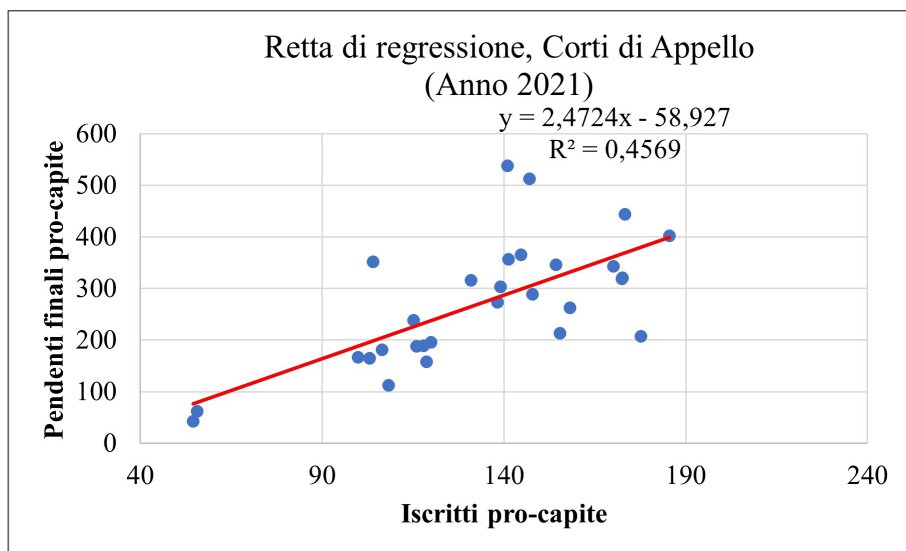


Grafico 12 - Nostre elaborazioni su dati del Ministero della Giustizia

Per verificare se il modello di regressione attuato è un buon modello previsionale, è necessario attuare una attenta analisi dei residui (cioè della parte di variabilità non spiegata dal modello) che rappresentano lo scarto tra la variabile Y considerata dipendente (variabile risposta) e la variabile stimata attraverso il modello.

In particolare, è necessario verificare se i residui sono incorrelati tra loro e con le variabili indipendenti, se sono influenzati dalla presenza di outlier e cioè di valori anomali che possono portare più su o più giù la retta di regressione, se si distribuiscono normalmente¹, etc.

Una delle ipotesi fondamentali alla base del modello lineare è quella di omoschedasticità dei residui e cioè di indipendenza dei residui dalle variabili esplicative; quando l'analisi dei residui mostra una tendenza degli stessi a covariare con una o più variabili esplicative i residui vengono detti eteroschedastici.

Nel caso di eteroschedasticità i modelli di regressione utilizzati sono detti lineari generali; essi permettono di stimare i parametri di regressione tenendo conto appunto della non omoschedasticità dei residui.

Un'altra ipotesi fondamentale alla base del modello di regressione lineare è che la variabile risposta si distribuisca secondo una distribuzione normale; tuttavia qualora la variabile risposta sia binaria (successi-insuccessi) o quantitativa discreta (es. numero di telefonate in un giorno) la sua distribuzione non è più normale ma rispettivamente bernoulliana² o poissoniana. In questi casi i modelli di *regressione* effettuati sono *lineari generalizzati* in quanto generalizzano il modello classico di regressione a casi di non normalità della variabile endogena.

Nei modelli lineari generalizzati (GLM) la variabile risposta può essere distribuita come una qualsiasi variabile casuale appartenente alla *famiglia esponenziale*³; una specifica funzione legame (funzione link) si preoccupa

¹ La distribuzione normale è una distribuzione continua di probabilità di un fenomeno statistico intorno alla media. È anche conosciuta come distribuzione gaussiana dal nome del matematico tedesco Gauss che ne è l'autore. È molto usata nell'analisi statistica, perché descrive efficacemente gran parte dei fenomeni naturali: le probabilità sono più elevate intorno alla media e si riducono man mano che ci si allontana dal valore medio a destra e a sinistra.

² Una distribuzione bernoulliana è una distribuzione di probabilità su due soli valori 0 ed 1 (fallimento, successo) mentre la distribuzione di Poisson esprime la probabilità per un numero di eventi che si verificano successivamente e indipendentemente in un dato intervallo di tempo (es. numero di telefonate ricevute in un giorno). Mentre la distribuzione normale è definita su tutto $\mathbb{R}$, le distribuzioni di Bernoulli e di Poisson sono definite in un sottospazio di $\mathbb{R}$ e soltanto per valori positivi.

³ Una famiglia esponenziale è un oggetto matematico che è, in probabilità e statistica, una

di mappare la variabile risposta dal suo sottospazio di definizione (esempio per la bernoulliana 0,1) a tutto lo spazio $\mathbb{R}$ (in cui la variabile normale è definita). Il modello lineare classico (distribuzione gaussiana della variabile risposta) diventa così un caso particolare dei GLM dove il legame funzionale è quello dell'identità⁴.

Nei modelli lineari generalizzati l'assunzione di normalità alla base del modello di regressione viene ripristinata attraverso la funzione link e l'ipotesi di omoschedasticità dei residui deve essere garantita.

4. Cenni alla verifica delle Ipotesi

La verifica delle ipotesi è una procedura inferenziale che ha lo scopo di stabilire se l'informazione ottenuta da una statistica campionaria può essere estesa ai parametri della popolazione.

Se consideriamo un campione di Tribunali e il totale (popolazione) di Tribunali, per esempio si potrebbe dire che l'efficienza nei settori civili e penali in termini di numero di procedimenti definiti negli uffici giudiziari potrebbe considerarsi migliore se ogni ufficio riducesse il tempo di definizione dei procedimenti.

Dai dati della Cepej (<<https://osservatoriocpi.unicatt.it>>) risulta che nel 2018 i procedimenti in primo grado in Italia hanno avuto una durata media stimata circa pari a 527 giorni.

Sulla base di questa informazione possiamo formulare un test di ipotesi "verosimile" assumendo un valore vicino a 527 giorni per la durata dei procedimenti in primo grado come parametro della popolazione; l'ipotesi che il parametro della popolazione sia uguale a un dato valore prende il nome di ipotesi nulla e viene indicata con H_0 .

Supponiamo pertanto che $H_0: \mu = 527$ dove rappresenta appunto la durata media dei procedimenti giudiziari nei Tribunali (parametro nella popolazione). Oltre all'ipotesi nulla ci dovrà essere una ipotesi alternativa, ad esempio $H_1: \mu \neq 527$ (nel caso di test bilaterale), e cioè una possibile ipotesi a cui si arriva qualora le evidenze campionarie indichino l'impossibilità ad accettare H_0 .

Un test di ipotesi consiste pertanto nella verifica della veridicità

classe di distribuzioni di probabilità che comprende la distribuzione binomiale, poisson, gamma, normale, etc..

⁴ Nel caso in cui la variabile risposta abbia una distribuzione normale la funzione link è pari a $g(\mu_i) = \mu_i = y_i$ ed è chiamata funzione link identità.

dell'ipotesi nulla o meno nella popolazione sulla base delle informazioni campionarie, nell'esempio la durata media in primo grado di un sottogruppo dei 140 Tribunali che possiamo indicare con $\bar{X}$ (media campionaria).

Si vuole “verificare” se un campione casuale di uffici di Tribunale contiene abbastanza evidenza per rifiutare l'ipotesi nulla in questione; pertanto si scelgono casualmente ad esempio 40 Tribunali, si calcolano le durate medie dei procedimenti e si vede se tale media campionaria è tanto inferiore o maggiore a 527 giorni.

Un test quindi è una statistica che fa corrispondere a ogni campione un valore che può essere considerato coerente con H_0 o non coerente con H_0 . Un test statistico dà luogo alla ripartizione dello spazio campionario in due sottoinsiemi complementari:

- *Regione di accettazione*, costituita da valori del test compatibili con l'ipotesi nulla H_0 ;
- *Regione critica o di rifiuto*, che raggruppa i valori del test considerati incompatibili con H_0 .

Quando il valore campionario della statistica test cade nella regione critica, l'evidenza empirica del fenomeno studiato porta a ritenere che l'ipotesi H_0 non possa essere considerata valida e che quindi non possa essere accettata come vera.

Pertanto, se la statistica calcolata cade nella zona di rifiuto decido di rifiutare H_0 ; in caso contrario decido di non rifiutare H_0 ; in alternativa all'approccio della regione di rifiuto si può riportare il *p-value* che indica la *probabilità di rifiutare H_0 quando essa è vera*. Se il p-value è inferiore o uguale a 0,05⁵ si rifiuta l'ipotesi nulla in quanto la differenza tra valore vero e valore stimato è statisticamente significativa e non può essere attribuita al caso, in caso contrario non si rifiuta l'ipotesi nulla.

Consideriamo lo stesso esempio di sopra in cui però si assume che ad oggi (nell'ipotesi che la durata dei procedimenti sia diminuita nel tempo) la media della durata nel totale dei Tribunali italiani sia pari a μ 465 giorni; si immagina di osservare in un campione di 40 Tribunali una durata media campionaria $\bar{X}$ di 480 giorni e una deviazione standard campionaria s pari a 65. Vogliamo

⁵ Per effettuare un test di ipotesi è necessario fissare un valore soglia che indica il *livello di significatività del test*, indicato con α , che indica quanto deve essere grande la differenza tra le medie considerate perché il risultato sia statisticamente significativo. Se il p-value è vicino al valore soglia è necessaria attenzione.

Il valore soglia pari a 0,05 si usa convenzionalmente e fu proposto da Fisher, essendo sufficientemente piccolo da poter concludere che sia piuttosto improbabile che la differenza osservata sia dovuta semplicemente al caso.

sapere se possiamo non rifiutare l'ipotesi nulla per cui la durata media dei procedimenti giudiziari è pari a 465, sulla base della media campionaria misurata sul campione dei 40 Tribunali pari a 480 giorni.

La statistica test è la statistica *t di student*⁶ che si ottiene sottraendo la media della popolazione dalla media campionaria e dividendo il tutto per la deviazione standard campionaria (*s*) fratto la radice quadrata della dimensione campionaria (40 Tribunali): $t = (\bar{X} - \mu)/(s/\sqrt{n})$.

In questo caso sapendo che sotto H_0 , cioè se l'ipotesi nulla fosse vera, la media della durata sarebbe pari a μ 465, la nostra *t* è pari a 1,46. Siccome il test è bilaterale in quanto l'ipotesi alternativa è che la durata media sia diversa da 465 (e può essere pertanto maggiore o minore) abbiamo due valori di *t* speculari, $t_1 = -1,46$ e $t_2 = 1,46$.

Sapendo il valore di *t*, è possibile calcolare il *p-value* (avendo a disposizione le tavole statistiche della distribuzione *t di student*) che in questo caso è pari a 0,15 (0,07 per $\mu > 465$ e 0,07 per $\mu < 465$) circa ed è pertanto maggiore, anche se di poco, di 0,05; ciò indica che, secondo la statistica test, il campione utilizzato si trova nella zona di non rifiuto di H_0 . L'ipotesi di una durata media dei procedimenti pari a 465 è pertanto considerata verosimile ovvero molto probabile (al 95%).

Per spiegare graficamente quello che abbiamo appena detto, riportiamo il grafico seguente (Grafico 13) in cui viene rappresentata l'area di non rifiuto dell'ipotesi nulla che è quella compresa tra le aree rosa pari a 0,025 ciascuna e che delineano a destra e a sinistra un'area complessiva (e quindi un *p-value*) pari a $p = 0,05$. La statistica test dell'esempio sopra assume un valore pari a 1,46 (< 2 , dove 2 è il valore che assumerebbe in corrispondenza di un *p-value* di 0,05 complessivo) e si colloca perciò nell'area di non rifiuto dell'ipotesi nulla; pertanto, l'ipotesi nulla non può essere rifiutata e appare verosimile che la durata media dei procedimenti giudiziari in Italia sia pari a 480 giorni.

Il grafico sottostante (Grafico 4) mostra le aree colorate di rosa che delimitano la regione di rifiuto e accettazione della statistica test *t*, il puntino blu a destra che segna il valore osservato della statistica, i *degrees of freedom*⁷

⁶ Nella teoria delle probabilità la *distribuzione di Student*, o *t di Student*, è una distribuzione di probabilità continua che interviene nella stima della media di una popolazione che segue la distribuzione normale ma che ha varianza incognita e viene utilizzata nei test per la significatività della differenza tra due medie. Tale distribuzione di probabilità governa il rapporto tra due variabili aleatorie, una che segue una distribuzione normale e l'altra che segue una distribuzione chi quadrato.

⁷ I gradi di libertà di una variabile aleatoria o di una statistica in genere esprimono il numero minimo di dati sufficienti a valutare la quantità d'informazione contenuta nella

(o gradi di libertà) che sono pari a $40-1=39$ e il p-value che è maggiore di 0,05 indicando una impossibilità a rifiutare l'ipotesi nulla.

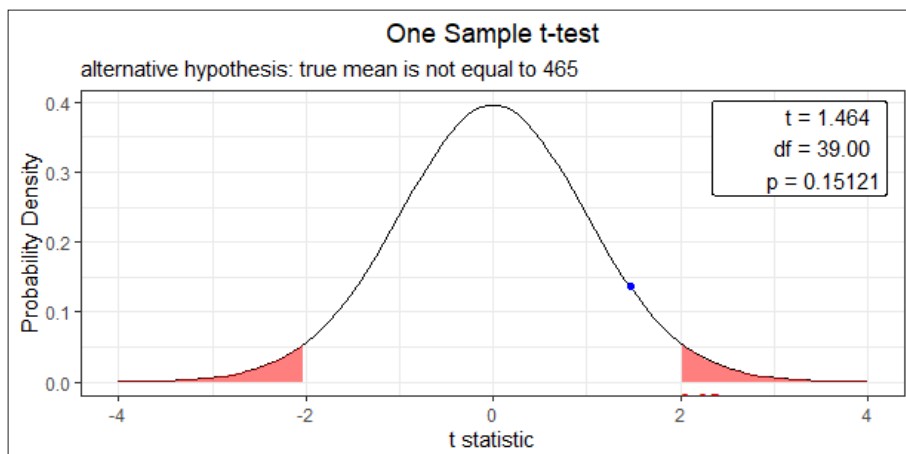


Grafico13

5. Cenni al campionamento e alle sue tecniche

Come è stato menzionato precedentemente l'inferenza statistica permette di estendere all'intera popolazione i risultati ottenuti su un campione.

Ci si potrebbe domandare tuttavia il perché è necessario considerare un campione di unità invece del totale di unità; i motivi sono molteplici e possono essere raggruppati in due tipologie: 1) necessità di un risparmio di costi e di tempi nelle varie fasi di progettazione, rilevazione, controllo e correzione ed elaborazione dei dati di indagine; 2) impossibilità di effettuare un'indagine totale, ad esempio perché la verifica renderebbe inutilizzabile il prodotto.

Tuttavia, la stima campionaria è comunque un'approssimazione statistica ed è soggetta ad errore.

Ovviamente l'attendibilità e la robustezza dei risultati dipendono anche dalla tipologia di campione realizzato che a sua volta dipende dalle caratteristiche della popolazione e dal fenomeno oggetto di studio. Di

statistica stessa. Infatti, se un dato non è indipendente, l'informazione che esso fornisce è già contenuta implicitamente negli altri dati (Wikipedia). Nell'esempio della statistica t di Student il parametro non indipendente è la media campionaria pertanto per calcolare i df è necessario sottrarre una unità dal numero totale di osservazioni, in questo caso pari a 40.

seguito vengono sintetizzate le tipologie più semplici di estrazione delle unità campionarie.

- *Campionamento casuale (random sampling)*. Ogni unità della popolazione ha la stessa probabilità di essere estratta nel campione.
- *Campionamento sistematico*. Le unità campionarie non vengono estratte con sorteggio ma vengono selezionate sistematicamente; la prima viene scelta casualmente, scegliendo un numero i da 1 a k , e le altre a partire dalla prima con un certo *passo* di campionamento k (il passo è dato dal rapporto tra numerosità della popolazione e numerosità campionaria (N/n)), la seconda unità pertanto sarà quella corrispondente al numero $i+k$, la terza a $i+2k$ etc.
- *Campionamento probabilistico (probabilistic sampling)*. Le unità della popolazione hanno probabilità diverse di essere estratte come campione.
- *Campionamento ragionato (judgmental sampling)*. La scelta del campione è affidata a chi cura l'indagine statistica in base alle sue conoscenze del fenomeno.
- *Campionamento per quote (convenience sampling)*. In questa selezione il campione ha la stessa struttura della popolazione nelle quote di uno o più attributi (es. sesso, istruzione, residenza).

Quando nella popolazione ci sono delle variabili correlate con la variabile oggetto di studio viene generalmente utilizzato un *campionamento stratificato* in cui si creano nella popolazione iniziale degli strati appunto in base alle variabili scelte e si effettua un campionamento casuale all'interno di ciascuno strato. Gli strati devono pertanto essere omogenei al proprio interno ed eterogenei (cioè differenti) tra di loro. Ad esempio, nell'ipotesi in cui si voglia studiare la capacità di definizione dei magistrati degli uffici giudiziari giudicanti, considerando che potrebbe essere oneroso raccogliere i dati per ciascun magistrato, si potrebbe effettuare un'indagine campionaria suddividendo le unità in strati a seconda di classi di sopravvenienze (es. da 0 a 200, da 200 a 500, etc..) dato che la quantità di fascicoli in carico a ciascun magistrato può influenzare il numero di definizioni. Gli strati saranno omogenei in quanto avranno magistrati simili per i fascicoli sopravvenuti e all'interno dei singoli strati verrà estratto casualmente un certo numero di magistrati e su questi verrà stimato il totale delle definizioni di tutti i magistrati.

Un altro tipo di *campionamento* molto usato è quello *a due o più stadi*, in cui il primo stadio di unità viene campionato casualmente e all'interno

di ciascuno stadio viene effettuato un altro campionamento casuale (ad esempio viene estratto casualmente un campione di uffici giudiziari e all'interno di ciascun ufficio viene estratto sempre casualmente un certo numero di magistrati).

Un caso particolare di campionamento a stadi è il *campionamento a grappolo* in cui al secondo stadio si considerano tutti i magistrati (2° stadio) all'interno degli uffici campionati (1° stadio).

In giustizia potrebbe sembrare inopportuno adottare un piano di campionamento se si considerano gli uffici giudiziari come unità primarie dato che essi non hanno una numerosità così elevata, tuttavia un'indagine campionaria avrebbe comunque senso qualora si volessero analizzare variabili relative a caratteristiche dei fascicoli, al lavoro dei magistrati presenti negli uffici, etc., effettuare rilevazioni su tutti gli uffici comporterebbe infatti un dispendio di costi e tempo consistente. Realizzare un campione statistico ha ancora più senso se le unità primarie non sono gli uffici ma il personale di magistratura per i motivi di numerosità sopra elencati.

Un aspetto importante legato alle indagini campionarie è l'*errore campionario* che indica l'influenza indotta dall'operazione di campionamento sulla varianza e sulla distorsione delle stime (cioè di quanto il valore stimato si distanzia dal valore vero nella popolazione obiettivo che è ignoto). L'errore campionario costituisce, insieme all'errore non campionario, l'errore totale. Mentre l'errore campionario non può essere presente in una indagine totale o censuaria, l'*errore non campionario* può essere anch'esso presente in una indagine campionaria e può essere commesso in ogni fase e sotto fase (i.e. progettazione iniziale, rilevazione, etc.) di una indagine statistica, consistendo ad esempio nel rifiuto a rispondere o nelle risposte errate da parte delle unità rispondenti. L'*errore non campionario* può essere di tre tipi:

- *Errore di copertura*: consiste nell'errore legato alla non corrispondenza fra la lista su cui si basa la selezione delle unità statistiche da contattare e la popolazione oggetto di indagine. Nello specifico si possono avere nell'indagine unità non appartenenti alla popolazione, in questo caso l'errore si dice di sovra copertura, oppure unità appartenenti alla popolazione ma non presenti negli archivi di base dell'indagine, in questo caso si parlerà di errore di sotto copertura. Questi ultimi sono più gravi e necessitano spesso di indagini di controllo affinché vengano individuati gli individui mancanti.
- *Errore di mancata risposta*: consiste nell'impossibilità o nel rifiuto a rispondere da parte delle unità statistiche contattate. Le mancate risposte possono essere totali se l'unità non partecipa all'indagine e

parziali se l'unità partecipa ma fornisce solo alcune risposte.

- *Errore di misurazione*: consiste nella differenza tra il vero valore della caratteristica da misurare su una data unità statistica e il valore effettivamente osservato all'indagine; tale differenza può essere causata dal rispondente stesso oppure dallo svolgimento delle fasi successive alla raccolta del dato (ad esempio errori durante la registrazione su supporto informatico o codifica di quesiti aperti).

La mancata risposta totale è spesso causata dal rifiuto di collaborare da parte dell'intervistato per motivi personali, dall'incapacità di comunicazione a causa per esempio di una diversa lingua, oppure si può avere una mancata risposta a causa del fatto che il questionario se cartaceo non è mai stato recapitato al diretto interessato.

Gli effetti della mancata risposta totale possono non essere importanti qualora le caratteristiche del non rispondente (ignote) siano simili a quelle della popolazione rispondente; in questo caso le variabili così come le unità sono *MCAR* (*missing completely at random*: la probabilità che un valore sia mancante è indipendente da qualsiasi altra variabile osservata o mancante). Se invece le unità che non partecipano all'indagine hanno caratteristiche particolari e diverse, si introduce una distorsione (*bias*) nei dati raccolti che si riproduce nelle stime.

Per illustrare un meccanismo MCAR in campo giudiziario, consideriamo ad esempio un questionario rivolto ad un campione di Tribunali in cui viene rilevata la percentuale di arretrato (pendenze ultratriennali) sulle pendenze totali nel contenzioso civile; i Tribunali che non hanno risposto all'indagine sono sia grandi che medi che piccoli in termini di organico della magistratura e con percentuali variabili di arretrato. Si può quindi ipotizzare che l'informazione mancante non dipenda dal numero di magistrati in organico, né dalla complessità dei fascicoli né da altre variabili osservate o meno.

Se il processo che causa dati mancanti non dipende dai valori mancanti di Y ma solo dai valori osservati di X e Y allora si dice *MAR* (*Missing at random*): in questo caso ci potrebbero essere differenze tra la variabile mancante e quella osservata ma esse possono essere spiegate da altre variabili osservate (la probabilità che un valore sia mancante è collegata ai valori osservati disponibili).

Nell'esempio precedente, la mancanza di alcuni uffici potrebbe essere spiegata dall'organico del personale amministrativo che deve essere, pertanto, una variabile osservata nel dataset.

Nel caso invece in cui la mancanza di un dato dipenda dalla tipologia

del dato stesso e da altre variabili, i valori mancanti non sono ignorabili e il meccanismo alla base si dice Missing Not At random (MNAR); ad esempio nel caso in cui si effettuasse un'indagine volta a indagare il reddito di un gruppo di lavoratori, il dato mancante potrebbe non essere casuale ma legato alla non propensione dell'individuo a fornire quel dato sensibile, con un effetto di auto selezione del campione dei rispondenti (bias di selezione).

Nel caso di mancata risposta totale, qualora sia necessario avere a disposizione la numerosità campionaria progettata inizialmente, può essere adottata una riponderazione delle unità del campione all'interno di post-strati creati individuando, laddove possibile, alcune caratteristiche della popolazione che influenzano la mancata risposta totale; si assume in questo caso che le unità campionate e non, siano omogenee nei post-strati.

La mancata risposta totale può avere effetti minori nel caso di indagini longitudinali in cui può mancare il record riferito a un dato anno ma le informazioni del record mancante si possono dedurre o supporre con una certa probabilità guardando alle informazioni della stessa unità riferite agli anni precedenti e/o successivi a disposizione nell'indagine.

Gli effetti di una mancata risposta parziale sono generalmente più contenuti rispetto a quelli di una mancata risposta totale; in questo caso si usano prevalentemente metodi di imputazione deterministica o probabilistica dell'informazione mancante, sfruttando altre informazioni presenti sulla stessa unità o derivanti da unità simili (vedi il seguito per i dettagli).

Le fasi di una indagine statistica

Il processo di progettazione e realizzazione di una indagine statistica è costituito da diverse fasi che vengono elencate di seguito (tratte da Fortini M., 2000).

- 1) *Obiettivi*: definizione del fenomeno oggetto di studio, individuando gli obiettivi globali e parziali in considerazione dei tempi e delle risorse disponibili.
- 2) *Disegno di indagine*: definizione della tipologia di indagine da adottare (ad esempio campionaria o totale) e in caso di indagine campionaria la *strategia di campionamento* (casuale semplice, stratificato, etc..) nonché la *tecnica di indagine* (intervista diretta faccia a faccia o telefonica, questionario postale auto compilato, etc..).

- 3) Messa a punto del *questionario*, della tipologia di domande e possibili risposte, senza il quale non è possibile applicare nessuna tecnica di indagine.
- 4) *Tempi e costi* che rappresentano due variabili che si influenzano reciprocamente e sono connesse fortemente alla qualità dell'informazione prodotta. Molte volte l'elemento di costo viene visto come un vincolo al quale la progettazione deve sottostare senza tener conto, in molti casi, delle risorse carenti che possono produrre errori nelle operazioni programmate e addirittura al fallimento degli obiettivi dell'indagine. D'altro canto, risorse illimitate possono indurre sprechi non sostenibili. È necessario pertanto stabilire il giusto trade-off tra tempi e costi per la buona riuscita dell'indagine e la garanzia della qualità della stessa nelle sue diverse dimensioni.
- 5) *Sistema dei controlli di qualità di una Indagine* consiste in un insieme di azioni predisposte nell'indagine finalizzate al trattamento dell'errore non campionario.
 - I controlli di qualità possono essere *preventivi*, ad esempio l'invio di una lettera di preavviso ai rispondenti o l'istituzione di un numero verde per le richieste di chiarimento possono ridurre le mancate risposte all'indagine.
 - Le azioni di controllo possono essere *in corso d'opera*: un esempio è l'uso dei programmi di registrazione controllata o l'applicazione di tecniche di identificazione automatica degli errori e una conseguente correzione per mezzo di un ritorno sul rispondente o altro.
 - Altre azioni di controllo sono invece *ex post* e sono pertanto di *valutazione* e hanno l'obiettivo di quantificare il livello di errore non campionario; esse consistono nella elaborazione dei dati durante l'esecuzione del processo di produzione e nella conduzione di indagini di controllo. Grazie a tali azioni è possibile avere una misura dell'errore tramite tassi di risposta e/o una misura della varianza semplice di risposta che si può ottenere tramite una re-intervista delle unità statistiche.
- 6) *Fasi operative*: che comprendono a loro volta le fasi di:
 - *Rilevazione*: le unità oggetto di indagine vengono contattate allo scopo di raccogliere l'informazione rilevante ai fini dello studio. Le modalità di contatto dipendono dalla tecnica di indagine adottata.

- *Registrazione dati su supporto magnetico*: consiste nel convertire le informazioni raccolte presso i rispondenti e disponibili su questionario cartaceo, su supporto informatico (CD Rom, DVD, etc.). Tale operazione di trascrizione solitamente avviene da parte di un operatore che digita sul computer esattamente ciò che legge sul questionario cartaceo. È possibile che il sistema avvisi qualora l'operatore commetta un errore di immissione del dato (come per esempio lasciare vuoto un campo obbligatorio), altre volte invece non avviene nessun controllo automatico, pertanto tale fase è soggetta a errori non campionari. Ovviamente qualora la rilevazione fosse di tipo CATI, intervista telefonica assistita da computer, o CAPI, intervista faccia a faccia assistita da computer, la fase di registrazione è assente in quanto le risposte vengono già inserite nel computer durante l'intervista.
- *Codifica dei quesiti aperti*: consiste nell'attività di trasposizione di informazioni pervenute sotto forma di linguaggio libero in un insieme finito di codici rispondenti ad una classificazione precostituita. Un esempio nel settore giudiziario è dato dalla codifica delle sezioni Sicid in codici informativi o in sezioni tabellari.
- *Revisione automatica o interattiva*: si definisce così la fase di individuazione e intervento di imputazione (cioè di sostituzione) sui valori mancanti o incongruenti nelle variabili rilevate, per mezzo di procedure informatizzate. Gli errori che vengono corretti sono essenzialmente quelli derivanti da violazioni di regole logico-formali, dette regole di compatibilità, relative ai limiti imposti sul campo di variazione delle singole variabili, alle relazioni intercorrenti tra le variabili. La modalità di revisione può essere anche interattiva: in questo caso viene automatizzata la sola fase di individuazione dell'errore, lasciando ad un operatore il compito di eseguire le correzioni al terminale. Gli errori che possono essere commessi in questa fase sono di due tipi, *errori sistematici* per i quali si può supporre che il valore da imputare sia unico, e gli *errori casuali*, per cui si può supporre che le possibili correzioni effettuabili siano varie.

I metodi di imputazione sono diversi ed hanno tutti l'obiettivo di apportare modifiche al dataset influenzando il meno possibile sulle stime di interesse.

In generale per effettuare imputazioni di valori mancanti o

incongrui si può ricontattare l'unità statistica in questione al fine di acquisire il valore vero, oppure si possono considerare informazioni relative a periodi precedenti sempre per quella stessa unità, oppure ancora si può ricorrere alla sostituzione dell'informazione con quella relativa ad unità simili a quella unità statistica per cui si è registrato l'errore.

Le procedure di imputazione *deterministica* correggono gli errori sistematici mettendo in atto una serie di regole appunto deterministiche del tipo SE-ALLORA-CORREZIONE.

Tra i metodi di imputazione deterministica vi sono:

- *L'imputazione con medie condizionate*: le unità vengono raggruppate in classi omogenee rispetto ad alcune variabili di interesse, si calcola la media sui rispondenti della variabile con non risposte o risposte errate entro ogni classe e si imputa il valore dell'unità mancante o con risposta incongrua con tale media che sarà diversa per ogni classe.
- *Imputazione con regressione*: consiste nell'applicare dei modelli di regressione nel caso in cui si ipotizzi una relazione esistente tra alcune variabili esplicative e la variabile risposta che è quella per cui si intende effettuare l'imputazione. I parametri stimati dal modello messo a punto sui rispondenti permettono di predire i valori della variabile risposta anche sulle unità non rispondenti per quella variabile.
- *Imputazione del donatore più vicino*: in questo caso si sostituisce un dato mancante con il corrispondente più vicino per i valori di alcune variabili ausiliarie; tale valore viene identificato adottando una funzione di distanza tra l'unità mancante e le altre unità rispondenti.

L'imputazione *probabilistica* va invece a correggere gli errori casuali, ha costi molto bassi e si applica generalmente a grandi moli di dati omogenei per alcune caratteristiche.

Esempi di imputazione di tipo probabilistica sono:

- *L'imputazione multipla*: in questo caso si va a generare un numero m di valori ($m > 2$) da imputare per ogni dato mancante in modo tale che i risultati delle m analisi distinte vengano poi combinati e si tenga conto dell'incertezza causata dai dati mancanti.
- *L'imputazione con regressione casuale*, in cui i valori mancanti vengono stimati con la regressione considerando anche la componente residuale del modello.

- 7) *Elaborazioni statistiche*: comprende l'analisi descrittiva ed esplorativa delle caratteristiche del fenomeno; in questa fase si applicano le metodologie statistiche per descrivere e sintetizzare i dati.
- 8) *Validazione*: si intende il processo attraverso il quale si valuta se l'informazione può essere considerata consona alle finalità per le quali è stata prodotta. La validazione può essere definita come l'insieme delle operazioni attraverso le quali si valuta lo scarto esistente tra gli obiettivi di qualità programmati in sede di progettazione dell'indagine statistica o amministrativa e i risultati effettivamente conseguiti. Pertanto gli obiettivi di qualità devono essere definiti in fase di progettazione e devono poter essere espressi in termini misurabili attraverso indicatori di qualità (stima indiretta della qualità dei dati) oppure attraverso un metodo meno economico ma efficace basato sulla stima diretta delle componenti di errore totale (ad esempio attraverso una re-intervista sul campione dei rispondenti e valutando poi la variabilità delle risposte sugli stessi individui nelle diverse occasioni di indagine).
- 9) *Diffusione*: consiste appunto nella diffusione della produzione di informazione statistica. Tale fase può sembrare accessoria ma è al contrario di grande importanza in quanto incide su alcune dimensioni della qualità dei dati (di cui verrà discusso in dettaglio nella sezione successiva). L'informazione ad esempio può essere diffusa in modalità cartacea, tramite CD-Rom se si tratta di file di dati individuali in cui vengono adottate alcune tecniche per garantire l'anonimato delle unità, in rete nel caso di dati più aggregati, tramite interrogazioni su dati elementari in cui il dato elementare non è visibile all'utente ma solo il risultato finale dell'interrogazione. In ogni caso la diffusione presuppone la garanzia della riservatezza del dato.

6. La qualità dei dati

La qualità dei dati viene normalmente definita come la capacità di soddisfare le esigenze informative ed è frutto della corretta applicazione di processi, di tecnologie accurate e personale qualificato.

Sia quando il dato è raccolto attraverso indagini appositamente predisposte sia quando deriva da banche dati amministrative occorre misurare e valutare la qualità del dato.

Come è già stato illustrato in precedenza, nella fase di progettazione e realizzazione di una indagine vi sono diverse regole e azioni finalizzate al trattamento dell'errore non campionario e al miglioramento della qualità dei dati finali.

Anche quando le elaborazioni statistiche provengono da fonte amministrativa, come spesso avviene per i dati giudiziari relativi ai procedimenti civili e penali nei quali si utilizzano i registri ufficiali in uso agli uffici, occorre valutare la qualità del dato e tener conto delle mancate annotazioni, degli errori presenti nei registri, della corretta interpretazione delle informazioni presenti e delle finalità per cui le stesse vengono raccolte.

La misurazione della qualità del dato avviene attraverso diverse dimensioni e metriche di seguito descritte:

- *Rilevanza*: è la capacità dell'informazione di soddisfare le esigenze conoscitive degli utenti, che possono essere anche organi di governo centrali o locali in base agli obiettivi dell'indagine prefissati.
- *Accuratezza*: è il grado di corrispondenza tra la stima ottenuta dall'indagine e il valore vero (ignoto) della caratteristica in oggetto nella popolazione obiettivo. Essa può essere misurata attraverso la stima dell'errore totale.
- *Tempestività*: è l'intervallo di tempo che intercorre fra il momento della diffusione dell'informazione prodotta e l'epoca di riferimento della stessa. La tempestività dell'informazione è strettamente legata ai costi dell'indagine ed è inversamente correlata all'accuratezza: una maggiore accuratezza normalmente richiede tempi più lunghi.
- *Accessibilità e chiarezza*: corrisponde alla semplicità per l'utente di reperire, acquisire e comprendere l'informazione disponibile in relazione alle proprie finalità.
- *Confrontabilità*: tale dimensione attiene alla possibilità di paragonare nel tempo e nello spazio le caratteristiche riguardanti il fenomeno di interesse. Si possono infatti verificare nel tempo delle modifiche concettuali nonché delle definizioni di alcune variabili.
- *Coerenza*: consiste nella possibilità di confrontare le statistiche derivanti da fonti diverse basate su definizioni, classificazioni e standard metodologici comuni.
- *Completezza*: è la capacità dei singoli processi di integrarsi per fornire un quadro informativo soddisfacente del dominio di interesse (i.e. flussi dei procedimenti per sede, distretto, regione, area geografica).
- *Tutela della riservatezza*: corrisponde alla garanzia dell'anonimato per ciascuno dei soggetti (individui, famiglie, etc.) che hanno fornito informazioni utili alla conduzione dell'indagine.

La valutazione della qualità del dato deve precedere l'elaborazione statistica e può accompagnarsi alla comparazione dei dati disponibili provenienti da diverse fonti al fine di identificare le eventuali problematiche presenti. Occorre quindi determinare e intraprendere le azioni correttive necessarie ad una più corretta rappresentazione del fenomeno oggetto di indagine.

7. Cenni alle principali tecniche di analisi multivariata

Tra le principali tecniche di analisi multivariata vi sono la cluster analysis, l'analisi fattoriale, l'analisi discriminante, l'analisi delle componenti principali e delle corrispondenze (Vitali O., 1993).

- *La cluster analysis* è una tecnica non supervisionata (che non può prendere esempio da nessun altro sottoinsieme di dati per fornire l'outcome desiderato) che ha l'obiettivo di creare dei gruppi di unità (soggetti o oggetti) tali che siano più omogenei al loro interno e più eterogenei tra di loro in base a determinate misurazioni della somiglianza tra le unità; l'omogeneità è misurata sulla base di alcune variabili di interesse che possono essere sia quantitative sia qualitative: ad esempio gruppi di studenti che hanno votazioni simili nelle materie umanistiche, gruppi di pazienti con effetti simili del trattamento, magistrati con un numero simile di procedimenti in carico etc.

Tra le principali tecniche di clustering si distinguono quelle basate sugli *algoritmi gerarchici*, rappresentabili anche graficamente tramite una struttura ad albero, e quelle basate su *algoritmi non gerarchici* in cui il numero di gruppi in cui raggruppare le osservazioni viene deciso a priori (i metodi dei k-centroidi o delle k-medie sono tra i più utilizzati in questo ambito). Nel caso di tecniche gerarchiche invece l'obiettivo di creare gruppi omogenei si fonde con quello di trovare il numero ottimale di cluster che permetta una rappresentazione del fenomeno sintetica e al tempo stesso interpretabile.

- *L'analisi discriminante* fu introdotta da R. A. Fisher nel 1936, che la utilizzò per attribuire dei reperti fossili ad uno di due gruppi: scimmie e umanoidi. Tale analisi ha l'obiettivo di individuare l'insieme di variabili che permette appunto di "discriminare" cioè distinguere le unità statistiche per poi associarle a cluster noti a priori, attraverso una regola di decisione (o regola di classificazione).
- *L'analisi in componenti principali* (ACP) è una tecnica di statistica

multivariata volta alla riduzione della *dimensionalità* del fenomeno garantendo la stessa *variabilità spiegata*. Tale analisi si basa su un insieme di variabili dette appunto componenti principali che nascono come combinazione lineare di un insieme più numeroso di variabili iniziali; le componenti principali hanno la caratteristica di rappresentare in una dimensione inferiore (essendo in numero inferiore) la stessa informazione o comunque gran parte dell'informazione prodotta dalle variabili originarie.

- L'*analisi delle corrispondenze multiple* (ACM), ha lo stesso obiettivo delle componenti principali ma lavora con sole variabili qualitative; lo scopo è individuare dei fattori soggiacenti alla struttura originaria dei dati che riassumano le relazioni di interdipendenza tra le variabili originarie. L'ACM si basa su tabelle di contingenza in cui righe e colonne sono associate tra loro e, non facendo nessuna ipotesi distributiva, permette di visualizzare in modo semplice e deduttivo le modalità delle variabili più associate tra loro che definiscono quindi i diversi fattori (assi cartesiani).
- L'*analisi fattoriale* ha come obiettivo quello di individuare variabili latenti e cioè non direttamente osservabili e misurabili a partire da una serie di variabili osservate; esempi di variabili latenti sono l'abilità nelle scienze umanistiche, la qualità della vita, etc. Proprio perché le variabili individuate sono latenti, esse possono aiutare a spiegare fenomeni complessivi latenti anch'essi, cioè non direttamente osservabili per natura, come ad esempio l'effettiva efficacia di un questionario indicando quindi se esso misura effettivamente ciò per cui è stato progettato. Vengono pertanto individuati dei fattori latenti anch'essi combinazioni di variabili originarie che hanno dei fattori comuni e dei fattori specifici; tali fattori possono essere proiettati su un grafico cartesiano per una migliore rappresentazione del fenomeno e una più semplice lettura.

8. Cenni alle Serie Storiche

Per serie storica si intende un insieme di valori di una stessa variabile rilevati nel tempo, cioè in istanti di tempo diversi, e di solito tra loro equidistanti.

Ad esempio, in giustizia, il numero di magistrati assunti anno per anno o la distribuzione annuale delle pendenze possono essere classificati come

serie storiche.

In una serie storica è presente una componente di fondo chiamata TREND della serie, insieme ad un'altra caratteristica che è quella relativa alla STAGIONALITÀ (Dagum, 2002).

Nel caso di trend accentuato, con la presenza contemporanea di valori molto piccoli e molto grandi, si preferisce sostituire ai valori i logaritmi; la trasformazione logaritmica permette di ingrandire le differenze piccole e diminuire le differenze grandi, evidenziando così un trend lineare nei logaritmi ed esponenziale quindi nei valori originari.

Per stagionalità si intende una componente oscillatoria, cioè un fattore che introduce variabilità, che fa ripetere un comportamento analogo nello stesso periodo di tempo (ad esempio un mese, un trimestre, etc.) di anni successivi. Basti pensare al calo della produttività di molte aziende nel periodo di agosto dovuto alle ferie dei dipendenti. Al fine di confrontare dati relativi a periodi diversi è necessario eliminare la stagionalità.

Uno dei metodi empirici più utilizzati per scomporre una serie storica e soprattutto per depurarla dalla sua componente stagionale, consiste nell'applicazione alla serie di una *Media Mobile*. La media mobile è una media aritmetica che viene applicata su un certo numero di termini della serie; ad esempio una media mobile a 3 termini consiste nella media aritmetica dei primi 3 termini: secondo, terzo e quarto termine; terzo, quarto e quinto termine e così via. Le medie mobili diventano a questo punto i nuovi dati e si ottiene l'effetto *smoothing* tradotto in italiano di "lisciamento" degli spigoli dovuti alla stagionalità.

Una volta che la stagionalità viene depurata dalla serie, è possibile calcolare indici di posizione come la media e la mediana e indici di variabilità come la varianza; tali indicatori possono essere calcolati sia in modo cross-section e cioè prendendo un anno o un periodo di riferimento, sia longitudinalmente e cioè su tutti i periodi o cmq su più periodi, anni o mesi etc., della serie storica. In quest'ultimo caso, oltre alla variabilità tra differenti individui viene considerata anche la variabilità tra gli stessi individui ai vari tempi.

Un'analisi in serie storica è utile non solo al fine di effettuare degli studi osservazionali sull'andamento di un fenomeno e sulla dinamicità delle caratteristiche e dei comportamenti degli individui ma anche al fine di effettuare delle previsioni.

La realizzazione di una previsione tramite l'analisi delle serie storiche non può prescindere dalle fasi alla base di qualsiasi ricerca statistica; esse consistono nell'analizzare approfonditamente il problema e stabilire gli obiettivi di studio, nella raccolta dei dati e nell'analisi della loro qualità con

conseguente controllo e correzione dei dati, nell'analisi esplorativa della serie storica al fine di tracciare il suo andamento e la sua struttura, nella scelta e applicazione del modello statistico e nella valutazione della sua bontà di adattamento, infine è necessaria una verifica e valutazione delle capacità predittive del modello stesso.

Nella fase esplorativa è necessario, anche attraverso l'uso di grafici e indici descrittivi, analizzare l'eventuale presenza del trend e della stagionalità, poi una volta scelto il modello è necessario valutare la bontà dello stesso e delle previsioni che si vogliono effettuare. Anche nelle Serie Storiche è possibile poi riscontrare i valori mancanti che è possibile trattare in modi diversi: se la percentuale di valori mancanti è elevata allora non si può considerare la variabile, non si deve invece eliminare mai l'osservazione a meno che non abbia valori mancanti su tutte le variabili, e in quest'ultimo caso è necessario sostituire l'informazione mancante scegliendo opportuni metodi di imputazione. Ad esempio, il metodo della media consiste nello scegliere una misura di centralità (media aritmetica, mediana, etc..) calcolata su tutta la serie storica, il valore trovato si sostituisce ai valori mancanti; tale tecnica ha il vantaggio di essere molto semplice ma dall'altra parte di ignorare l'ordinamento dei dati e produrre una sottostima della periodicità a causa dell'introduzione di un valore stabile. Un altro metodo di imputazione molto utilizzato si basa sull'interpolazione lineare tra due valori contigui nell'ipotesi che ci sia un andamento lineare rispetto al tempo nei pressi del valore mancante; tale tecnica ha il vantaggio di essere semplice e di rispettare l'ordinamento temporale. Esistono anche altri metodi di imputazione come il metodo di Buck non trattati in questa sede.

Il grafico seguente riporta un esempio di serie storica relativa all'età media dei vincitori di concorso in magistratura per genere e anno di entrata in magistratura.

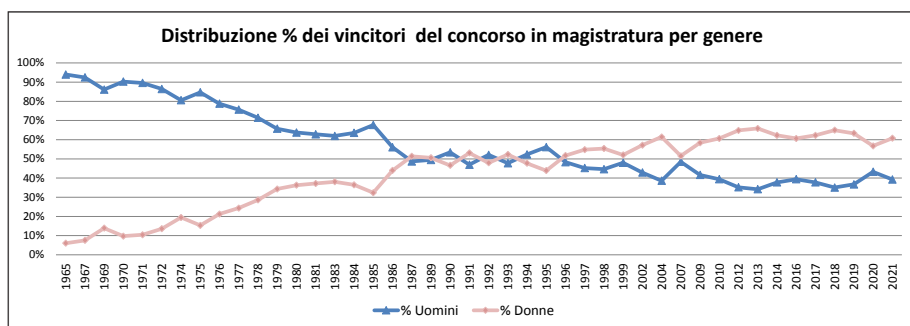


Grafico 14

9. *Software per la Statistica*

Di seguito vengono riportati i software più comunemente utilizzati per le elaborazioni statistiche in vari settori di ricerca e quindi anche in quello giudiziario.

- ✓ *SAS* (Statistical Analysis System) è una piattaforma che consente di analizzare dati statistici attraverso dei pacchetti integrati che consentono l'inserimento, la ricerca e la gestione dei dati, la generazione di report, grafici, analisi statistica e matematica per la pianificazione, la previsione e il supporto alle decisioni.
- ✓ *Spss* (Statistical Package for Social Science) nasce nell'ambito delle scienze sociali e nel tempo sono state sviluppate numerose funzioni di statistica, tale da renderlo utile per qualsiasi ambito come quello delle scienze mediche/epidemiologiche, economiche, demografiche, etc. La prima versione di SPSS è stata realizzata nel 1968; oggi è uno strumento di analisi che permette di svolgere numerose operazioni come accedere, controllare e analizzare rapidamente basi dati abbastanza voluminose. Il programma è dotato di una interfaccia grafica e permette, oltre ad effettuare operazioni e modifiche di dataset, anche di condurre analisi statistiche descrittive e multivariate.
- ✓ *R* è un linguaggio di programmazione e un ambiente di sviluppo specifico per l'analisi statistica dei dati. Venne scritto inizialmente dal matematico e statistico canadese Robert Gentleman, e dallo statistico neozelandese Ross Ihaka; dal 1997 ai due sviluppatori si è aggiunto un gruppo di programmatori tra cui John Chambers con diritti di scrittura sul progetto. *R* è un software *freeware*, è possibile pertanto scaricarlo liberamente nelle sue versioni più o meno aggiornate ed è disponibile per diversi sistemi operativi (ad esempio Unix, GNU/Linux, MacOS, Microsoft Windows). La popolarità di tale software è dovuta anche all'ampia disponibilità di moduli organizzati in un apposito sito chiamato CRAN (Comprehensive R Archive Network); tali moduli permettono l'applicazione di pacchetti con funzioni e modelli statistici aggiuntivi ma anche la comunicazione con database o con sistemi GIS. Il linguaggio *R* è fornito con una interfaccia a riga di comando ma sono disponibili anche diverse interfacce grafiche. A partire dal

2009 viene pubblicato *The R Journal*, con una serie di articoli che coprono argomenti che dovrebbero essere di interesse di utenti o sviluppatori di R; the R journal ha sostituito la rivista elettronica R news che veniva pubblicata dal 2001 con cadenza irregolare e relativa all'utilizzo di alcuni moduli di R per alcuni problemi di statistica applicata. Il software R contiene pacchetti funzionali alla gestione di dataset, al loro join, all'applicazione di operazioni di modifica dei dataset, alla creazione di grafici ma anche, come accennato sopra, all'applicazione di modelli statistici. R inoltre, essendo un linguaggio di programmazione permette la messa a punto di algoritmi che possono appunto includere funzioni statistiche di stima al loro interno.

- ✓ *Stata* è un software statistico ad ampio spettro di applicazione, molto usato in ambito epidemiologico e medico ma anche statistico (e statistico giudiziario), funzionale sia alla gestione e alla elaborazione di dati che alla rappresentazione grafica, all'analisi statistica, alle simulazioni e alla programmazione personalizzata. Sono stati sviluppati molti pacchetti per le analisi statistiche che possono essere caricati nella libreria di Stata ed essere utilizzati in Stata.

Capitolo VI

La statistica nell'organizzazione degli uffici giudiziari

Gli uffici giudiziari necessitano di dati statistici per l'organizzazione delle attività. In questo capitolo si farà cenno alle principali attività organizzative nelle quali la statistica giudiziaria riveste un ruolo fondamentale. In particolare, si illustreranno: il documento organizzativo principale che descrive il funzionamento di un ufficio giudiziario, ovvero le tabelle di organizzazione degli uffici giudicanti e il progetto organizzativo degli uffici requirenti; il programma di gestione volto allo smaltimento dell'arretrato; gli obiettivi previsti dal Piano Nazionale di Ripresa e Resilienza e all'Ufficio del processo; la digitalizzazione; la valutazione di professionalità del magistrato.

1. Le Tabelle di organizzazione degli uffici giudiziari giudicanti

La struttura organizzativa degli uffici giudicanti (Tribunali e Corti d'Appello, Tribunali e Uffici di Sorveglianza, Tribunali per i minorenni, Corte di Cassazione) deve essere nota a tutti i cittadini nel rispetto del principio di precostituzione del giudice naturale di cui all'art. 25 della Costituzione: il giudice che si occuperà di un determinato processo deve essere individuato sulla base di criteri oggettivi conoscibili a priori.

Al fine del rispetto di tale principio, il CSM indica in una apposita circolare le modalità di predisposizione del documento organizzativo (progetto tabellare) da parte del capo dell'ufficio giudiziario; tale circolare viene aggiornata con periodicità quadriennale e avvia il processo di formazione delle cosiddette tabelle.

La proposta tabellare presentata dal capo dell'ufficio si compone del documento organizzativo generale (Dog) e del progetto tabellare.

Il "sistema tabellare" rappresenta il cardine della struttura organizzativa degli uffici giudicanti; le tabelle riportano:

- la ripartizione di ogni ufficio in settori (civile e penale) ed in sezioni,
- la distribuzione dei singoli magistrati in sezione,
- le concrete attribuzioni di ciascun magistrato,
- i criteri di assegnazione degli affari giudiziari,
- la designazione dei magistrati ai quali è attribuito il compito di direzione di una sezione o di coordinamento della sezione o dell'ufficio Gip/Gup,
- l'assegnazione alle sezioni dei presidenti,
- il calendario delle udienze,
- i criteri per la sostituzione del giudice astenuto, ricusato o impedito.

Nel documento organizzativo generale vengono chiarite le ragioni delle scelte organizzative; esso viene predisposto partendo dalla analisi dello stato dei servizi, dei carichi di lavoro e dei flussi delle pendenze risultanti dai dati statistici per ufficio e, ove esistano, per ciascuna delle diverse sezioni dell'ufficio; contiene l'indicazione delle eventuali cause di disfunzione rilevate a seguito dei monitoraggi eseguiti all'interno di ciascun settore o sezione.

In tal modo le determinazioni organizzative assunte dai dirigenti degli uffici giudicanti possono essere conosciute e verificate da tutti gli attori del procedimento tabellare e, da ultimo, dal Consiglio Giudiziario e poi dal CSM che deve valutare se approvarle o meno.

L'ultima Circolare *sulle tabelle* è disponibile sul sito www.csm.it ed è stata deliberata prevedendo ulteriori linee direttrici:

- Razionalizzazione della disciplina (benessere organizzativo, tutela della maternità, della genitorialità e dei compiti di cura);
- Procedimento di formazione: coinvolgimento dell'avvocatura e del Procuratore della Repubblica;
- Responsabilizzazione dei magistrati con funzioni direttive e semi-direttive (anche al fine di evitare il conferimento di incarichi di collaborazione sovrapponibili alle funzioni semidirettive);
- Trasparenza delle decisioni;
- Valorizzazione della pari dignità delle funzioni e dell'equa distribuzione degli affari;
- Valorizzazione dell'ufficio per il processo;
- Miglioramento dei modelli organizzativi;
- Adeguamento della disciplina alle esigenze specifiche della Corte di Cassazione.

Nell'ambito della procedura di formazione del progetto tabellare è importante soffermarsi su alcuni aspetti ed articoli della circolare consiliare che incidono nella rappresentazione statistica dei flussi.

Innanzitutto, come è già stato detto, l'ufficio è suddiviso in settori: civile e penale. In molti uffici viene effettuata una ulteriore distinzione all'interno del settore civile tra lavoro e previdenza e i restanti affari del contenzioso, volontaria giurisdizione, fallimenti ed esecuzioni civili.

Nel progetto tabellare occorre determinare il numero delle sezioni e dei magistrati assegnati a ciascuna sezione tenendo conto delle specifiche esigenze di ciascun ufficio, allo scopo di assicurare l'efficienza dell'organizzazione. L'articolo 49 della circolare precisa che la costituzione di ogni sezione del tribunale ordinario, fatta eccezione per la sezione Gip/Gup, richiede l'assegnazione di non meno di cinque giudici, escluso il Presidente di sezione, ai sensi dell'articolo 46, co. 5, regio decreto 30 gennaio 1941, n. 12.

Il numero di sezioni è funzione, quindi, della dimensione dell'ufficio e del numero di Presidenti di sezione previsto. Gli uffici piccoli normalmente sono organizzati in un'unica sezione promiscua oppure in una sezione civile e una sezione penale.

Nell'organizzazione degli uffici va però favorito lo sviluppo di competenze specialistiche per materie omogenee e predeterminate, anche all'interno delle singole sezioni. L'art. 56 della circolare precisa che la costituzione di sezioni specializzate rappresenta il modello organizzativo più adeguato a garantire professionalità maggiormente qualificate, tale da rendere più efficace e celere la risposta all'istanza di giurisdizione.

La specializzazione va preferita sia negli uffici organizzati in più sezioni civili e penali, sia negli uffici più piccoli attraverso la creazione di gruppi di lavoro all'interno della stessa sezione.

Presso i tribunali che hanno un organico di giudici del lavoro non inferiore a cinque magistrati può essere istituita una autonoma sezione civile per la trattazione in via esclusiva della materia lavoro.

In tutti i casi nei quali il numero dei magistrati assegnati all'ufficio e il numero degli affari giudiziari lo consente, è istituita una sezione per la trattazione dei procedimenti relativi alle persone e ai rapporti di famiglia, con eventuale ulteriore specializzazione per la trattazione delle materie di competenza del giudice tutelare.

Per le materie previste dall'articolo 3 del decreto legislativo 27 giugno 2003, n. 168 (così come sostituito dall'articolo 2, comma 1, lett. d), del decreto legge 24 gennaio 2012, n. 1, convertito, con modificazioni, nella legge 24 marzo 2012, n. 27), operano le sezioni specializzate in materia di impresa presso i Tribunali e le Corti d'appello di Ancona, Bari, Bologna,

Bolzano, Brescia, Cagliari, Campobasso, Catania, Catanzaro, Firenze, Genova, L'Aquila, Milano, Napoli, Palermo, Perugia, Potenza, Roma, Torino, Trento, Trieste e Venezia.

Per le materie previste dall'articolo 3 del decreto-legge 17 febbraio 2017, n. 13, convertito, con modificazioni, in legge 13 aprile 2017, n. 46 e successive modifiche, operano le sezioni specializzate presso i Tribunali distrettuali in materia di immigrazione, protezione internazionale e libera circolazione dei cittadini dell'Unione Europea.

La sezione Gip/Gup è istituita in tutti i tribunali organizzati in più di due sezioni; negli altri uffici vengono individuati i magistrati addetti a tali funzioni.

Il capo dell'ufficio deve inserire informaticamente le informazioni organizzative, indicando le sezioni e le competenze di ciascuna sezione attraverso la selezione delle materie e dei reati all'interno di un raggruppamento già previsto dal sistema informativo (denominato "catalogo materie" e contenente i singoli codici oggetto del civile e la tabella dei reati del penale).

Tali informazioni informatiche consentono, quindi, la distinzione dei flussi dei procedimenti complessivi di ciascun ufficio giudiziario tra sezioni tabellari, nonché la estrapolazione di dati statistici più dettagliati.

2. L'organizzazione degli uffici di Procura

Gli aspetti organizzativi sono una componente essenziale funzionale all'attuazione dei beni giuridici costituzionalmente rilevanti dell'esercizio imparziale dell'azione penale, della celerità del procedimento, dell'effettività dell'azione penale e del diritto di difesa, anche negli uffici requirenti.

Con delibera del 16 dicembre 2020 è stata rivisitata la Circolare sull'organizzazione degli uffici requirenti, partendo dai principi costituzionali declinati negli artt. 105, 107, 108 e 112 Cost., dalle delibere del 2007 e del 2009 e dall'analisi delle problematiche emerse nell'applicazione della circolare precedente.

Premesso che la procedura di formazione del progetto organizzativo della Procura è simile a quello degli uffici giudicanti, gli aspetti su cui si è concentrata la circolare consiliare riguardano principalmente:

- L'interpello come strumento per l'assegnazione dei magistrati ai gruppi di lavoro, per l'assegnazione degli incarichi di coordinamento ai Procuratori aggiunti e per gli incarichi di collaborazione;

- Il metodo partecipato per l'adozione del progetto organizzativo. Nello specifico, viene delineata la centralità dell'assemblea dei magistrati dell'ufficio come momento preliminare e necessario di condivisione ai fini dell'adozione del progetto organizzativo. Inoltre, altro essenziale aspetto è il confronto con il dirigente del corrispondente ufficio giudicante, onde favorire l'adozione di provvedimenti condivisi, attesa la interconnessione fra le attività dei due uffici.
- La regolamentazione delle funzioni semidirettive; la previsione dell'obbligo di svolgimento di una quota di lavoro "giudiziario" in senso stretto per i procuratori aggiunti; il divieto di esonero per i magistrati con funzioni di collaborazione non titolari di funzioni semidirettive.
- La trasparenza e imparzialità nell'attività del dirigente.
- La previsione di espressi criteri per l'assegnazione degli affari, la coassegnazione e l'autoassegnazione; la necessità di specifica motivazione in caso di deroga a criteri automatici di assegnazione.
- I criteri di assegnazione dei procedimenti all'interno dei gruppi di lavoro e le regole dell'auto-assegnazione tenendo conto, da un lato, dell'esigenza di salvaguardare l'attribuzione al Procuratore della Repubblica della titolarità dell'esercizio dell'azione penale, che trova il proprio fondamento nella normativa primaria; dall'altro, dell'esigenza concomitante di una chiara regolamentazione di tutte le circostanze in cui siano derogati i criteri di attribuzione dei procedimenti individuati dal medesimo Procuratore della Repubblica attraverso il suo progetto organizzativo, secondo regole trasparenti ispirate al buon andamento dell'azione giudiziaria, all'imparzialità e all'efficienza.
- L'individuazione di criteri predefiniti per l'assegnazione dei magistrati alle DDA e per le coassegnazioni dei procedimenti di competenza della DDA.
- La definizione del ruolo dei Consigli giudiziari nell'esame dei progetti organizzativi.
- La regolamentazione dell'organizzazione della DNAA (Direzione Nazionale Antimafia e Antiterrorismo), anche con riferimento alla assegnazione dei magistrati ai gruppi di lavoro, alla assegnazione degli affari, al ruolo dei Procuratori aggiunti.

3. Programmi di gestione

Programmi di gestione civili

L'art. 37 co. 1 del D.L. n. 98/2011 convertito con l. n. 111/2011 ha previsto l'obbligo per i dirigenti degli uffici giudiziari di redigere annualmente il programma di gestione degli affari civili.

Il Programma di gestione ha lo scopo di monitorare l'attività degli uffici giudiziari giudicanti di primo e secondo grado fornendo allo stesso tempo uno strumento di supporto agli uffici stessi per rendere efficace la programmazione della loro attività.

La norma prevede che il capo dell'ufficio giudiziario con il programma determina sia gli obiettivi di riduzione della durata dei procedimenti concretamente raggiungibili nell'anno in corso, sia gli obiettivi di rendimento dell'ufficio, con l'indicazione, per ciascuna sezione o, in mancanza, per ciascun magistrato, dei risultati attesi sulla base dell'accertamento dei dati relativi al quadriennio precedente e di quanto indicato nel programma di cui all'articolo 4 del decreto legislativo 25 luglio 2006, n. 240, e, comunque, nei limiti dei carichi esigibili di lavoro individuati dai competenti organi di autogoverno, nonché l'ordine di priorità nella trattazione dei procedimenti pendenti, individuati secondo criteri oggettivi ed omogenei che tengano conto della durata della causa, anche con riferimento agli eventuali gradi di giudizio precedenti nonché della natura e del valore della stessa.

Dal 2011 diverse circolari del CSM, disponibili on line, hanno fornito indicazioni agli uffici sulle modalità di predisposizione di tale documento organizzativo che deve essere depositato entro il 31 gennaio di ogni anno.

Il programma di gestione viene attualmente predisposto sulla base di un apposito format predisposto dal CSM. Una parte importante del format riguarda l'indicazione dell'*obiettivo di smaltimento*, ovvero del numero di fascicoli pendenti al 30.6. dell'anno in corso - ultratriennali nel caso di Tribunali o ultrabiennali nel caso di Corti di Appello - che l'ufficio intende smaltire entro il 31 dicembre dell'anno successivo. Una seconda sezione riguarda l'indicazione del *carico esigibile* che rappresenta la "capacità di lavoro" dei magistrati che fisiologicamente consente di coniugare qualità e quantità del lavoro in un dato periodo di tempo, da individuarsi alla luce della concreta situazione dell'ufficio presso il quale prestano servizio i magistrati. Il Carico Esigibile rappresenta il numero di procedimenti che il capo dell'ufficio si aspetta siano definiti in un anno solare da un magistrato addetto a tempo pieno nell'ufficio, in una data sezione e/o in una data macromateria. Si tratta quindi di un dato pro-capite, per ufficio, per sezione,

per macromateria e/o per macromateria all'interno della sezione, funzionale per la valutazione della capacità di programmazione dello smaltimento delle pendenze da parte dell'ufficio. Esso viene determinato partendo dal numero di procedimenti mediamente definiti nel quadriennio precedente e tenendo conto del numero di unità di magistrati Full Time Equivalent (FTE) addetti a ciascun settore o sezione o macromateria.

Le altre sezioni del format civile riguardano gli *obiettivi di qualità*, la *valorizzazione dei criteri di priorità* e le *modalità di monitoraggio attuate del programma di gestione*.

Il capo dell'ufficio predispone il programma attraverso una procedura partecipata, utilizzando i dati statistici messi a disposizione da parte dell'Ufficio Statistico del CSM ed estratti con gli strumenti disponibili, ovvero con il datawarehouse ministeriale dal quale vengono estratti i flussi complessivi e con il database consiliare che consente la distinzione di tali flussi in funzione dell'organizzazione tabellare degli uffici in sezioni.

Il programma di gestione viene depositato e successivamente analizzato dal Consiglio Giudiziario e dal CSM per la sua approvazione.

La recente modifica normativa ha previsto che il programma di gestione tenga conto di quanto indicato nel programma delle attività annuali (di cui all'articolo 4 del decreto legislativo 25 luglio 2006, n. 240), al fine di integrare i diversi documenti di programmazione richiesti.

Programmi di gestione penale

La recente modifica all'art. 37 co. 1 del D.L. n. 98/2011 ha reso obbligatorio per i dirigenti degli uffici giudiziari giudicanti anche il programma di gestione nel settore penale, prevedendo che contenga – oltre agli obiettivi di riduzione della durata dei procedimenti concretamente raggiungibili nell'anno in corso e agli obiettivi di rendimento dell'ufficio già citati per i programmi di gestione civili - i criteri di priorità nella trattazione dei procedimenti pendenti, sulla base delle disposizioni di legge e delle linee guida elaborate dal CSM.

Il CSM già da qualche anno aveva chiesto ai capi degli uffici giudicanti di predisporre un programma di gestione anche per il settore penale al fine di programmare le attività tenendo conto anche delle esigenze di contenimento della durata dei processi penali e della opportunità di distribuire le risorse disponibili in funzione del carico complessivo di lavoro dell'ufficio nei due settori.

Anche per la predisposizione dei programmi di gestione penale è stato predisposto un format nel quale viene richiesto al capo dell'ufficio

di assicurare un giusto equilibrio tra le diverse esigenze di contenere la durata media dei procedimenti, di perseguire un obiettivo di rendimento dell'ufficio che tenga conto dei carichi esigibili e dei criteri di priorità nella trattazione dei procedimenti penali, di perseguire standard qualitativi di rendimento.

Le informazioni richieste e necessarie alla programmazione, mutuando quanto richiesto ai fini della predisposizione del programma di gestione civile, riguardano la durata media dei procedimenti, gli obiettivi di smaltimento dell'ufficio con riferimento alla materia oggetto del procedimento o al numero degli imputati o all'anno di iscrizione del procedimento, il carico esigibile da ciascun magistrato, gli obiettivi di qualità, i criteri di priorità, l'attività di monitoraggio adottata.

Anche per il settore penale, il capo dell'ufficio predispone il programma attraverso una procedura partecipata, utilizzando i dati statistici messi a disposizione da parte dell'Ufficio Statistico del CSM; successivamente il programma di gestione viene depositato e analizzato dal Consiglio Giudiziario e dal CSM per la sua approvazione.

4. Valutazioni di professionalità e standard di rendimento

Il D.L. vo n. 160/2006, come modificato dagli artt. 1 e 2 della L. n. 111/2007, prevede che i magistrati siano sottoposti a valutazione di professionalità con cadenza quadriennale, a decorrere dalla data della nomina e fino al superamento della settima valutazione.

Sulla base di quanto disciplinato nella suddetta norma, la valutazione di professionalità riguarda la capacità, la laboriosità, la diligenza e l'impegno. In particolare:

“ a) la capacità, oltre che alla preparazione giuridica e al relativo grado di aggiornamento, è riferita, secondo le funzioni esercitate, al possesso delle tecniche di argomentazione e di indagine, anche in relazione all'esito degli affari nelle successive fasi e nei gradi del procedimento e del giudizio ovvero alla conduzione dell'udienza da parte di chi la dirige o la presiede, all'idoneità a utilizzare, dirigere e controllare l'apporto dei collaboratori e degli ausiliari;

b) la laboriosità è riferita alla produttività, intesa come numero e qualità degli affari trattati in rapporto alla tipologia degli uffici e alla loro condizione organizzativa e strutturale, ai tempi di smaltimento del

lavoro, nonché all'eventuale attività di collaborazione svolta all'interno dell'ufficio, tenuto anche conto degli standard di rendimento individuati dal Consiglio superiore della magistratura, in relazione agli specifici settori di attività e alle specializzazioni;

c) la diligenza è riferita all'assiduità e puntualità nella presenza in ufficio, nelle udienze e nei giorni stabiliti; è riferita inoltre al rispetto dei termini per la redazione, il deposito di provvedimenti o comunque per il compimento di attività giudiziarie, alla partecipazione alle riunioni previste dall'ordinamento giudiziario per la discussione e l'approfondimento delle innovazioni legislative, per la conoscenza dell'evoluzione della giurisprudenza;

d) l'impegno è riferito alla disponibilità per sostituzioni di magistrati assenti e alla frequenza di corsi di aggiornamento organizzati dalla Scuola superiore della magistratura; nella valutazione dell'impegno rileva, inoltre, la collaborazione alla soluzione dei problemi di tipo organizzativo e giuridico."

La valutazione di professionalità si conclude con un giudizio:

- a) "positivo", quando la valutazione è sufficiente in relazione a tali parametri. In tal caso, il magistrato consegue la valutazione di professionalità ed è sottoposto a successiva valutazione quadriennale;
- b) "non positivo", quando emergono carenze in relazione a uno o più dei medesimi parametri. In tal caso il magistrato è soggetto ad ulteriore valutazione per il periodo di un anno;
- c) "negativo", quando risultano gravi carenze in relazione a due o più dei parametri o il perdurare di carenze in uno o più dei parametri che hanno portato ad un precedente giudizio "non positivo". In tal caso il magistrato è soggetto ad ulteriore valutazione per un periodo di due anni. (art. 11, D.lgs. n. 160/2006). Se il CSM esprime un secondo giudizio negativo il magistrato è dispensato dal servizio.

Il magistrato deve quindi allegare alla sua autorelazione le statistiche sui flussi estratte dai sistemi informativi in uso negli uffici giudiziari.

Nell'ambito della laboriosità, il CSM avrebbe dovuto individuare degli standard di rendimento specifici per attività e specializzazione.

Anche l'attuale Circolare del CSM prevede che gli indici di laboriosità riguardino: il numero di procedimenti e processi definiti, i flussi in entrata degli affari, la complessità dei procedimenti assegnati e trattati; il rispetto degli standard medi di definizione dei procedimenti, determinati

annualmente dal CSM Consiglio e comunicati, tramite una scheda individuale, ai magistrati in valutazione, ai capi degli uffici ed ai Consigli giudiziari.

Tali standard sono individuati in base alla media statistica della produzione dei magistrati dell'ufficio di cui il magistrato sottoposto a valutazione fa parte ed assegnati a funzioni, sezioni, gruppi di lavoro omogenei a quest'ultimo, ma tenendo anche conto dei flussi in entrata, della complessità degli affari trattati, dell'attività di collaborazione alla gestione dell'ufficio, degli esoneri, delle assenze.

Nonostante il lavoro svolto dalle Commissioni incaricate all'individuazione degli standard di rendimento e le proposte operative formulate, tale parametro di riferimento non è ancora stato individuato da parte del CSM.

5. Il Piano Nazionale di Ripresa e Resilienza

Il Piano Nazionale di Ripresa e Resilienza (PNRR) dell'Italia, trasmesso ufficialmente dalla Presidenza del Consiglio dei Ministri alla Commissione europea il 30 aprile 2021, prevede investimenti e riforme per *191,5 miliardi di euro* articolati in 6 Missioni (M) con *16 Componenti* (C). Le 16 Componenti rappresentano gli ambiti in cui aggregare progetti di investimento e progetti di riforma del PNRR.

Il PNRR per il Ministero della Giustizia individua *tre linee di intervento*.

Ufficio per il Processo e Capitale umano (M1C1)-Investire nel capitale umano per rafforzare la struttura organizzativa denominata «Ufficio per il processo» e superare le disparità tra gli uffici giudiziari, migliorarne la performance sia in primo che in secondo grado e accompagnare - completandolo - il processo di transizione digitale del sistema giudiziario.

Digitalizzazione (M1C1)-Investire nella trasformazione digitale, attraverso la digitalizzazione dei fascicoli e l'adozione di strumenti avanzati di analisi dati.

Edilizia giudiziaria (M2C3)-Riqualificare il patrimonio immobiliare dell'amministrazione giudiziaria (dati tratti dal sito del Ministero della Giustizia).

Gli obiettivi del PNRR coinvolgono principalmente due aggregati: la *durata prognostica* o *disposition time* che, secondo quanto deciso dalla CEPEJ, viene stimata con il rapporto tra pendenti alla fine del periodo

e definizioni nell'anno, e l'*arretrato ultratriennale/ultrabiennale* a seconda che si parli di Tribunali o di Corti di Appello.

Nello specifico della prima misura (disposition time) l'obiettivo è quello di riduzione del 40% della durata dei procedimenti civili (56% in Tribunale e Corte di Appello e 25% in Cassazione) entro la metà del 2026 e del 25% della durata dei procedimenti penali entro lo stesso periodo (28% in Tribunale e Corte e 0% in Cassazione).

Per quanto concerne la seconda misura (l'arretrato), l'obiettivo del PNRR è quello dell'abbattimento dell'arretrato civile del 65% in primo grado e del 55% in appello entro la fine del 2024 e l'abbattimento dell'arretrato civile del 90% in tutti i gradi di giudizio entro la metà del 2026. L'abbattimento deve essere calcolato considerando come baseline di riferimento l'anno 2019.

È importante notare che ci sono delle differenze negli aggregati "Disposition time" e "Arretrato"¹.

Per quanto riguarda il settore civile per il calcolo del Disposition time sono considerati i procedimenti contenziosi secondo la classificazione proposta dalla Commissione Europea per l'Efficienza della Giustizia (Cepej): per il Tribunale sono considerati gli affari civili contenziosi con l'esclusione delle separazioni e dei divorzi consensuali, le controversie agrarie e in materia di lavoro, la previdenza e l'assistenza obbligatoria inclusi i procedimenti speciali e gli Accertamenti tecnici preventivi (ATP). Sono esclusi i ruoli degli Affari di Volontaria giurisdizione, dei procedimenti speciali e sommari, nonché le procedure esecutive e concorsuali con l'eccezione delle istanze di fallimento.

Per le Corti di Appello, l'aggregato di riferimento include gli affari civili contenziosi escludendo le separazioni e i divorzi consensuali, le controversie agrarie e in materia di lavoro e previdenza e assistenza obbligatoria, sono escluse le altre tipologie di procedimento con l'eccezione dei procedimenti di Equa riparazione che vanno considerati nell'aggregato.

Per il calcolo dell'Arretrato, la Cepej ha accolto i criteri derivanti dalle disposizioni di leggi nazionali (legge Pinto) e utilizzati per la produzione delle statistiche ufficiali. Pertanto, per il Tribunale sono considerati "Arretrato" tutti i procedimenti pendenti da oltre tre anni in area Sicid (Affari contenziosi, Controversie agrarie, Controversie in materia di lavoro, previdenza, assistenza obbligatoria, Affari di volontaria giurisdizione, procedimenti speciali e sommari), con l'esclusione della

¹ Informazioni estratte dalla circolare DGStat sugli indicatori di raggiungimento degli obiettivi previsti dal PNRR.

materia del Giudice Tutelare, dell'ATP e dell'attività di "ricevimento e verbalizzazione di dichiarazione giurata".

Per la Corte di Appello sono considerati "Arretrato" tutti i procedimenti pendenti da oltre due anni.

In sintesi:

- Le misure considerate nell'aggregato Disposition time ma non nell'aggregato Arretrato sono l'Accertamento Tecnico Preventivo ex art. 445 bis c.p.c. e le istanze di Fallimento nei Tribunali;
- Le misure considerate nell'aggregato Arretrato ma non nell'aggregato Disposition time sono il ruolo dei procedimenti speciali sommati solo per quanto concerne i Tribunali, il ruolo della Volontaria giurisdizione, escluso la materia del giudice tutelare e le separazioni e i divorzi consensuali per Tribunali e Corti d'Appello.

Per quanto concerne il settore penale, per il calcolo del Disposition time, i criteri richiesti dalla Commissione europea coincidono con quelli utilizzati per le statistiche ufficiali nazionali: per il Tribunale sono considerati i procedimenti relativi alle sezioni Gip-Gup autore noto, Dibattimento in primo grado, Dibattimento in II grado, e Assise, sono escluse le decisioni interlocutorie del Gip, i procedimenti di competenza del Tribunale del riesame e i procedimenti per l'applicazione di misure di prevenzione. Per la Corte di Appello l'aggregato di riferimento è dato dai procedimenti relativi alle sezioni ordinaria, assise e minorenni.

Il capo dell'ufficio predispone un progetto organizzativo ai fini del raggiungimento degli obiettivi previsti dal PNRR e dell'ottimizzazione delle risorse assegnate dal Ministero all'Ufficio del Processo. In tale documento individua le criticità dell'ufficio, i principali obiettivi sui quali intervenire, le azioni che si intendono proporre, i criteri e le modalità di distribuzione delle risorse, i servizi di monitoraggio statistico continuo, di supporto e di coordinamento delle attività amministrative inerenti al PNRR.

Il Ministero monitorerà periodicamente il raggiungimento degli obiettivi di riduzione della durata e di abbattimento dell'arretrato sulla base di quanto previsto dalla Commissione Europea.

6. *L'ufficio del Processo*

Al fine del miglioramento del servizio giustizia e del raggiungimento degli obiettivi previsti da PNRR, il Ministero ha posto in essere un programma di interventi organizzativi e normativi per fornire a Tribunali e Corti di Appello la cornice normativa, le risorse finanziarie e gli strumenti informatici per avviare l'organizzazione di strutture di staff in grado di affiancare il giudice nelle attività di ufficio.

È stato attuato in molti uffici giudiziari *l'Ufficio per il processo*, la misura organizzativa più rilevante in termini di investimento finanziario ma anche per finalità e obiettivi assegnati nell'ambito del PNRR, che consente di supportare processi di innovazione attraverso la costituzione di veri e propri staff di supporto al magistrato e alla giurisdizione.

L'Ufficio per il processo è una innovazione organizzativa che permette di creare maggiori sinergie tra il lavoro dei magistrati e quello delle cancellerie ai fini di un miglioramento qualitativo del lavoro.

I compiti dello staff sono infatti sia quelli di studio, ricerca, redazione di bozze di provvedimenti, stesura di relazioni, massimizzazione di sentenze, ma anche quelli di raccordo con le cancellerie e le segreterie, di preparazione dell'udienza con il magistrato, di assistenza in generale al capo dell'ufficio e ai presidenti di sezione per le attività di innovazione, di monitoraggio statistico e organizzativo, di supporto alla creazione di indirizzi giurisprudenziali e di banca dati, etc.

Come misura a sostegno dell'Ufficio del Processo, è stato messo a punto il cosiddetto *Progetto Task Force* ovvero il "Progetto unitario per la diffusione dell'Ufficio per il Processo e l'implementazione di modelli operativi innovativi negli uffici giudiziari per lo smaltimento dell'arretrato" volto a promuovere l'efficienza della prestazione degli Uffici giudiziari attraverso l'implementazione di modelli organizzativi innovativi, come appunto l'Ufficio per il Processo, finalizzati ad incrementare l'efficienza del lavoro del giudice.

A tal fine le principali azioni avviate riguardano la messa a punto di progetti pilota per la creazione di Uffici per il processo; progetti di coordinamento tra uffici, università e ordini degli avvocati per la "massimazione" di indirizzi giurisprudenziali; la diffusione del software di Consolle Assistente presso gli uffici pilota.

L'obiettivo è anche quello di implementare negli uffici modelli operativi finalizzati alla gestione delle pratiche e delle modalità di definizione attraverso delle sperimentazioni in uffici pilota di modelli

innovativi individuati a partire da linee guida comuni e la diffusione negli stessi uffici pilota dei dati statistici necessari all'implementazione del modello operativo.

Il Progetto, le cui beneficiarie sono le università, ha l'obiettivo di rafforzare le sinergie tra sistema della giustizia e sistema di ricerca universitaria facendo in modo che il sistema universitario si adegui maggiormente alle esigenze del sistema giudiziario nella sua offerta formativa.

Ciò si aggiunge ovviamente agli obiettivi pratici dell'aumento della produttività degli uffici, della riduzione dei tempi di lavorazione e definizione delle pratiche, dell'individuazione di metodi uniformi per ridurre l'arretrato civile, del miglioramento della qualità del lavoro dei giudici e dei cancellieri anche attraverso un maggior uso tecnologico da parte degli stessi (informazioni tratte dal sito del Ministero della Giustizia).

7. Digitalizzazione e data lake

Come accennato nei paragrafi in precedenza, nelle linee di progetto collegate al PNRR il Ministero della Giustizia ha intrapreso già da tempo la strada della digitalizzazione attraverso il Processo civile telematico (PCT) introdotto già nel 2010 e obbligatorio a partire dal 2014 in primo e secondo grado e il Processo penale telematico sul quale si sta avendo una accelerazione in tempo attuale.

Il PCT costituisce una serie di attività tipicamente processuali che erano precedentemente realizzate in forma cartacea e successivamente costituite in via telematica grazie all'utilizzo di opportuni strumenti informatici. Dati e documenti informatici divengono pertanto oggetto di interscambio, firmati digitalmente e veicolati su canali sicuri e criptati e il fascicolo informatico si sostituisce così definitivamente al fascicolo cartaceo. Il sistema realizza una operabilità bidirezionale tra applicazioni esterne al dominio Giustizia (come ad esempio sistemi di avvocatura e di altre PPAA.) e le applicazioni presso gli uffici giudiziari.

Tramite il PCT è possibile consultare on line il fascicolo processuale, comunicare telematicamente con altri uffici giudiziari, nonché gestire via web l'intero settore civile tramite i registri SIECIC per le procedure esecutive e fallimentari e SICID per i procedimenti contenziosi, di volontaria giurisdizione, lavoro e previdenza.

Il PCT si basa sulla scansione degli atti depositati nei processi civili di Tribunali e Corti negli ultimi 10 anni e delle pronunce della Corte di

Cassazione, permettendo così l'accesso da remoto agli atti e ai documenti con conseguente riduzione dei tempi di accesso e di lavorazione delle richieste.

Oltre ad eliminare la componente cartacea con il PCT infatti viene eliminata anche la gestione di archivi cartacei producendo pertanto un benefico impatto dal punto di vista dello snellimento archivistico e degli oneri di ricerca di archivio, permettendo inoltre un accesso più rapido alle informazioni, ai dati e ai documenti e in continuità operativa ovvero che può avvenire simultaneamente da parte di tutti gli utenti del sistema giustizia e degli altri fruitori del servizio.

La digitalizzazione sta creando e porterà sempre di più a un ampliamento della capacità informativa e di analisi degli strumenti anche attraverso le tecnologie di intelligenza artificiale applicate alla giustizia, come ad esempio i Data Lake. L'approccio Data Lake, a differenza di quanto avviene con i tradizionali sistemi di Big Data Analytics e Data Warehouse, non necessita di una strutturazione ex ante del dato. Infatti, mentre nell'attuale Data Warehouse civile è stata definita a priori la struttura del database, nel Data Lake i dati sono acquisiti nel formato nativo permettendo l'accesso a una quantità maggiore di informazioni derivanti dalle diverse tipologie di dati e interrogazioni potenzialmente possibili, indipendentemente dalla sorgente che li ha generati; la capacità di accogliere dati strutturati, semi-strutturati e destrutturati è il punto di forza di tale tecnologia.

Per quanto concerne il Processo penale telematico (PPT), la digitalizzazione ha riguardato l'informatizzazione dei registri di cancelleria realizzata dal Ministero della Giustizia con il SICP (Sistema Informativo della Cognizione Penale), un registro unico nazionale composto da Re.Ge. Web, il registro generale delle notizie di reato, da SIRIS/ARES che consente ricerche nel Re.ge storico e nel Re.Ge Web e Consolle del magistrato e che contiene delle query standardizzate con cui il magistrato può monitorare i dati relativi al suo ruolo e alle relative scadenze.

Oltre al SICP, nell'ambito del Processo penale telematico sono stati sviluppati altri sistemi come il SIES (Sistema Integrato esecuzione sorveglianza), il SIUS (Sistema Informativo Uffici di Sorveglianza), il SIAMM per le spese di giustizia e la gestione del personale, il SIPPI, che è il sistema per le misure di prevenzione e che di recente è evoluto nel S.I.T. M.P., il SIGMA per gli uffici minorili civile e penale, il SNT-Sistema Notifiche Telematiche, il TIAP e l'AUROR@ che sono due gestori documentali.

Con il PPT si vuole realizzare un Work Flow digitalizzato che permetta di ampliare la tipologia di provvedimenti e di soggetti abilitati alla redazione, firma e deposito digitale e telematico dei provvedimenti al fine di rendere

digitali tutti i flussi procedurali.

Il Portale Deposito atti Penali (PDP), si pone l'obiettivo di ampliare la tipologia di atti e di soggetti abilitati al deposito (e all'accesso/visione) telematico, comprendendovi tutti gli atti di parte formati e redatti da avvocati ed ausiliari (periti, consulenti tecnici, interpreti).

La Direzione Generale dei Sistemi informativi sta valutando come uniformare e integrare i diversi sistemi informativi disponibili nel settore penale.

Raffaello Lupi

Postfazione

Statistica e studio giuridico sociale della funzione giudiziaria

1 *Statistica giudiziaria come previsione e interpretazione*

Il volume di Filomeno e Rocchetti consente alcune riflessioni sugli studi sociali, e sul loro rapporto col diritto, esaminato da chi scrive nel volume *Studi sociali e diritto*, anch'esso pubblicato su Roma 3 Press <https://romatrepress.uniroma3.it/libro/studi-sociali-e-diritto/>. La statistica giudiziaria si inserisce infatti a pieno titolo negli studi sociali, esaminando vicende umane, come la funzione pubblica di giustizia; esso guarda in modo secondario agli aspetti previsionali della statistica, quali i sondaggi di opinione o le intenzioni di voto elettorali, rilevate in base a una campionatura da cui effettuare proiezioni su tendenze e preferenze della società in generale. Nella statistica giudiziaria prevalgono invece, per la natura stessa della disciplina, gli aspetti interpretativi di eventi già verificatisi, classificati in modo da averne chiavi di lettura. La relativa finalità è valutare l'andamento della funzione giudiziaria, per averne elementi di indirizzo e coordinamento, sia a livello generale che di singoli uffici giudiziari. Per questo, come rileva anche l'introduzione del volume, è fondamentale la corretta rilevazione e classificazione dei dati nella quotidianità del lavoro degli uffici. Oltre alla raccomandazione alla diligenza degli operatori chiamati a questo lavoro di raccolta, richiamerei anche l'importanza dei criteri con cui si decide di effettuare la rilevazione, cioè cosa si decide di registrare o meno, magari perché sfuggente, anche se importante.

Si tratta di valutazioni in cui si combinano l'importanza del dato statistico, la sua pratica rilevabilità e la sua interpretabilità ai fini delle valutazioni di cui sopra. Alcuni dati rilevanti, ai fini del controllo sociale su questa funzione, hanno carattere qualitativo, e sono poco inseribili nei modelli statistici; un esempio di dato qualitativo difficilmente rilevabile è l'articolazione o la stringatezza delle udienze, o delle motivazioni delle sentenze rispetto alle allegazioni delle parti e alla natura delle controversie. Un magistrato più scrupoloso e diligente, che fa meno udienze, ma più

approfondite e diciamo così *partecipate*, rischia forse di essere giudicato quantitativamente meno efficiente di un collega più attento alla velocità che alla qualità. Si tratta a ben vedere di due punti di vista da contemperare nell'esercizio di qualsiasi funzione pubblica, cui si collega un'ineliminabile discrezionalità, da non rimuovere, ma da analizzare; sul tema rinvio alle pagg.70-74 del sopra citato *Studi sociali e diritto*, nonché -vista la sede- ad un volume che a mio avviso vale la pena di recuperare nelle biblioteche, cioè *Studi sul potere discrezionale del giudice civile*, di Alessandro Raselli (Giuffrè, edizione 1975). Un'accurata programmazione dei dati da rilevare ai fini della valutazione statistica dovrebbe sforzarsi di trovare strumenti per intercettare anche i casi in cui apparenti carenze quantitative sono compensate da profili qualitativi, sul piano dell'impegno e dello sforzo richiesto nelle specifiche pratiche trattate. Si tratterebbe idealmente di ostacolare comportamenti opportunistici, dove il contenuto dell'attività è falsato dalla volontà di *fare statistica*; un esempio è la scena del film *Il medico della mutua* in cui Alberto Sordi mette tre punti di sutura ad un taglietto, clinicamente superflui, per raggiungere la soglia necessaria a considerare l'intervento come *un'operazione*, ai fini della rendicontazione mutualistica. Sono aspetti cui si dovrebbe riservare una qualche considerazione, sia pure con un impegnativo esercizio d'immaginazione, già ex ante, nella programmazione delle rilevazioni statistiche. A tal fine, in una qualche misura, è auspicabile che nell'elaborare le proprie griglie di raccolta dei dati su un certo fenomeno sociale, lo statistico si avventuri anche nell'analisi del medesimo. Anche i relativi cultori devono sapersi mettere in una prospettiva statistica, e viceversa, in una proficua osmosi che non è avvenuta, in materia tributaria, nella costruzione degli studi di settore per la stima dei ricavi e compensi di oltre il 95% delle imprese italiane; ne sono derivati alcuni paradossi metodologici che hanno delegittimato lo strumento agli occhi degli operatori del settore, portando alla sua abrogazione, a vantaggio di *indici di affidabilità economica* altrettanto sfuggenti rispetto al senso comune. Il fatto è che la statistica, ed anche l'intelligenza artificiale, come accenneremo al punto seguente, non possono essere la ciambella di salvataggio di funzioni incapaci di spiegare se' stesse. La conoscenza del settore è quindi la base di partenza per un esercizio di immaginazione dell'interpretazione futura dei dati, stabilendo le tipologie di quelli da rilevare; c'è un'interazione tra questi due profili, cioè la necessità di fare mente locale sulle tipologie di liti giudiziarie, in modo da prevedere le informazioni più significative ai fini della valutazione indicata sopra.

2. *Statistica, controllo sociale sulla giustizia e intelligenza artificiale*

Una volta raccolti i dati di cui al punto precedente, si tratta di procedere alla loro aggregazione, presentazione e modalità di divulgazione; anche quest'ultima comporta un'interpretazione del dato, che a ben vedere *non parla mai da solo*. Esso rileva invece ai fini di un controllo sociale, svolto a vari livelli d'interesse, intrecciato con le percezioni esperienziali degli interessati al tema, nel nostro caso la giustizia. Non si tratta di un controllo sociale tecnico giuridico, come quello di legittimità e di opportunità su casi singoli, ma relativo all'amministrazione della giustizia nel suo complesso, ed a sue articolazioni per area territoriale. Si tratta di un controllo definibile, nelle grandi linee, di *economicità* ed *efficienza*, da svolgere tenendo conto delle risorse disponibili, in termini finanziari e di personale. Un'interpretazione più complessa, tipicamente giuridico sociale, riguarda l'impatto della funzione di giustizia sulle aree sottostanti della vita reale; mi riferisco all'impatto del *mondo legale* sul mondo reale, in termini di stimolo a comportamenti socialmente virtuosi, e di dissuasione rispetto a comportamenti opportunistici e devianti. Sono valutazioni che potrebbero essere svolte per cause seriali, come gli sfratti per morosità, l'infortunistica stradale, il recupero crediti, le separazioni familiari, etc; un esempio penalistico di interazione tra mondo legale e mondo reale riguarda la microcriminalità, formalmente in calo salvo interrogarsi sulla mancanza di incentivi a denunciare per la modesta prospettiva di individuare i colpevoli o ottenere ristori, salve necessità assicurative o di formalizzare la perdita di disponibilità di beni mobili registrati, soprattutto veicoli.

In base a queste ulteriori considerazioni, che sfuggono inevitabilmente alle metodologie statistiche, l'opinione pubblica forma le proprie percezioni dei vari settori della macchina statale. In questa prospettiva mi sembra di poter escludere l'idea dello *statistico tuttologo*, che passa dall'evasione fiscale, alla sanità, agli incidenti stradali, alle classifiche di povertà, e nel nostro caso alla giustizia. Credo invece che per categorizzare statisticamente un fenomeno sociale occorra una combinazione tra lo studio e l'esperienza continuativa col medesimo, per capire cosa si domandano gli interessati all'argomento. Torna il circolo virtuoso tra statistici e studiosi sociali del settore, nel nostro caso inevitabilmente anche giuristi, già indicato al punto precedente e riguardante anche i meccanismi di intelligenza artificiale; di essa il volume non si occupa direttamente, ma oggi se ne parla molto e vi fa giustamente riferimento anche la prefazione della dr.ssa Ciriello. Anche per costruire modelli di intelligenza artificiale, come nella statistica sociale,

occorre individuare in base all'intelligenza naturale passaggi rilevanti e ricorrenti del comportamento umano. In relazione ad essi l'intelligenza artificiale va prima costruita dalla suddetta osmosi tra informatici, statistici e studiosi del settore su cui intervenire. In un secondo passaggio essa va utilizzata dall'operatore, nel nostro caso ipoteticamente il giudice. Sotto entrambi gli aspetti occorre padroneggiare i processi mentali da riprodurre, in modo da scomporli e ricostruirli. Il vantaggio dell'intelligenza artificiale è eliminare i condizionamenti emotivi contingenti che caratterizzano la gestione di argomenti parzialmente dissonanti, in fatto e in diritto; è un'emotività frequente in materia giudiziaria, come conferma l'influenza che l'ascoltatore subisce di fronte alle arringhe processuali dell'accusa o della difesa, in un medesimo processo. La rilevazione statistica può quindi aiutare a costruire meccanismi di intelligenza artificiale che contemperino in modo parallelo gli elementi di decisione, riducendo l'emotività immanente nell'intelligenza naturale. Anche qui la statistica aiuta a preparare il terreno per un'ineliminabile attività interpretativa, ma più scomponibile, consapevole e trasparente nel gestire variabili squisitamente umane, e quindi sfuggenti. Statistica e intelligenza artificiale vanno quindi gestite evitando la tentazione dell'ipocrisia tecnica (*juristen bese christen* scriveva Lutero), portata per comodità a *far dire ai dati* o all'intelligenza artificiale quello che pare loro più opportuno; ricordo i commenti ai risultati elettorali dove *tutti dicono di aver vinto* o la battuta *se torturi i dati abbastanza finiranno per confessare quello che vuoi*. Né la statistica né l'intelligenza artificiale possono insomma totalmente sostituire il discernimento umano, nel nostro caso quello del giudice, cui però tornano utili per dare un peso specifico comparativo ai vari elementi da rilevare e da interpretare.

Appendice

Tabella 21. Distribuzione della dimensione dei Tribunali e delle Procure della Repubblica a dicembre 2022

Distretto	Sede	Organico Tribunali	Organico Procure della Repubblica	Dimensione Tribunali	Dimensione Procure della Repubblica
Ancona	Ancona	33	13	Medio piccoli	Medio piccoli
Ancona	Ascoli Piceno	14	5	Piccoli	Piccoli
Ancona	Fermo	14	5	Piccoli	Piccoli
Ancona	Macerata	22	9	Medio piccoli	Piccoli
Ancona	Pesaro	16	6	Piccoli	Piccoli
Ancona	Urbino	6	3	Piccoli	Piccoli
Bari	Bari	97	41	Medio grandi	Grandi
Bari	Foggia	69	28	Medio grandi	Medio grandi
Bari	Trani	40	13	Medio piccoli	Medio piccoli
Bologna	Bologna	82	32	Medio grandi	Grandi
Bologna	Ferrara	22	9	Medio piccoli	Piccoli
Bologna	Forlì	22	8	Medio piccoli	Piccoli
Bologna	Modena	41	14	Medio piccoli	Medio piccoli
Bologna	Parma	29	11	Medio piccoli	Piccoli
Bologna	Piacenza	17	7	Piccoli	Piccoli
Bologna	Ravenna	24	10	Medio piccoli	Piccoli
Bologna	Reggio nell'Emilia	29	11	Medio piccoli	Piccoli
Bologna	Rimini	25	10	Medio piccoli	Piccoli
Bolzano	Bolzano	39	12	Medio piccoli	Medio piccoli
Brescia	Bergamo	56	20	Medio grandi	Medio piccoli
Brescia	Brescia	74	30	Medio grandi	Medio grandi
Brescia	Cremona	21	8	Medio piccoli	Piccoli

Distretto	Sede	Organico Tribunali	Organico Procure della Repubblica	Dimensione Tribunali	Dimensione Procure della Repubblica
Brescia	Mantova	22	9	Medio piccoli	Piccoli
Cagliari	Cagliari	62	22	Medio grandi	Medio grandi
Cagliari	Lanusei	6	3	Piccoli	Piccoli
Cagliari	Oristano	17	7	Piccoli	Piccoli
Caltanissetta	Caltanissetta	34	19	Medio piccoli	Medio grandi
Caltanissetta	Enna	18	8	Piccoli	Piccoli
Caltanissetta	Gela	14	6	Piccoli	Piccoli
Campobasso	Campobasso	12	7	Piccoli	Piccoli
Campobasso	Isernia	9	4	Piccoli	Piccoli
Campobasso	Larino	9	3	Piccoli	Piccoli
Catania	Caltagirone	13	5	Piccoli	Piccoli
Catania	Catania	117	47	Grandi	Grandi
Catania	Ragusa	26	9	Medio piccoli	Piccoli
Catania	Siracusa	37	16	Medio piccoli	Medio piccoli
Catanzaro	Castrovillari	30	11	Medio piccoli	Piccoli
Catanzaro	Catanzaro	52	32	Medio grandi	Grandi
Catanzaro	Cosenza	38	14	Medio piccoli	Medio piccoli
Catanzaro	Crotone	23	6	Medio piccoli	Piccoli
Catanzaro	Lamezia Terme	17	5	Piccoli	Piccoli
Catanzaro	Paola	17	6	Piccoli	Piccoli
Catanzaro	Vibo Valentia	22	8	Medio piccoli	Piccoli
Firenze	Arezzo	23	9	Medio piccoli	Piccoli
Firenze	Firenze	88	34	Medio grandi	Grandi
Firenze	Grosseto	18	8	Piccoli	Piccoli
Firenze	Livorno	26	9	Medio piccoli	Piccoli
Firenze	Lucca	29	11	Medio piccoli	Piccoli

Distretto	Sede	Organico Tribunali	Organico Procure della Repubblica	Dimensione Tribunali	Dimensione Procure della Repubblica
Firenze	Pisa	28	11	Medio piccoli	Piccoli
Firenze	Pistoia	20	8	Piccoli	Piccoli
Firenze	Prato	26	10	Medio piccoli	Piccoli
Firenze	Siena	19	8	Piccoli	Piccoli
Genova	Genova	95	36	Medio grandi	Grandi
Genova	Imperia	21	12	Medio piccoli	Medio piccoli
Genova	La Spezia	21	8	Medio piccoli	Piccoli
Genova	Massa	15	6	Piccoli	Piccoli
Genova	Savona	24	9	Medio piccoli	Piccoli
L'Aquila	Avezzano	11	5	Piccoli	Piccoli
L'Aquila	Chieti	15	5	Piccoli	Piccoli
L'Aquila	Lanciano	7	3	Piccoli	Piccoli
L'Aquila	L'Aquila	14	7	Piccoli	Piccoli
L'Aquila	Pescara	27	13	Medio piccoli	Medio piccoli
L'Aquila	Sulmona	6	3	Piccoli	Piccoli
L'Aquila	Teramo	23	10	Medio piccoli	Piccoli
L'Aquila	Vasto	7	3	Piccoli	Piccoli
Lecce	Brindisi	38	14	Medio piccoli	Medio piccoli
Lecce	Lecce	66	22	Medio grandi	Medio grandi
Messina	Barcellona Pozzo di Gotto	17	6	Piccoli	Piccoli
Messina	Messina	49	23	Medio piccoli	Grandi
Messina	Patti	19	6	Piccoli	Piccoli
Milano	Busto Arsizio	36	13	Medio piccoli	Medio piccoli
Milano	Como	30	12	Medio piccoli	Piccoli
Milano	Lecco	15	6	Piccoli	Piccoli

Distretto	Sede	Organico Tribunali	Organico Procure della Repubblica	Dimensione Tribunali	Dimensione Procure della Repubblica
Milano	Lodi	20	7	Piccoli	Piccoli
Milano	Milano	290	94	Metropolitani	Metropolitani
Milano	Monza	60	18	Medio grandi	Medio piccoli
Milano	Pavia	37	14	Medio piccoli	Medio piccoli
Milano	Sondrio	12	6	Piccoli	Piccoli
Milano	Varese	26	9	Medio piccoli	Piccoli
Napoli	Avellino	38	13	Medio piccoli	Medio piccoli
Napoli	Benevento	35	15	Medio piccoli	Medio piccoli
Napoli	Napoli	315	114	Metropolitani	Metropolitani
Napoli	Napoli Nord in Aversa	86	32	Medio grandi	Medio grandi
Napoli	Nola	51	17	Medio grandi	Medio piccoli
Napoli	Santa Maria Capua Vetere	91	28	Medio grandi	Medio grandi
Napoli	Torre Annunziata	51	16	Medio grandi	Medio piccoli
Palermo	Agrigento	32	14	Medio piccoli	Medio piccoli
Palermo	Marsala	24	9	Medio piccoli	Piccoli
Palermo	Palermo	129	71	Grandi	Grandi
Palermo	Sciacca	10	5	Piccoli	Piccoli
Palermo	Termini Imerese	25	10	Medio piccoli	Piccoli
Palermo	Trapani	25	13	Medio piccoli	Medio piccoli
Perugia	Perugia	37	16	Medio piccoli	Medio piccoli
Perugia	Spoleto	16	6	Piccoli	Piccoli
Perugia	Terni	20	8	Piccoli	Piccoli
Potenza	Lagonegro	19	5	Piccoli	Piccoli
Potenza	Matera	16	6	Piccoli	Piccoli
Potenza	Potenza	35	15	Medio piccoli	Medio piccoli

Distretto	Sede	Organico Tribunali	Organico Procure della Repubblica	Dimensione Tribunali	Dimensione Procure della Repubblica
Reggio Di Calabria	Locri	30	8	Medio piccoli	Piccoli
Reggio Di Calabria	Palmi	31	11	Medio piccoli	Medio piccoli
Reggio Di Calabria	Reggio di Calabria	63	33	Medio grandi	Grandi
Roma	Cassino	25	10	Medio piccoli	Piccoli
Roma	Civitavecchia	26	10	Medio piccoli	Piccoli
Roma	Frosinone	23	8	Medio piccoli	Piccoli
Roma	Latina	43	15	Medio piccoli	Medio piccoli
Roma	Rieti	13	6	Piccoli	Piccoli
Roma	Roma	373	107	Metropolitani	Metropolitani
Roma	Tivoli	30	12	Medio piccoli	Piccoli
Roma	Velletri	43	15	Medio piccoli	Piccoli
Roma	Viterbo	21	8	Medio piccoli	Piccoli
Salerno	Nocera Inferiore	33	12	Medio piccoli	Piccoli
Salerno	Salerno	75	33	Medio grandi	Grandi
Salerno	Vallo della Lucania	14	4	Piccoli	Piccoli
Sassari	Nuoro	16	8	Piccoli	Piccoli
Sassari	Sassari	28	10	Medio piccoli	Piccoli
Sassari	Tempio Pausania	15	7	Piccoli	Piccoli
Taranto	Taranto	62	20	Medio grandi	Medio piccoli
Torino	Alessandria	29	13	Medio piccoli	Medio piccoli
Torino	Aosta	8	5	Piccoli	Piccoli
Torino	Asti	24	11	Medio piccoli	Medio piccoli
Torino	Biella	12	4	Piccoli	Piccoli
Torino	Cuneo	26	11	Medio piccoli	Medio piccoli

Distretto	Sede	Organico Tribunali	Organico Procure della Repubblica	Dimensione Tribunali	Dimensione Procure della Repubblica
Torino	Ivrea	23	10	Medio piccoli	Piccoli
Torino	Novara	19	8	Piccoli	Piccoli
Torino	Torino	163	65	Grandi	Grandi
Torino	Verbania	14	5	Piccoli	Piccoli
Torino	Vercelli	19	8	Piccoli	Piccoli
Trento	Rovereto	9	3	Piccoli	Piccoli
Trento	Trento	22	11	Medio piccoli	Piccoli
Trieste	Gorizia	12	6	Piccoli	Piccoli
Trieste	Pordenone	21	9	Medio piccoli	Piccoli
Trieste	Trieste	27	12	Medio piccoli	Piccoli
Trieste	Udine	36	16	Medio piccoli	Medio piccoli
Venezia	Belluno	12	6	Piccoli	Piccoli
Venezia	Padova	47	18	Medio piccoli	Medio piccoli
Venezia	Rovigo	19	7	Piccoli	Piccoli
Venezia	Treviso	41	14	Medio piccoli	Piccoli
Venezia	Venezia	62	28	Medio grandi	Grandi
Venezia	Verona	52	19	Medio grandi	Medio piccoli
Venezia	Vicenza	43	17	Medio piccoli	Medio piccoli

Riferimenti Bibliografici

- Agresti, A. (2002). *Categorical Data Analysis*, (2nd ed.). Wiley Series in Probability and Statistics;
- Dagum, E.B. (2002). *Analisi Delle Serie Storiche: Modellistica, Previsione e Scomposizione (Italian Edition)*, Springer;
- Everitt, B. S., Landau S., Leese M., Stahl D. (2011). *Cluster Analysis*, (5th ed.). Wiley Series in Probability and Statistics.
- Fortini, M., *Linee guida metodologiche per rilevazioni statistiche Nozioni metodologiche di base e pratiche consigliate per rilevazioni statistiche dirette o basate su fonti amministrative* (<https://digilander.libero.it/rosarb/Lineeguida_rilevazioni_Istat.pdf>), linee guida Istat 2000.
- Guarnera, U., *Alcuni metodi di imputazione delle mancate risposte parziali per dati quantitativi. Il software QUIS*. Istat paper 2018.
- ISTAT, *L'Italia in 150 anni. Sommario di statistiche storiche 1861-2010*, 2011.
- Manzari, A., *Valutazione comparativa di alcuni metodi di imputazione singola delle mancate risposte parziali per dati quantitativi*. Istat paper 2004.
- Russo, A., Russo, M. (2002). *Statistica descrittiva*, Philos.
- Sciacca, M., Verzelloni, L., Miccoli, G. (2013). *Giustizia in bilico*, Aracne, collana Strumenti di giustizia.
- Vitali, O. (1993). *Statistica per le scienze applicate* (2nd ed.), Cacucci.
- Zellini, P. (2018). *La dittatura del calcolo*, IBS.
- Altre informazioni relative ai singoli argomenti sono state dedotte dai siti della Cepej* (<<https://www.coe.int/en/web/cepej/about-cepej>>), *del Ministero della Giustizia* (<<https://webstat.giustizia.it/>>), *del Consiglio Superiore della Magistratura (Csm,* <www.csm.it>) *dell'Istat* (<www.istat.it>, <<https://gino.istat.it/incidenti/>>), *dell'Eurostat* (<<https://ec.europa.eu/eurostat>>) e indicati opportunamente nel testo.

Ringraziamenti

Un ringraziamento va al prof. ordinario di statistica Antonello Maruotti (LUMSA) e a Vito Ayroldi per aver letto il contributo e aver dato le loro osservazioni a riguardo.

La statistica giudiziaria è diventata supporto fondamentale per l'organizzazione degli uffici giudiziari. La diffusione di una cultura del dato giudiziario è cruciale per migliorare la qualità delle informazioni al fine di ottenere una efficace ed efficiente gestione dei dati disponibili. Per raggiungere tale obiettivo è necessaria un'adeguata conoscenza delle fonti, della tipologia dei dati disponibili e delle modalità di estrazione. Mediante l'ausilio di misure, indicatori e strumenti statistici è possibile sintetizzare la mole di dati raccolti consentendone il concreto impiego nella predisposizione di procedure e documenti organizzativi all'interno degli uffici giudiziari. Il testo è un volume indispensabile per chi, senza temere i risultati della scienza che raccoglie i dati per trarne informazioni, voglia con essi confrontarsi in maniera matura e consapevole.

Maria Filomeno

È nata a Putignano nel 1972, si è laureata in Scienze statistiche ed economiche all'Università di Bari e ha conseguito il dottorato di ricerca in Statistica biomedica presso l'Università di Milano. In servizio per oltre 15 anni come funzionario statistico presso il Ministero della Giustizia, attualmente ricopre il ruolo di Direttore dell'Ufficio Statistico del Consiglio Superiore della Magistratura.

Irene Rocchetti

È nata a Roma nel 1982, si è laureata in Scienze statistiche per le analisi demografiche e sociali (laurea triennale) e in Scienze demografiche per le politiche sociali e sanitarie (laurea magistrale) e ha conseguito il dottorato in Demografia presso l'Università La Sapienza di Roma. In servizio come ricercatore metodologico per cinque anni presso l'Istat, attualmente ricopre il ruolo di funzionario dell'Ufficio Statistico del Consiglio Superiore della Magistratura.

