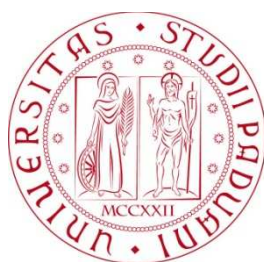


Università degli Studi di Padova  
Dipartimento di Scienze Statistiche  
Corso di Laurea Magistrale in  
Scienze Statistiche



# Evoluzione degli Intimate Partner Homicides in Italia: il ruolo dei Centri Antiviolenza

*Relatore:*

Prof.: Stefano Mazzuco

*Correlatore:*

Prof.ssa: Alessandra Minello

*Laureanda:*

Clara Biancuzzo

Matricola N. 2014983

Anno accademico 2023/2024



# Indice

<b>Introduzione</b>	<b>5</b>
<b>1 Concetti introduttivi</b>	<b>7</b>
1.1 Il fenomeno della violenza di genere . . . . .	7
1.1.1 Gli <i>Intimate Partner Homicides</i> . . . . .	10
1.1.2 I Centri Antiviolenza . . . . .	10
<b>2 L'analisi di interesse</b>	<b>13</b>
2.1 Obiettivo dell'analisi . . . . .	13
2.2 I dati a disposizione . . . . .	13
2.3 Operazioni sui dati a disposizione . . . . .	14
2.3.1 Operazioni preliminari sui <i>dataset</i> distinti . . . . .	16
2.3.2 Risoluzione del problema delle "relazioni mancanti" . . . . .	17
2.3.3 Risoluzione del problema dei duplicati e unione dei due <i>dataset</i> . . . . .	20
<b>3 L'evoluzione degli IPH italiani</b>	<b>23</b>
<b>4 Il ruolo dei Centri Antiviolenza nell'evoluzione degli IPH italiani</b>	<b>43</b>
4.1 Costruzione del <i>dataset</i> . . . . .	43
4.2 Scelta dei modelli statistici . . . . .	46
4.3 Procedura d'analisi e discussione dei risultati . . . . .	50
<b>Appendice</b>	<b>63</b>
<b>Bibliografia</b>	<b>75</b>
<b>Sitografia e Filmografia</b>	<b>77</b>



# Introduzione

A partire dai dati a disposizione in merito agli omicidi avvenuti in Italia nel periodo 1992-2015 e alla relativa presenza di centri antiviolenza sul territorio, questo elaborato si propone di condurre un'analisi georeferenziata allo scopo di indagare il ruolo dei centri antiviolenza sul fenomeno dei cosiddetti *Intimate Partner Homicides* (IPH). Nello specifico, lo studio si concentra dunque sugli IPH femminili (preponderanti rispetto al totale di IPH), al fine di dedicare particolare attenzione ad un tema ben più ampio e quanto mai attuale: la violenza di genere.

In funzione dell'obiettivo d'analisi, l'intero elaborato si struttura come segue. Nel primo capitolo si illustra brevemente il contesto storico-culturale entro cui verrà condotta l'indagine. E nello specifico, si dà una definizione formale dei due macro fenomeni in esame: *Intimate Partner Homicides* femminili e centri antiviolenza. Il secondo capitolo inizia con la definizione più strutturata dell'obiettivo d'analisi, alla quale segue un'illustrazione dei dati resi disponibili e delle specifiche operazioni effettuate su questi ultimi, affinché fosse possibile condurre realmente l'indagine. Il terzo capitolo viene dedicato allo studio degli IPH femminili italiani attraverso un'analisi descrittiva volta ad identificarne le caratteristiche proprie. Infine, il quarto capitolo viene interamente dedicato al tema centrale: lo studio georeferenziato - tramite impiego di modelli parametrici - del ruolo dei centri antiviolenza sulla dinamica evolutiva registrata per gli *Intimate Partner Homicides* femminili italiani nel corso del periodo 1992-2015.



# Capitolo 1

## Concetti introduttivi

In questo capitolo si illustra brevemente il contesto storico-culturale entro cui verrà condotta l'indagine a seguire. E nello specifico, si dà una definizione formale dei due macro fenomeni d'interesse: *Intimate Partner Homicides* femminili e centri antiviolenza.

### 1.1 Il fenomeno della violenza di genere

La Convenzione di Istanbul (2011), punto di riferimento per la legislazione internazionale in materia di lotta alla violenza di genere, inquadra «la violenza contro le donne» come un atto che costituisce «una violazione dei diritti umani e una forma di discriminazione nei confronti delle donne». Forme di tale violenza comprendono qualsiasi atto in grado di causare «danni o sofferenze fisiche, sessuali, psicologiche o economiche, comprese le minacce di tali atti, la privazione della libertà coercitiva o arbitraria, sia nella vita pubblica che in quella privata».

Nonostante i termini "violenza di genere" e "violenza contro le donne" siano spesso usati come sinonimi, in questo elaborato si preferisce l'uso del termine "genere". Quest'ultimo, infatti, ha una sua rilevanza, in quanto rimarca il fatto che molte forme di violenza sulle donne siano radicate nelle disuguaglianze di potere tra i generi.

L'attenzione politica e mediatica nei confronti della violenza di genere è cresciuta sensibilmente solo negli ultimi decenni. Questo è stato possibile, grazie ad una mobilitazione femminile e femminista che ha anticipato e, di

fatto, sostituito per anni il ruolo delle istituzioni. Ad oggi, tuttavia, il peso del fenomeno ha raggiunto un'entità tale da non poter più essere ignorato. L'ISTAT, nel *report* - aggiornato al 2023 - in merito alla violenza di genere<sup>1</sup>, rileva che il 31.5% delle donne italiane in età compresa tra 16 e 70 anni (6 milioni e 788 mila) ha subito una qualche forma di violenza di genere nel corso della propria vita. Ed in particolare, il 13.6% (2 milioni 800 mila) ha subito violenze fisiche o sessuali da parte di *partner* o *ex-partner*. Specie nel contesto delle relazioni intimo-affettive, alle forme di violenza già citate, si aggiungono quelle psicologiche ed economiche: il 26.1% delle donne ha subito violenza psicologica o economica<sup>2</sup> dal *partner*, mentre la percentuale sale a 46.1 se si considerano gli *ex-partners* (dato aggiornato al 2014). Ma i dati evidenziano anche quanto la violenza di genere sia ancora un fenomeno sommerso. È infatti elevata la quota di donne che non parla della violenza subita, di chi non denuncia e, più in generale, delle vittime che non cercano alcun aiuto. Mentre sono ancora poche quelle che si rivolgono ad un centro antiviolenza o ad un qualsiasi altro tipo di servizio specializzato (rispettivamente, il 3.7% nel caso di violenza nella coppia e l'1% nel caso di violenze esterne alla coppia)<sup>3</sup>. Fabio Roia - Presidente vicario del Tribunale di Milano - individua il punto centrale: «Generalmente le donne non hanno la volontà di denunciare, spesso sono proprio gli operatori di polizia o dei centri a spiegare alle vittime del fatto che si tratti di violenza.»<sup>4</sup>. Questa affermazione trova conferma nel fatto che solo un terzo delle donne vittime di violenza (fisica o sessuale) da parte del *partner* ritiene di essere stata effettivamente vittima di un reato. A ciò si aggiunge un aspetto psicologico importante: subire per anni una situazione di maltrattamento fa sì che la vittima tenda ad assorbirla nel tempo - quasi a normalizzarla - e questo rafforza ancor di più la condizione di dipendenza tra le due figure coinvolte (Cristiani<sup>5</sup> in "Un altro domani", 2023). In particolare, le donne non denunciano perchè hanno imparato a gestire la situazione da sole (39.6% per violenze da *partner* e

---

<sup>1</sup><https://www.istat.it/it/violenza-sulle-donne/il-fenomeno/violenza-dentro-e-fuori-la-famiglia/numero-delle-vittime-e-forme-di-violenza>

<sup>2</sup>Con il termine "violenza psicologica ed economica", l'ISTAT fa riferimento a «tutti i comportamenti di umiliazione, svalorizzazione, controllo ed intimidazione, nonché di privazione o limitazione nell'accesso alle proprie disponibilità economiche o della famiglia».

<sup>3</sup><https://www.istat.it/it/violenza-sulle-donne/il-fenomeno/violenza-dentro-e-fuori-la-famiglia/consapevolezza-e-uscita-dalla-violenza>

<sup>4</sup>"Un altro domani", 2023.

<sup>5</sup>Silvia Cristiani - psicologa CIPM (Centro Italiano per la Promozione della Mediazione) Liguria).



39.5% per violenze esterne), non ritengono grave quanto accaduto (rispettivamente, 31.6% e 42.2%), per paura (rispettivamente, 10.1% e 5.0%), per timore, vergogna e/o imbarazzo (rispettivamente, 7.1% e 7.0%), per sfiducia nelle forze dell'ordine (rispettivamente, 5.9% e 8.0%) e, nel caso di violenza interna alla coppia, perchè non volevano che il *partner* venisse arrestato (13.8%)<sup>3</sup>. Da qui si intuisce il ruolo cruciale della consapevolezza - femminile e maschile - in merito a questo fenomeno e a tutte le sue possibili declinazioni. Il maltrattante, infatti, non è mai una persona che si presenta come tale, è anzi colui che «[...] inizia con delle premesse molto speciali ed apparentemente positive, che però conducono verso un'esclusività e una dinamica che va sempre più verso il controllo» (Cristina Carelli - coordinatrice generale CADMI<sup>6</sup> - in "Un altro domani", 2023). Al tempo stesso, queste persone non riescono a pensarsi diversamente in un rapporto, se non in quella modalità controllante. E, anzi, ritengono che quella modalità controllante sia un elemento di qualità della relazione, perchè la vivono come una dimostrazione d'affetto nei confronti del *partner* (Giulini<sup>7</sup> in "Un altro domani", 2023). Si tratta di persone che non sono state educate ai sentimenti e per questo è un problema culturale. Roia prosegue: «La violenza di genere è un fenomeno strutturale, ha radici sociali, culturali, trova un fattore genetico di concausa in leggi che sono arrivate troppo tardi [...]». Per tale motivo non si può considerare fatto isolato un qualsiasi episodio di violenza di genere; poichè questo, in realtà, si inquadra all'interno di un contesto culturale ben preciso, la cui origine è un «ritenuto e anche avvallato predominio di genere che l'uomo ha storicamente esercitato nei confronti delle donne». Una società civile deve quindi rendere prioritaria la prevenzione cosiddetta "primaria" delle condotte violente, deve cioè creare una cultura operativa che dia spazio alla sofferenza e alla vulnerabilità quando questa sofferenza e vulnerabilità dei soggetti possono diventare un qualcosa di lesivo per l'altro. «E questo significa dare spazio alla consapevolezza della gravità di queste condotte per chi le subisce, portare testimonianza su quanto è grave e su quanto può essere devastante per una persona avere esperienze di quel tipo, alla consapevolezza di quali sono i confini della nostra relazione con gli altri, di quali sono i limiti nei comportamenti che agiamo nei confronti dell'altro. Questa è prevenzione

---

<sup>6</sup>CADMI è l'acronimo della Casa di Accoglienza delle Donne Maltrattate di Milano.

<sup>7</sup>Paolo Giulini - criminologo clinico/presidenza CIPM (Centro Italiano per la Promozione della Mediazione).

primaria, non dare più per scontato tutto ciò. Noi oggi siamo arrivati ad un livello tale di costruzione di civiltà che possiamo permetterci di non dare più per scontato queste cose.» (Giulini in "Un altro domani", 2023).

### 1.1.1 Gli *Intimate Partner Homicides*

La violenza di genere trova quella che è la sua realizzazione più estrema negli omicidi di donne, e più precisamente, nei "femminicidi". Sebbene questo termine venga utilizzato nel linguaggio comune con una certa flessibilità, Russel e Harmes (2001) ne danno un'immediata definizione formale: «uccisione di donne, da parte di uomini, perchè donne»; ponendo l'accento fin da subito sulla connotazione della vittima nella sua dimensione di genere. In altre parole, il femminicidio può definirsi un omicidio commesso da un uomo nei confronti di una donna "in quanto tale" (Minello e Dalla Zuanna, 2020). Mentre con il termine *Intimate Partner Homicide* - IPH - si intende un omicidio commesso per mano di *partner* o *ex partner*; e dunque accaduto in un contesto altamente specifico che è quello di una relazione intima affettiva presente o pregressa (World Health Organization, 2012). Si noti che la definizione di IPH è generica in merito al sesso di autore e vittima. Tuttavia, i dati sugli omicidi italiani «nascondono una forte dinamica di genere», e, nello specifico,

- studiare gli omicidi di donne vuol dire focalizzarsi sui delitti avvenuti in ambito familiare o nel contesto di relazione intima affettiva (IPH) (Minello e Dalla Zuanna, 2020);
- studiare gli IPH significa focalizzarsi prevalentemente su delitti di donne (nel 78% dei casi di IPH le vittime sono donne (UNODC, 2016)).

A completamento, si noti dunque che i femminicidi risultano solo in parte sovrapponibili con gli IPH femminili.

### 1.1.2 I Centri Antiviolenza

I Centri Antiviolenza - insieme alle Case Rifugio - costituiscono «il fulcro della rete territoriale di presa in carico delle donne vittime di violenza. Si tratta di servizi specializzati che lavorano sulla base di una metodologia dell'accoglienza basata su un approccio di genere e sui principi della Convenzione di

Istanbul»<sup>8</sup>.

In Italia, centri di questo tipo hanno subito una forte evoluzione nel corso del tempo. Le prime esperienze di assistenza alle vittime risalgono agli anni '70, quando la diffusione del movimento femminista porta alla nascita di gruppi di autocoscienza femminile. Ed è proprio in quegli anni che i dibattiti interni al movimento creano una forte consapevolezza in merito al fenomeno della violenza di genere. In un contesto storico-culturale complesso da questo punto di vista - si pensi che la legge contro la violenza sessuale risale al tardo 1996 - sorgono dunque i primi centri antiviolenza. Tuttavia, solo a partire dagli anni '90, se ne registra una crescita, seppure contenuta: nell'ordine di una settantina di realtà autogestite che collaborano tra loro sul territorio nazionale. Parallelamente si assiste anche all'evoluzione del quadro giuridico nazionale - e internazionale - in materia di violenza di genere, ma rimane un'evoluzione che "rincorre". Risale infatti al 2011 - con formale adozione da parte dell'Italia nel 2013 - la Convenzione del Consiglio d'Europa sulla prevenzione e la lotta alla violenza contro le donne e la violenza domestica, primo trattato internazionale con l'intento di disporre di uno strumento giuridico volto a proteggere le donne contro qualsiasi forma di violenza. Ed è a partire dal 2017 - con la nascita del progetto di monitoraggio dei centri antiviolenza, finanziato dal "piano di azione nazionale contro la violenza di genere" del Dipartimento per le Pari Opportunità - che il ruolo dello Stato italiano comincia a farsi rilevante, in termini di risorse e di favoreggiamento all'istituzione di nuovi luoghi di assistenza. Ad oggi, i centri antiviolenza sono divenute realtà strutturate, che giocano un ruolo essenziale nella strategia nazionale di prevenzione e protezione dalla violenza di genere.

L'operato dei centri antiviolenza (CAV) si fonda su tre principi: anonimato, riservatezza e gratuità delle prestazioni. Il CAV può pertanto definirsi un "servizio" pubblico e gratuito, dedicato a qualsiasi donna necessiti di aiuto, consulenza e/o di interventi di protezione tempestiva, in quanto vittima di violenza. Tuttavia, il termine "servizio" tra virgolette è rilevante: «i centri antiviolenza non sono semplicemente dei servizi, bensì sono considerati dei luoghi in cui si esercita sia progettualità politica che culturale, in cui il nocciolo è dato dalla libertà femminile» (Ciaravolo, 2021). Ciò, dunque,

---

<sup>8</sup><https://www.istat.it/it/violenza-sulle-donne/la-fuoriuscita-dalla-violenza/centri-antiviolenza>

appare in linea con il fatto che le figure operanti in questi centri debbano essere necessariamente donne. Questo è fondamentale poiché «il rapporto che si instaura tra donne facilita l'*empowerment* della donna vittima di violenza, permette il raggiungimento di una maggiore consapevolezza di sé e sopprime il concetto di potere, generalmente impresso nella figura dell'uomo, che caratterizza queste dinamiche violente» (Carelli in "Un altro domani", 2023).

## Capitolo 2

# L'analisi di interesse

In questo capitolo si definisce formalmente l'obiettivo d'analisi e, a seguire, si illustrano nel dettaglio i dati a disposizione e le operazioni che sono state impiegate su questi ultimi affinché fosse condurre l'indagine di interesse.

### 2.1 Obiettivo dell'analisi

Lo studio del ruolo dei centri antiviolenza nell'evoluzione degli *Intimate Partner Homicides* (IPH) italiani attraverso analisi descrittive e modelli statistici parametrici adeguati rappresenta l'obiettivo fondante di questa analisi.

### 2.2 I dati a disposizione

La Professoressa Minello ha messo a disposizione per l'analisi di interesse tre *dataset*. Nello specifico, *Autori* e *Vittime* contengono informazioni, validate dal Ministero dell'Interno, in merito alla totalità degli omicidi avvenuti in Italia dal 1992 al 2016. I dati in questione provengono direttamente dai registri delle forze dell'ordine, documenti d'archivio ufficiali nei quali è previsto che vengano riportate tutte le denunce di omicidio. Si tratta pertanto di informazioni altamente sensibili il cui accesso in forma disaggregata è stato consentito ai Professori Minello e Dalla Zuanna per scopi scientifici, in particolare per la redazione di "Morire in Italia: omicidi di donne, omicidi in famiglia" (2020).

All'interno del *dataset Autori* sono presenti 49 variabili relative agli omicidi ed in particolare agli autori di tali reati. Di seguito se ne da una bre-

ve illustrazione raggruppandole, per una maggiore chiarezza espositiva, in macro classi:

- generiche coordinate spazio-temporali relative all'evento, quali data e luogo dell'omicidio;
- informazioni relative alla dinamica dell'evento, quali ambito criminale, circostanza, movente, arma/i, numero di persone coinvolte, etc.;
- informazioni anagrafiche in merito ai singoli autori, quali data di nascita, genere, nazionalità, etc.;
- variabili create a partire da quelle inizialmente disponibili (ad es. età dell'autore o tipologia di omicidio sulla base del movente).

Inoltre si evidenzia fin da subito la presenza di una variabile che ricoprirà un ruolo fondamentale per l'analisi di interesse, ossia il tipo di relazione che lega la vittima (o le vittime) all'autore (o agli autori).

All'interno del *dataset Vittime* sono presenti 23 variabili relative agli omicidi ed in particolare alle vittime di tali reati. Nello specifico, si ritrovano le generiche coordinate spazio-temporali e le informazioni relative alla dinamica dell'evento presenti anche in *Autori*, alle quali si aggiungono le analoghe informazioni anagrafiche in merito alle singole vittime.

Il terzo *dataset* - *CAV* - risulta dall'elaborazione personale della Professoressa Minello e riguarda i centri anti violenza. Più precisamente, a partire dagli anni di apertura di questi ultimi, è stata creata una tabella a doppia entrata con frequenze risultanti che corrispondono al numero di centri italiani presenti nelle diverse combinazioni di Provincia - Anno. L'arco temporale preso in considerazione va dal 1965 al 2017, sovrapponendosi ampiamente a quello degli altri due *dataset*.

## 2.3 Operazioni sui dati a disposizione

Alla luce dell'obiettivo di analisi (§ 2.1) e dei dati a disposizione (§ 2.2), appare evidente la necessità di isolare gli IPH dal resto degli omicidi, in modo da poterne studiare prima la dinamica a se stante e poi l'evoluzione in relazione alla presenza dei centri anti violenza.

Da un punto di vista teorico si è partiti dalla definizione stessa di IPH (vedi § 1.1.1) declinandola ulteriormente - ai fini dell'analisi - agli omicidi di donne per mano di uomini, accaduti in un contesto di relazione intima affettiva presente o pregressa. Tuttavia, nella pratica, l'individuazione di questa tipologia di omicidi risulta particolarmente complessa. Di seguito si elencano i principali problemi riscontrati:

- *Autori* e *Vittime* si presentano come due *dataset* almeno parzialmente distinti per struttura e contenuto;
- la presenza di più autori, vittime e/o armi del delitto da origine ad un numero considerevole di duplicati;
- l'informazione in merito al tipo di relazione che intercorre tra l'autore e la vittima comincia ad essere registrata a partire dal 1999.

Per poter comprendere le procedure risolutive adottate è necessario chiarire ulteriormente quella che è la struttura dei due *dataset* relativi alla totalità degli omicidi italiani. *Autori* e *Vittime* condividono alcune informazioni (ad esempio, il codice identificativo dell'omicidio) mentre altre sono, per così dire, "esclusive" dei singoli insiemi di dati. E se la selezione degli IPH di interesse dipende proprio da quest'ultime, si intuisce fin da subito che una tale conformazione distinta può rivelarsi piuttosto problematica. La soluzione più ovvia sembrerebbe quella di unire i due *dataset* con l'ausilio di una variabile comune, tuttavia c'è un altro aspetto che va a complicare ulteriormente la situazione impedendo una semplice operazione di *merge*. La presenza di duplicati in *Autori* deriva dalla registrazione di più autori e/o armi del delitto mentre in *Vittime* questa stessa presenza deriva dalla registrazione di più vittime e/o armi del delitto. Ed è proprio tale asimmetria a dare origine ad una numerosità differente dei due *dataset*. Si rivela dunque necessaria una procedura in più step:

1. operazioni preliminari sui *dataset* distinti;
2. risoluzione del problema delle "relazioni mancanti";
3. risoluzione del problema dei duplicati  
e conseguente unione dei due *dataset*.

### 2.3.1 Operazioni preliminari sui *dataset* distinti

Nel corso di questa prima fase si opera sui singoli *dataset* con l'obiettivo di indagare tutte le variabili a disposizione ed effettuare, di conseguenza, una selezione preliminare di eventuali variabili e osservazioni.

Le prime operazioni effettuate su *Autori - dataset* con 16692 osservazioni e 49 variabili - consistono nella selezione dei soli uomini (tramite la variabile Sesso) e nell'esclusione di tutti gli omicidi riconducibili alla criminalità organizzata o ad atti di terrorismo (tramite la variabile Gruppo). Si procede poi alla gestione dei dati mancanti, ove possibile: l'arco temporale viene ridotto a 1992-2015 poichè le osservazioni relative all'ultimo anno di rilevazione (2016) presentano NA in corrispondenza della quasi totalità delle variabili. Argomento determinante a sostegno di questa scelta è l'assenza dei codici identificativi riferiti ai singoli autori (variabile Persona) e omicidi (variabile Coppie), il cui controllo incrociato avrebbe permesso una ricostruzione quantomeno parziale dei delitti perpetrati. A titolo esemplificativo, nel caso di tre record con comune codice Coppie e diverso codice Persona, sarebbe possibile affermare con certezza che si sta facendo riferimento allo stesso omicidio commesso tuttavia da almeno<sup>1</sup> tre autori differenti. Esaminando ulteriormente il *dataset* si sono rese necessarie imputazioni a partire da altre variabili<sup>2</sup> e creazioni di modalità aggiuntive, entro cui collocare i dati effettivamente mancanti di variabili categoriali. L'insieme delle operazioni sopracitate consente di pervenire ad una riduzione<sup>3</sup> significativa dei valori mancanti, oltre che ad una gestione più utile e funzionale di quelli rimasti. Seguono poi una serie di altre azioni preliminari volte a rendere il *dataset* *Autori* pronto per la fase successiva. Nello specifico, si eliminano le variabili costanti, quelle con un numero di modalità troppo elevato (ad esempio, la variabile Luogoevento presenta 2792 modalità) e quelle ritenute portatrici di informazione "ridondante" (ad esempio, le variabili relative alla nazionalità

---

<sup>1</sup>Si ricorda la previa operazione di selezione in base al genere che ha escluso tutte le autrici donne dal *dataset*. Ne deriva dunque la perdita della totalità degli omicidi con un solo autore di genere femminile ma anche l'esclusione di singoli record nel caso di delitti con più autori, di cui almeno uno di essi è di genere maschile.

<sup>2</sup>Operazioni di questo tipo sono realizzabili grazie alla configurazione del *dataset* stesso che contiene "versioni diverse" delle stesse variabili. Ad esempio, accorpando le modalità della variabile Arma con riferimento alle categorie Marzio, si ri-ottiene la variabile Armacod; è perciò possibile fare imputazione di valori in Armacod a partire da Arma.

<sup>3</sup>Come anticipato in § 2.3 in merito alla relazione che lega vittima e autore, alcuni NA sono presenti in maniera sistematica; pertanto l'eliminazione di tutte le osservazioni con dati mancanti comporterebbe una pericolosa distorsione dei risultati.



dell'autore - Nazaut e Nazautore - coincidono). Si crea inoltre la variabile relativa all'età dell'autore (Eta\_autore) a partire dalle informazioni disponibili (variabili Datanascita e Dataevento).

Per *Vittime - dataset* con 18847 osservazioni e 23 variabili - la procedura è analoga e compatibile con quella appena descritta. Le prime operazioni effettuate consistono nella selezione delle sole donne (tramite la variabile Sesso) e nell'esclusione di tutti gli omicidi riconducibili alla criminalità organizzata o ad atti di terrorismo (tramite la variabile Gruppoambitocriminale). Si procede poi alla gestione dei dati mancanti. Quest'ultima, decisamente meno complessa rispetto alla precedente, si compone di un'unica operazione: creazione di modalità aggiuntive entro cui collocare i dati effettivamente mancanti delle variabili categoriali. L'arco temporale viene ridotto a 1992-2015 coerentemente con quanto fatto in *Autori*. Si sottolinea che all'interno di questo *dataset* le osservazioni relative all'ultimo anno di rilevazione (2016) non presentavano un numero di NA particolarmente elevato (come invece accadeva in *Autori*); tuttavia si procede comunque alla riduzione del periodo poichè non sarebbe possibile alcuna ricostruzione degli omicidi con le sole informazioni disponibili in *Vittime*. Seguono poi una serie di altre azioni preliminari volte a rendere il *dataset* pronto per la fase successiva. Nello specifico, si eliminano le variabili costanti, quelle con un numero di modalità troppo elevato (ad esempio, la variabile Luogoevento) e quelle ritenute portatrici di informazione "ridondante" (ad esempio, la variabile qualitativa Straniero può essere ottenuta accorpando in tre modalità - straniero, italiano, mancante - il supporto della variabile Nazionalita). Si crea inoltre la variabile relativa all'età della vittima (Eta\_vittima) a partire dalle informazioni disponibili (variabili Datanascita e Dataevento).

Alla fine di questa prima fase, il *dataset Autori* contiene 11841 osservazioni e 30 variabili mentre il *dataset Vittime* contiene 4425 osservazioni e 15 variabili.

### 2.3.2 Risoluzione del problema delle "relazioni mancanti"

In questa seconda fase si vuole risolvere il problema dei dati mancanti - presenti in modo sistematico - nella variabile relativa al tipo di legame che intercorre tra vittima e autore.

Sono tuttavia d'obbligo le seguenti precisazioni tecnico-pratiche. L'informazione di interesse è contenuta all'interno della variabile Relazioneesoggetti,

presente in entrambi i *dataset*. Quest'ultima riesce a descrivere gran parte dei legami possibili grazie ad un numero piuttosto elevato di modalità: i 35 livelli individuano nel dettaglio relazioni familiari presenti e pregresse (ad esempio, Cognato e Ex-Cognato), affettive-intime presenti e pregresse (ad esempio, Coniuge e Ex-Coniuge), lavorative, amicali e addirittura indirette. Ciò detto lo studio degli IPH sembrerebbe sovrapporsi perfettamente alla trattazione delle sole modalità che identificano relazioni affettive-intime presenti o pregresse. Tuttavia, da un punto di vista pratico, emergono alcune criticità. La prima - ormai ben nota - corrisponde al problema dei dati mancanti: l'informazione in merito al tipo di legame tra vittime e autore dell'omicidio comincia ad essere rilevata a partire dal 1999. C'è poi una questione relativa all'attendibilità dell'informazione: trattandosi di dati provenienti da documenti ufficiali (si veda § 2.2), il coniuge viene registrato come ex coniuge solo in seguito all'emissione di una sentenza di divorzio da parte dell'autorità giudiziaria italiana. Dunque il fenomeno delle relazioni pregresse tende ad essere sottostimato per costruzione. In ultimo c'è una problematica riguardante le modalità di diffusione dei risultati: per questioni legate alla sensibilità del dato, informazioni di questo tipo possono essere diffuse solo in forma aggregata. Diventa pertanto obbligatorio rifarsi alla classificazione ISTAT delle tipologie di legame potenzialmente intercorribili tra vittima e autore dell'omicidio. Questa prevede sei macro categorie: *partner*, *ex partner*, parente, altro conoscente, autore sconosciuto alla vittima e autore non identificato; ed è già presente all'interno del *dataset Autori* con l'etichetta "Istatrel".

Riassumendo: l'informazione specifica sul tipo di relazione che lega vittima e autore (variabile *RelazioneSoggetti*) deve essere necessariamente inserita in una categoria più ampia, a prescindere dall'eventuale presenza (sistemica e/o casuale) di dati mancanti in merito. Per di più, l'obiettivo di selezionare gli IPH restringe il campo delle classi di interesse alle sole *partner* e *ex partner* ma è noto che quest'ultime - così distinte - non risultino particolarmente affidabili per motivi legati alle modalità di raccolta dei dati. Di conseguenza, lo studio degli IPH sembrerebbe sovrapporsi piuttosto alla trattazione delle relazioni affettive-intime senza alcuna declinazione più specifica in merito.

Si noti che quanto detto suggerisce un'idea di procedura. Se l'obiettivo è quello di individuare e isolare le osservazioni caratterizzate da una qual-

siasi tipologia di relazione affettiva-intima, la soluzione al problema dei dati mancanti non si ottiene tentando di imputare un legame specifico (ossia, una modalità della variabile Relazione-soggetti o Istatrel) bensì provando ad escludere tutte quelle osservazioni ritenute "incompatibili" con la tipologia di relazione di interesse. Da un punto di vista pratico, tale operazione viene effettuata utilizzando l'informazione presente all'interno dei *dataset*. Se ne riporta di seguito un elenco distinto per *Autori* e *Vittime*.

*Dataset Autori*. Si procede a:

1. l'esclusione di tutti gli omicidi caratterizzati da una relazione tra autore e vittima di tipo lavorativo o indiretto, tramite la variabile Femminicidio4 (variabile dicotomica con modalità "femminicidio" e "altro tipo di omicidio");
2. l'esclusione di tutti gli omicidi con eutanasia come movente, tramite la variabile Femminicidio3 (variabile categoriale relativa alla tipologia di omicidio);
3. la selezione dei soli omicidi con circostanza di interesse<sup>4</sup>, tramite la variabile Circostanza (variabile categoriale);
4. la selezione dei soli omicidi con movente di interesse<sup>5</sup>, tramite la variabile Movente (variabile categoriale);
5. l'esclusione di tutti gli omicidi caratterizzati da una relazione tra autore e vittima non di interesse per l'analisi, tramite la variabile Relazione-soggetti (vedi sopra).

*Dataset Vittime*. Si procede a:

1. la selezione dei soli omicidi con circostanza di interesse, tramite la variabile Circostanza;
2. la selezione dei soli omicidi con movente di interesse, tramite la variabile Movente;

---

<sup>4</sup>La circostanza di interesse - Familiari-Affettivi - è stata individuata esaminando che modalità della variabile (Circostanza) presentano gli omicidi caratterizzati da una relazione tra autore e vittima di tipo affettivo-intimo.

<sup>5</sup>I moventi di interesse sono stati individuati esaminando che modalità della variabile (Moenente) presentano gli omicidi caratterizzati da una relazione tra autore e vittima di tipo affettivo-intimo.

3. l'esclusione di tutti gli omicidi caratterizzati da una relazione tra vittima e autore non di interesse per l'analisi, tramite la variabile *Relazione* soggetti.

La bontà di una procedura di questo tipo deve essere necessariamente valutata. Una metodologia possibile consiste nel visualizzare graficamente l'evoluzione del numero di omicidi nel corso del tempo, al fine di scongiurare la presenza di andamenti anomali. Si sceglie tuttavia di effettuare questa valutazione in una fase successiva, nello specifico, sul *dataset* risultante dall'unione di *Autori* e *Vittime* (si rimanda a § 3-4).

Alla fine di questa seconda fase, il *dataset Autori* contiene osservazioni e variabili mentre il *dataset Vittime* contiene osservazioni e variabili.

### 2.3.3 Risoluzione del problema dei duplicati e unione dei due *dataset*

Nel corso di questa terza fase si opera in primis sui singoli *dataset* con l'obiettivo di risolvere il problema dei duplicati. Fatto ciò, l'unione di *Vittime* e *Autori* in un unico insieme di dati diventa possibile attraverso una semplice operazione di *merge*.

Si chiarisce innanzitutto il concetto di duplicato. Il termine "duplicato" indica un'osservazione che si ripete due o più volte all'interno del *dataset* e che presenta dunque valore comune almeno in corrispondenza della variabile identificativa dell'unità statistica (in questo caso, variabile *Coppie*). Nello specifico, come già anticipato in § 2.3, la presenza di osservazioni ripetute in *Autori* deriva dalla registrazione di più autori e/o armi del delitto mentre in *Vittime* questa stessa presenza deriva dalla registrazione di più vittime e/o armi del delitto. Ed è proprio tale asimmetria che da origine ad un numero di duplicati differente: 83 in *Autori* e 138 in *Vittime*<sup>6</sup>.

Con l'intento di pervenire ad un *dataset* composto da un unico record per ogni omicidio, si configurano procedure di risoluzione diverse in base alla tipologia del duplicato: più autori, più vittime, più armi o combinazioni delle precedenti. Ne consegue la necessità di determinare con certezza la "causa" che sta all'origine delle osservazioni ripetute, così da poterne tenere conto in

---

<sup>6</sup>Per determinarne l'entità del fenomeno da un punto di vista pratico si procede al conteggio dei codici identificativi degli omicidi (variabile *Coppie*) con frequenza assoluta maggiore di uno.

maniera adeguata. Questo tipo di operazione si declina in un esame incrociato dei due *dataset*, e più precisamente, in un controllo dei valori assunti dalle variabili Arma (arma del delitto), Codicearianna (codice identificativo della vittima), Coppie (codice identificativo dell'omicidio) e Persona (codice identificativo dell'autore). Il criterio generale a cui si fa riferimento - indipendentemente dalla tipologia di duplicato - deriva dalla seguente riflessione. Se lo studio degli IPH rappresenta l'obiettivo finale, allora esiste una sola osservazione di interesse nell'insieme delle osservazioni ripetute: la combinazione autore - vittima caratterizzata da un legame di tipo affettivo-intimo è unica. In un omicidio con una vittima e più autori, ad esempio, solo uno tra gli autori coinvolti può intrattenere con la vittima la tipologia di legame ritenuto di interesse per l'analisi. Tuttavia, anche record di questo tipo non sono immuni al problema dei duplicati che derivano dalla registrazione di più armi.

Si procede a illustrare le procedure di risoluzione adottate in base alla tipologia di duplicato.

*Duplicato dovuto alla presenza di più autori*

1. Si seleziona il record di interesse nell'insieme delle osservazioni ripetute con l'ausilio delle variabili presenti nel *dataset*<sup>7</sup>.
2. Si verifica la correttezza del valore assunto dalla variabile relativa al numero totale di autori (variabile Numautcoppia) per non perdere l'informazione utile apportata dal duplicato.

*Duplicato dovuto alla presenza di più vittime*

1. Si seleziona il record di interesse nell'insieme delle osservazioni ripetute con l'ausilio delle variabili presenti nel *dataset*.
2. Si verifica la correttezza del valore assunto dalla variabile relativa al numero totale di vittime (variabile Numvittcoppia) per non perdere l'informazione utile apportata dal duplicato.

*Duplicato dovuto alla presenza di più armi del delitto.* Per questa tipologia di osservazioni ripetute, la procedura adottata deriva da una riflessione teorica

---

<sup>7</sup>Nella maggior parte dei casi l'operazione di selezione avviene sulla base della variabile RelazioneSoggetti. Tuttavia, in caso di dati mancanti, questa stessa operazione viene effettuata a partire da un esame incrociato delle variabili Circostanza, Movente, Eta\_Autore e Eta\_Vittima.

aggiuntiva: ogni arma è caratterizzabile attraverso un certo grado di pericolosità. Pertanto l'idea di soluzione consiste nell'individuare tra le opzioni l'arma "più pericolosa", ossia quella che con più probabilità ha causato la morte della vittima. In aggiunta, si crea una variabile dicotomica ( $n\_Armi$ ) che assume valore unitario in corrispondenza degli omicidi con più armi del delitto. Quest'ultima permette infatti di non perdere l'informazione utile fornita dal duplicato, tenendo conto di un'eventuale ruolo che può giocare una combinazione di armi rispetto al caso *standard* di un'arma unica. In sintesi:

1. Si seleziona il record caratterizzato dall'arma più pericolosa nell'insieme delle osservazioni ripetute.
2. Si procede all'imputazione del valore unitario in corrispondenza della variabile  $n\_Armi$  per non perdere l'informazione utile apportata dal duplicato.

Si noti che in presenza di più tipologie di duplicati basta seguire in successione un insieme selezionato delle procedure sopracitate.

Alla fine di questa fase, entrambi i *dataset* (*Autori* e *Vittime*) contengono 1643 osservazioni uniche<sup>8</sup>.

Prima di procedere all'effettiva operazione di *merge* si eliminano le variabili comuni da uno dei due dataset. L'eliminazione non è stata fatta in precedenza perchè - fino ad ora - queste variabili hanno giocato un ruolo fondamentale nei controlli incrociati volti a ricostruire le dinamiche degli omicidi. L'unica variabile esclusa dalla procedura appena citata è *Coppie*, ossia il codice identificativo dell'omicidio, dal momento che l'unione di *Autori* e *Vittime* in un unico insieme di dati viene effettuata proprio sulla base di tale variabile.

Il *dataset* risultante - ottenuto a seguito di numerose operazioni finalizzate a isolare gli IPH dal resto degli omicidi - contiene 1643 osservazioni e 16 variabili.

---

<sup>8</sup>Si noti che una volta risolto anche il problema dei duplicati, la numerosità dei due *dataset* finalmente coincide.

## Capitolo 3

# L'evoluzione degli IPH italiani

Questo capitolo è interamente dedicata allo studio degli *Intimate Partner Homicides* italiani attraverso un'analisi descrittiva del *dataset* che li contiene (§ 2.3).

Si ricorda che l'insieme di dati in questione è il risultato di numerose operazioni effettuate su due *dataset* di partenza ben più ampi, che includevano al loro interno informazioni in merito alla totalità degli omicidi avvenuti in Italia dal 1992 al 2015. Tali operazioni sono servite a isolare una specifica tipologia di omicidi: gli IPH, ossia omicidi di donne per mano di uomini accaduti in un contesto di relazione intima affettiva presente o pregressa. In particolare, si dispone di 1643 osservazioni sulle quali sono state rilevate 16 variabili:

- **codice\_IPH**: codice identificativo dell'IPH (variabile quantitativa)
- **anno\_evento**: anno di accadimento dell'IPH (variabile quantitativa con supporto  $\{1992, \dots, 2015\}$ )
- **regione\_evento**: regione in cui è avvenuto l'IPH (variabile qualitativa con 20 modalità corrispondenti alle 20 regioni italiane)
- **provincia\_evento**: provincia in cui è avvenuto l'IPH (variabile qualitativa con modalità corrispondenti alle province italiane nel periodo di rilevazione)
- **capoluogo**: corrispondenza tra la provincia e il luogo preciso in cui è avvenuto l'evento (variabile dicotomica: "sì" o "no")

- **rel\_ISTAT**: categoria entro cui rientra il tipo di relazione che intercorre tra la vittima e l'autore dell'IPH, secondo la classificazione ISTAT (variabile qualitativa con 3 modalità: "partner", "ex partner" o "mancante")
- **movente**: movente dell'autore dell'IPH (variabile qualitativa con 6 modalità)
- **cod\_arma**: categoria dell'arma del delitto secondo la classificazione Marzio (variabile qualitativa con 6 modalità)
- **n\_armi**: registrazione di più di un'arma del delitto (variabile dicotomica: "sì" o "no")
- **eta\_vittima**: età - in anni compiuti - della vittima all'evento IPH (variabile quantitativa)
- **nazionalita\_vittima**: nazionalità della vittima (variabile qualitativa con 16 modalità)
- **eta\_autore**: età - in anni compiuti - dell'autore all'evento IPH (variabile quantitativa)
- **nazionalita\_autore**: nazionalità dell'autore dell'IPH (variabile qualitativa con 16 modalità)
- **altri\_crim**: commissione di altri crimini da parte dell'autore dell'IPH (variabile dicotomica: "sì" o "no")
- **num\_vittime**: numero totale di vittime coinvolte (variabile quantitativa)
- **num\_autori**: numero totale di autori coinvolti (variabile quantitativa)

Un generico controllo delle variabili ha evidenziato la presenza di NA in **eta\_vittima** e **eta\_autore**<sup>1</sup>. Trattandosi di dati mancanti con numerosità relativamente ridotta, si procede in entrambi i casi all'imputazione del valore mediano. La differenza empirica osservata tra media e mediana suggerisce la

---

<sup>1</sup>Si ricorda che le età anagrafiche di vittime e autori sono state calcolate in § 2.3.1 a partire da variabili con valori mancanti. Tuttavia in quella fase non si è proceduto ad alcuna imputazione, poiché il valore medio o mediano si sarebbe riferito alla totalità degli omicidi e non ai soli IPH.



presenza di valori estremi; i quali, hanno - per costruzione - un minor effetto sul calcolo dell'età mediana.

Segue l'analisi descrittiva vera e propria, allo scopo di indagare le caratteristiche proprie degli IPH italiani ma anche la loro evoluzione nel corso del periodo considerato.

Come prima indagine si sceglie di riferire quella condizionata alla componente temporale. In questo caso l'obiettivo è duplice: nel corso del periodo 1992-2015 si vuole

- indagare l'evoluzione degli omicidi italiani con *focus* specifico sugli IPH;
- valutare se - ed eventualmente in che modo - le distribuzioni relative alle variabili disponibili differiscono in base all'anno di accadimento degli IPH.

Prima di esporre quanto ottenuto divengono necessarie alcune premesse. Un'analisi in funzione del tempo non può essere condotta ignorando il contesto storico entro cui si colloca. Quest'ultimo viene dunque schematizzato in Figura 3.1 con l'ausilio di una linea del tempo, popolata dagli eventi che potrebbero aver influito sull'andamento del fenomeno oggetto di studio.

Per di più è bene ricordare che i dati disponibili (*dataset Autori e Vittime*) avevano un problema di "relazioni mancanti" in modo sistematico: il tipo di legame tra vittima e autore comincia ad essere rilevato gradualmente a partire dal 1999. In fase di creazione del *dataset* contenente gli IPH sono state applicate una serie di operazioni volte a risolvere questa problematica (per un approfondimento in merito si rimanda a § 2.3.3); tuttavia la bontà della procedura seguita deve essere necessariamente valutata. Una strada possibile - come già anticipato in § .2.3.3 - consiste nel visualizzare graficamente proprio la dinamica temporale degli IPH, così da scongiurare la presenza di chiare anomalie nell'andamento. Si presterà dunque particolare attenzione a quest'ultimo aspetto durante l'analisi dei grafici che seguono.

Il primo *focus* viene posto sull'andamento degli omicidi italiani nel corso del periodo 1992-2015. Studi in merito hanno rilevato la presenza di un trend decrescente unito ad una forte dinamica di genere: «[...] questo trend decrescente è dovuto alla diminuzione degli omicidi di uomini, che storicamente sono sempre stati la parte più consistente degli omicidi. Il rischio

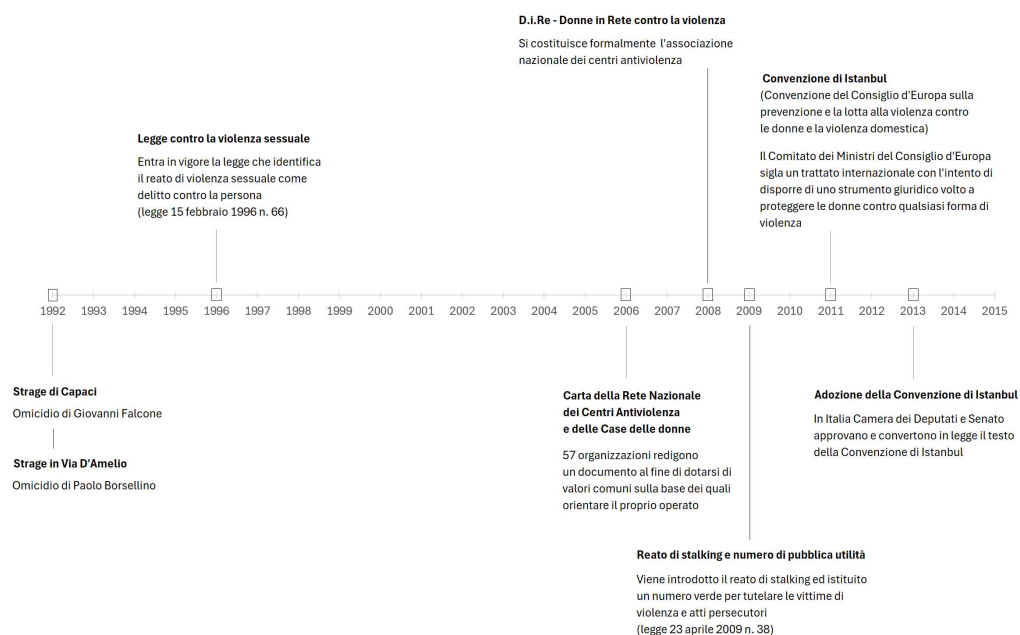


Figura 3.1: *Timeline* degli eventi potenzialmente rilevanti - Italia, 1992-2015.

per una donna di essere uccisa, partendo da livelli estremamente più bassi, nel corso degli anni è però diminuito in misura molto minore» (Minello e Dalla Zuanna, 2020). Un tale scenario trova conferma empirica anche a partire dai dati disponibili per questa analisi, come si può vedere in Figura 3.2. Ciò che risulta particolarmente interessante è il legame - individuato in letteratura - tra tipologia di omicidio e genere della vittima. Da un lato Minello e Dalla Zuanna (2020) operano la seguente distinzione: «A morire di criminalità organizzata sono prevalentemente gli uomini, le vittime di omicidi familiari sono invece prevalentemente le donne». Dall'altro, Barbagli e Colombo (2011) osservano un calo considerevole delle morti per mano di mafia, camorra e/o 'ndrangheta a fronte di una diminuzione ben più contenuta degli omicidi familiari. La lotta alla criminalità organizzata gioca un ruolo decisivo in questa dinamica poichè è proprio a partire dal 1992 - anno delle stragi di Capaci e Via D'Amelio - che la magistratura comincia a fare grandi passi avanti nel contrasto alle organizzazioni mafiose. Ciò detto, una sola conclusione appare immediata: quell'andamento decrescente con forte dinamica di genere che caratterizza gli omicidi italiani nel periodo 1992-2015 si motiva attraverso l'indagine della relazione tra tipologia di omicidio e sesso

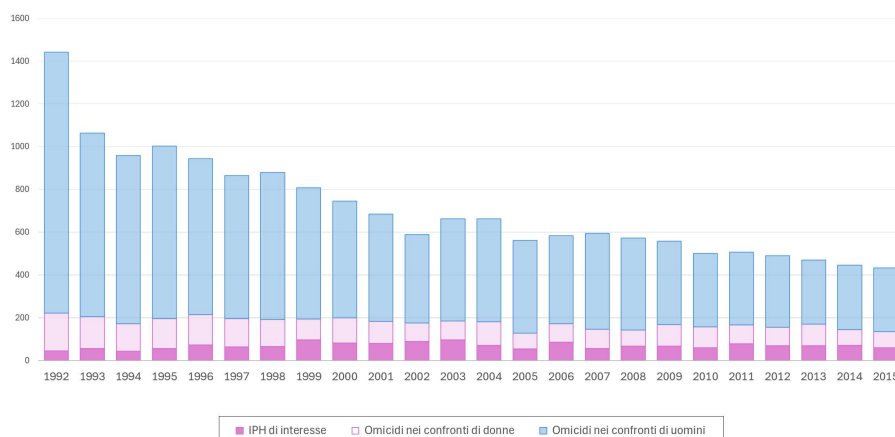


Figura 3.2: Numero totale di omicidi, stratificato per genere della vittima - Italia, 1992-2015.

In fucsia: proporzione di IPH rispetto alla totalità degli omicidi di donne.

della vittima.

In Figura 3.2 è possibile indagare anche l'evoluzione del numero di IPH (per il grafico in formato di serie storiche si rimanda a Figura 4.2 in Appendice). Questo andamento rimane crescente fino al 2004 per poi farsi oscillatorio attorno al valore medio di 67 omicidi l'anno. Tuttavia, un esame di quantità assolute non basta a descrivere in modo esaustivo la dinamica oggetto dell'analisi. Da qui l'obbligo di indagare versioni alternative: il numero di IPH rapportato al totale di omicidi commessi nei confronti di donne (Figura 3.3) e alla popolazione femminile ritenuta potenzialmente a rischio di sperimentare l'evento<sup>2</sup> (Figura 3.4). Nello specifico, l'incidenza sulla popolazione a rischio mostra un andamento in linea con quello osservato per il numero di IPH. Mentre se si indaga la percentuale ricoperta dagli IPH rispetto alla totalità degli omicidi di donne, un trend appena crescente sembra sostituirsi alla dinamica più stabile che caratterizzava il periodo 2005-2015 nelle Figure 3.4 e 4.2. È dunque un calo delle morti femminili per mano di autori "esterni" (ossia, non in relazione intima affettiva con la vittima) a determinare la lieve diminuzione nel numero degli omicidi di donne osservata in Figura 3.2.

<sup>2</sup>La popolazione cosiddetta "a rischio di sperimentare l'evento IPH" si compone di tutte le donne residenti in Italia con un'età non inferiore ai 14 anni. Tale limite è stato scelto individuando - all'interno del *dataset* di partenza - l'età minima registrata per le vittime di IPH.



Figura 3.3: Percentuale di IPH rispetto alla totalità degli omicidi di donne - Italia, 1992-2015.



Figura 3.4: Proporzione di IPH rispetto alla totalità di donne a rischio - Italia, 1992-2015.

Come noto, la dinamica evolutiva degli IPH che si registra a partire dalle informazioni disponibili può subire diverse tipologie di condizionamenti. Il primo risiede nell'influenza caratterizzante - per definizione - i dati in formato di serie storiche, ossia nell'effetto del contesto storico e socioculturale entro cui si colloca il fenomeno oggetto di studio (per una rappresentazione schematica degli eventi storici potenzialmente rilevanti si veda Figura 3.1). Mentre il secondo condizionamento trova origine proprio nei limiti strutturali dei *dataset* disponibili e si traduce, a livello pratico, nell'effetto della procedura di imputazione adottata per risolvere il problema delle "relazioni mancanti" (si veda § 2.3.3). Un'analisi esaustiva deve dunque valutare modalità e misura di tali influenze. Per quanto riguarda quella potenzialmente esercitata dal contesto storico, sembra che gli unici eventi in grado di sortire effetti importanti sulla dinamica degli omicidi italiani - e di riflesso, su quella degli IPH - siano le stragi di Capaci e Via D'Amelio. Tuttavia è d'obbligo una precisazione: quanto affermato non vuole escludere in alcun modo la possibilità di influenza da parte di altri eventi, quali leggi e operati dei centri antiviolenza volti a contrastare qualsiasi forma di aggressività contro le donne. L'effetto di questi ultimi potrebbe infatti svelarsi in forme meno evidenti e/o in tempi più dilatati. Nello specifico, si ricorda che indagare il ruolo dei centri antiviolenza nell'evoluzione degli IPH rappresenta l'obiettivo fondante dell'intera analisi; motivo per il quale si rimanda una tale discussione alla sezione dedicata (§ 4). Segue un'esame della possibile influenza da parte del contesto socioculturale, inteso come quadro che concerne insieme le strutture di una data società e i correlati aspetti culturali<sup>3</sup>. Con il loro andamento piuttosto stabile - che si protrae per gli ultimi 11 anni di rilevazione - e la loro crescita in termini di percentuale ricoperta sul totalità degli omicidi di donne, i dati disponibili sugli IPH confermano quanto asserito in letteratura: « La violenza di genere è un fenomeno che in Italia è strutturale: ha radici sociali, culturali, trova un fattore genetico di concausa in leggi che sono arrivate troppo tardi. [...] Questo si inquadra in un contesto ben preciso la cui origine è un ritenuto e anche avvallato predominio di genere che l'uomo ha storicamente esercitato nei confronti delle donne. » (Roia in "Un altro domani", 2023). Si valuta, infine, la procedura di imputazione delle "relazioni mancanti". Come noto, il tipo di legame tra vittima e autore comincia ad essere rilevato gradualmente a partire dal 1999. L'assenza di

---

<sup>3</sup><https://www.treccani.it/vocabolario/socioculturale/>

chiare anomalie nella dinamica evolutiva degli IPH (Figura 4.2) permette di escludere fin da subito un "effetto procedura" importante. Tuttavia, i valori registrati per alcuni anni - 1992, 1994, 1999, 2003 - inducono a pensare che possa esserci comunque del condizionamento alle operazioni eseguite durante la fase di imputazione. Da qui la decisione di operare in modo distinto a seconda dell'obiettivo specifico per il quale si conduce l'analisi. Avendo escluso la presenza di un forte "effetto procedura",

- in questa sezione - dedicata a indagare le caratteristiche proprie degli IPH italiani e la loro evoluzione nel tempo - si procede con l'analisi descrittiva dell'intero insieme di dati (periodo di rilevazione: 1992-2015);
- nella sezione successiva - dedicata a indagare il ruolo dei centri anti-violenza nell'evoluzione degli IPH - si sceglie di condurre due analisi parallele, utilizzando insiemi di dati relativi a periodi differenti (nello specifico, 1992-2015 e 2004-2015), così da valutare l'entità dell' "effetto procedura".

Il secondo *focus* è posto sulla dinamica evolutiva delle variabili disponibili. Un esame approfondito di tutte le distribuzioni condizionate alla componente temporale permette di individuare quattro gruppi:

1. variabili la cui distribuzione rimane pressochè invariata nel tempo: `regione_evento`, `capoluogo`, `cod_arma`, `n_armi`, `num_vittime` e `num_autori`;
2. variabili la cui distribuzione cambia nel tempo a causa di valori che si rivelano mancanti in maniera sistematica: `movente`;
3. variabili la cui distribuzione cambia nel tempo a causa di una modifica nell'insieme delle modalità: `provincia_evento`;
4. variabili la cui distribuzione cambia leggermente nel tempo: `eta_vittima`, `eta_autore`, `altri_crim`, `nazionalita_vittima` e `nazionalita_autore`.

Sono d'obbligo alcune precisazioni. In merito al movente dell'autore di IPH: 3 (su 6) modalità cominciano a registrare frequenze non nulle solo a partire dal 2000. In merito alla provincia di accadimento dell'omicidio: nel 2002 l'istituzione di 4 nuove province (Barletta-Andria-Trani, Monza e Brianza, Fermo e Sud Sardegna) modifica quella che era la precedente suddivisione

territoriale.

Poichè l'obiettivo è comune - indagare gli IPH in funzione delle variabili disponibili - si sceglie di inserire lo studio dell'andamento temporale registrato dai casi rilevanti (gruppo 4.) all'interno dell'analisi descrittiva che segue.

Il primo risultato si riferisce all'entità del fenomeno di interesse: con riferimento al periodo 1992-2015, gli IPH commessi da uomini nei confronti di donne costituiscono "solo" il 9.65% degli omicidi italiani. Questa proporzione rispetto al totale - già intuibile da un primo confronto tra numerosità iniziali dei *dataset Autori* e *Vittime*<sup>4</sup> e numerosità delle osservazioni di interesse - conferma in maniera evidente la rarità del fenomeno oggetto di studio. Si ricorda infatti che l'IPH può essere interpretato come un caso particolare di femminicidio, a sua volta solo in parte sovrapponibile con il generico omicidio nei confronti di donne (per un approfondimento si veda § 1.1.1). Purtroppo, a partire dai dati a disposizione per l'analisi, non è possibile quantificare l'ammontare assoluto o relativo dei femminicidi. Si possono tuttavia rapportare gli IPH commessi da uomini nei confronti di donne rispetto alla totalità degli omicidi nei confronti di donne, per avere un'ulteriore indicazione quantitativa in merito. Con riferimento al periodo 1992-2015 emerge che gli IPH di interesse costituiscono il 39% degli omicidi nei confronti di donne. A fini esplicativi, la Figura 3.5 sintetizza in un unico grafico tutti i precedenti risultati, descrivendo la composizione degli omicidi italiani per genere delle vittime ed evidenziando, nello specifico, la percentuale ricoperta dagli IPH di interesse rispetto alla totalità degli omicidi nei confronti di donne.

Una volta verificata la rarità del fenomeno oggetto di studio anche dal punto di vista empirico, si conduce l'analisi descrittiva vera e propria. Segue dunque una selezione di grafici con l'obiettivo comune di indagare gli IPH in funzione delle variabili disponibili.

Fin da subito appare evidente una disparità tra le regioni italiane (vedi Figura 3.6): la Lombardia si distingue in maniera significativa, collocandosi al primo posto per numero di IPH registrati nel corso del periodo di riferimento (262 su 1643, ovvero il 16%). Segue la Sicilia con quasi il 10% degli IPH

---

<sup>4</sup>Si noti che il totale delle vittime (17018) non coincide con la numerosità iniziale dei *dataset* a disposizione poichè quest'ultima fa riferimento ad un periodo più ampio e, al tempo stesso, include duplicati (vedi § 2.3).

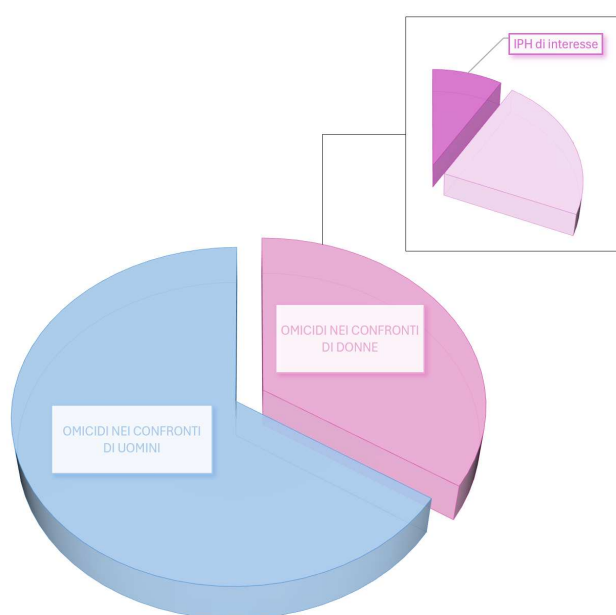


Figura 3.5: Omicidi per genere della vittima - Italia, 1992-2015.  
Nel riquadro (in alto a destra): IPH di interesse rispetto alla totalità degli omicidi nei confronti delle donne.



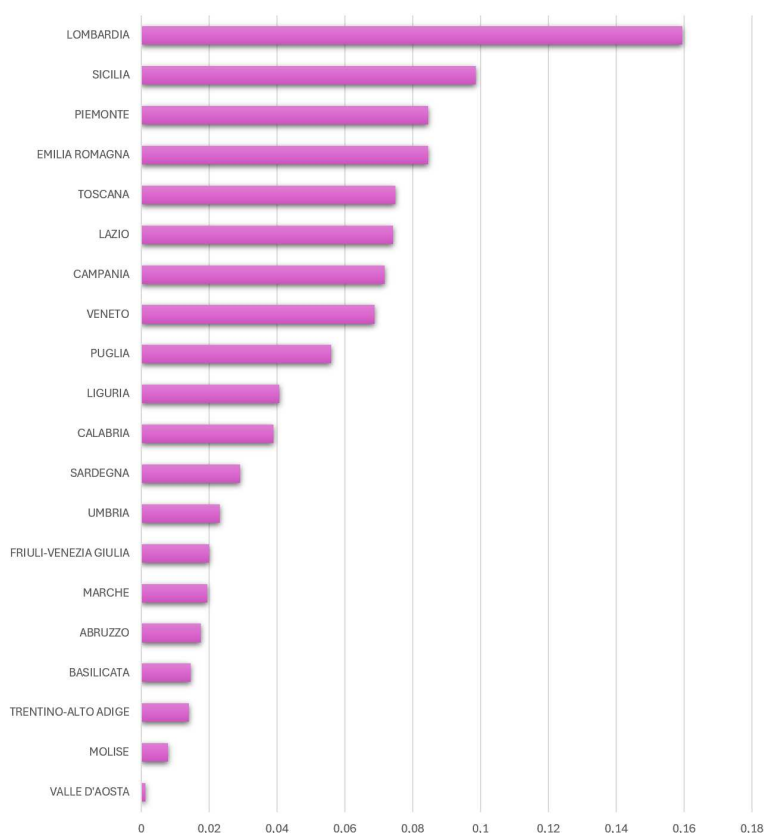


Figura 3.6: IPH per regione (proporzione rispetto al totale di IPH) - Italia, 1992-2015.

totali, poi il Piemonte e l'Emilia-Romagna con una percentuale condivisa dell'8%. Le proporzioni minori - al di sotto dell'1% - sono invece registrate, nell'ordine, da Molise (13 IPH) e Valle d'Aosta (2 IPH).

Un altro dato interessante: solo il 32,5% degli IPH interessa le città italiane che fanno provincia. Pertanto, con riferimento al periodo 1992-2015, è possibile affermare che omicidi di questo tipo tendono a essere commessi in paesi e cittadine piuttosto che in grandi città. La spiegazione a tale fenomeno dipende sicuramente dall'insieme di più fattori, di seguito se ne citano alcuni possibili.

- La minor densità di popolazione che caratterizza paesi e cittadine determina una modalità di "abitare" diversa rispetto a quella che si sperimenta nelle città grandi: ad una preponderanza di appartamenti e

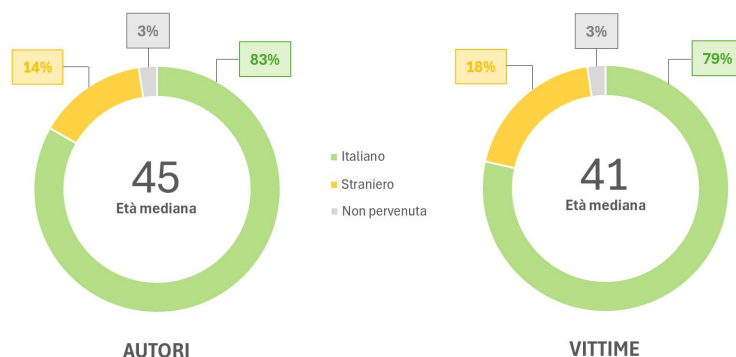


Figura 3.7: Età mediana e nazionalità di autori (sinistra) e vittime (destra) - Italia, 1992-2015.

villette a schiera si sostituiscono case medio-grandi, spesso anche molto distanti tra loro, ad uso di un singolo nucleo familiare. Una tale conformazione potrebbe favorire l'accadimento di IPH poiché riduce le possibilità di intervento da parte di vicini e/o condomini.

- Le città grandi ospitano molte persone che vi si trasferiscono per motivi di studio e/o lavoro. Dunque, partendo dal presupposto che la motivazione per cui ci si sposta possa definire un certo ordine delle priorità, ha senso ipotizzare che tali individui siano - almeno temporaneamente - focalizzati sulla propria carriera (universitaria e/o lavorativa) piuttosto che sulle dinamiche relazionali di coppia.

Si prosegue con l'analisi delle due figure coinvolte: autori (uomini) e vittime (donne). Le informazioni disponibili in merito sono state sintetizzate in Fig. 3.7.

Gli autori di IPH hanno un'età compresa tra 17 e 97 anni, con un valore mediano pari a 45; mentre per le vittime l'intero supporto appare leggermente spostato verso età più giovani: 14-95 anni, con un valore mediano pari a 41. La differenza (di circa due anni) che si osserva tra media e mediana suggerisce una presenza di *outliers*, confermata empiricamente da ulteriori analisi esplorative (ad esempio, si veda Fig. 4.3 in Appendice). È dunque l'età mediana, per costruzione meno sensibile a valori estremi, l'indicatore di sintesi più idoneo all'analisi. Da un esame approfondito in merito alla composizione per età di vittime e autori (si veda Figura 3.8 e si vedano Figure

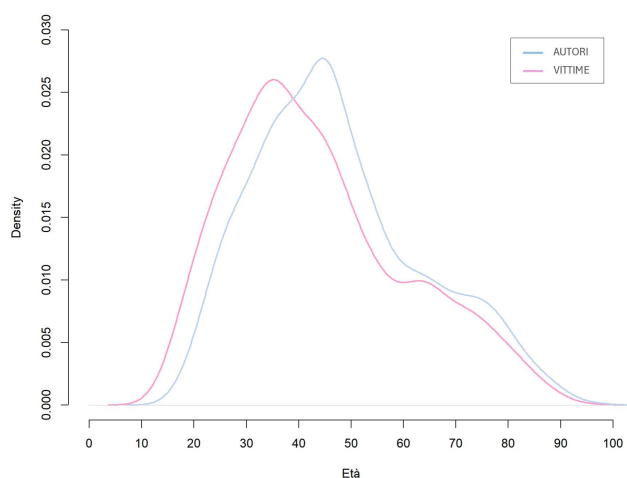


Figura 3.8: Densità di probabilità relativa alla variabile età di autori (linea azzurra) e vittime (linea rosa) - Italia, 1992-2015.

4.4 e 4.3 in Appendice) emergono risultati interessanti: la forma delle due distribuzioni - seppur traslata - si rivela molto simile. Con riferimento agli autori che hanno un'età compresa tra 25 e 80 (corrispondenti all'89% della totalità), l'85% non supera i 65 anni. Proporzioni simili si osservano per le vittime: delle donne con un'età compresa tra 20 e 75 (corrispondenti all'89% della totalità) l'84% non supera i 60 anni. Ciò detto, il fenomeno sembra riguardare particolarmente specifiche e distinte classi d'età: [35-50] per gli autori e [30-45] per le vittime. A questi risultati è possibile affiancarne altri ancor più informativi mettendo in relazione i singoli coinvolti in uno specifico IPH. In particolare emerge che l'89% delle vittime ha un'età inferiore, uguale o superiore al più di cinque anni rispetto a quella del proprio omicida. La maggior parte (1200 su 1643) ricade tuttavia nella prima casistica, delineando uno scenario preponderante: la donna vittima di IPH tende ad essere più giovane del proprio aggressore, sebbene in generale l'*age gap* sia ridotto<sup>5</sup>. Quest'ultima osservazione appare in linea con quanto ci si potrebbe aspettare a livello teorico, poichè, trattandosi di omicidi avvenuti in un contesto di relazione intima affettiva, le differenze in età osservate tra vittime e autori corrispondono a quelle che sono le differenze in età esistenti tra *partners*. A

<sup>5</sup>Dall'analisi si evince che nel 49% dei casi la differenza d'età tra vittima e aggressore non supera i 5 anni.

completamento, si ricorda la necessità di un esame condizionato alla variabile temporale. Questa indagine rileva una crescita del valore mediano all'evento IPH (si veda Figura 4.5 in Appendice): da 36 (dato relativo al 1992) a 42 anni (dato relativo al 2015) per le vittime, da 44 (dato relativo al 1992) a 47 anni (dato relativo al 2015) per gli autori. Tuttavia, l'andamento oscillatorio e una mancata conoscenza in merito all'entità dell' "effetto procedura" (procedura di imputazione delle "relazioni mancanti"), inducono a limitarsi alla seguente verifica. I risultati dell'esame condizionato alla componente temporale sono in linea con quelli ottenuti in precedenza. Nello specifico,

- il fenomeno sembra riguardare particolarmente specifiche e distinte classi d'età: [35-50] per gli autori e [30-45] per le vittime;
- la donna vittima di IPH tende ad essere più giovane del proprio aggressore, sebbene in generale l'*age gap* sia ridotto.

Per quanto concerne invece la nazionalità delle figure coinvolte, è evidente fin dalle prime analisi una netta prevalenza di italiani. Nello specifico gli autori di IPH stranieri costituiscono solo il 14% della totalità, percentuale alla quale va ad aggiungersi un 3% di dati mancanti, cosicché più dell'80% risulta interamente popolato da uomini italiani. Proporzioni simili si osservano per le vittime che appaiono così ripartite: 3% di dati mancanti, 18% di donne straniere e 79% di donne italiane. Con i dati a disposizione (si rimanda all'elenco delle variabili contenute nel *dataset* di cui sopra) sarebbe possibile indagare in modo più dettagliato la cittadinanza di vittime e autori stranieri, tuttavia le numerosità esigue non permetterebbero di trarre risultati particolarmente affidabili in merito. Dunque, per completezza espositiva ci si limita ad illustrare le modalità delle variabili `nazionalita_vittima` e `nazionalita_autore`, segnalando che non sono presenti livelli con frequenza pari a zero e che, in entrambi i casi, la nazionalità straniera più diffusa risulta essere quella rumena. Le variabili qualitative in questione - già presenti nei rispettivi *dataset* di partenza - hanno le seguenti modalità: Albania, Cina, Italia, Marocco, Perù, Romania, Tunisia, EU 15 (Regno Unito, Germania, Francia, Belgio, Spagna o Austria), Altri Paesi Europei (Polonia, Portogallo, Repubblica Moldava, Svizzera, Bulgaria, Ucraina o Ungheria), Ex Jugoslavia (Croazia, Repubblica di Serbia, Slovenia, Bosnia-Erzegovina, Repubblica Macedonia o Kosovo), Altri Sviluppatori (Federazione Russia, Stati Uniti o Liechtenstein), Altri America Sud (Ecuador, Colombia, Argentina,

Brasile o Cile), Altri Asia (India, Pakistan, Filippine o Cambogia), Altri Africa (Senegal, Egitto, Capo Verde, Costa D'Avorio, Somalia, Zaire Repubblica Democratica del Congo, Nigeria, Sierra Leone, Algeria, Liberia, Ciad, Ghana, Congo, Etiopia o Angola), Altri (Libano, Sri Lanka, Mauritius, Giordania, Repubblica Domenicana, Isola di Cuba o El Salvador) e Non Rilevata. A questi risultati è possibile affiancarne altri ancor più informativi mettendo in relazione la nazionalità dei singoli coinvolti in uno specifico IPH<sup>6</sup>. Le percentuali che ne derivano confermano quanto rilevato poc'anzi: la cittadinanza prevalente si traduce in un numero molto elevato di omicidi (1219 su 1643) nei quali vittima e autore risultano essere entrambi italiani. Al contrario, solo l'11%<sup>7</sup> dei casi totali coinvolge sia donne che uomini di nazionalità straniera. L'analisi condotta permette inoltre di appurare che il numero di IPH con autori italiani e vittime straniere è pressochè il doppio rispetto a quello di IPH con autori stranieri e vittime italiane (nello specifico, 121 contro 58). A completamento si vuole ricordare la presenza, per quanto in numerosità esigua (68 su 1643), di dati mancanti; osservando che se neppure un'eventuale imputazione di questi ultimi sarebbe in grado di alterare lo scenario predominante (ossia, IPH con vittima e autore di nazionalità italiana), nulla di analogo si può asserire in merito alle altre casistiche. Infine, si ricorda la necessità di un esame condizionato alla variabile temporale. Questa indagine rileva una crescita graduale nella proporzione di autori e vittime stranieri rispetto al totale degli IPH<sup>8</sup> (si vedano Figure 4.7 e 4.8 in Appendice). Tuttavia, affermazioni in merito all'entità di un tale aumento non sembrano particolarmente attendibili, causa la presenza sistematica di valori mancanti, l'andamento oscillatorio e il problema relativo all'intensità dell' "effetto procedura" (procedura di imputazione delle "relazioni mancanti"). Ci si limita dunque a incrociare i risultati dell'esame condizionato alla componente temporale con quelli ottenuti in precedenza. Nello specifico,

- sebbene si registri una crescita graduale nella proporzione di autori e vittime stranieri rispetto al totale degli IPH, la nazionalità predo-

---

<sup>6</sup>Per una rappresentazione sintetica di tali risultati si rimanda a Fig. 4.6 in Appendice.

<sup>7</sup>Percentuale calcolata rispetto al totale delle osservazioni, incluse quelle con modalità "Non Rilevata" in corrispondenza delle variabili `nazionalita_vittima` e/o `nazionalita_autore`.

<sup>8</sup>Si noti che in questo caso non è possibile costruire alcun tasso di incidenza, in quanto autori e vittime di nazionalità straniera potrebbero non essere residenti in Italia.

minante rimane quella italiana per entrambe le figure coinvolte e per tutto il periodo 1992-2015;

- i dati mancanti si distribuiscono - in modo non omogeneo - sui primi dieci anni di rilevazione (1992-2011), pertanto il periodo 2002-2015 non risente di una tale problematica.

Seguono alcune considerazioni specifiche in merito alla dinamica caratterizzante gli IPH. Con una percentuale del 45% il movente *passionale* si impone come più diffuso. Tuttavia termini quali "delitto passionale" non sono propri dell'ordinamento giuridico italiano bensì di un linguaggio giornalistico che individua il movente in una rappresentazione potenzialmente violenta dell'amore (per un approfondimento si veda §). In questo caso l'uso dell'aggettivo vuole essere esclusivamente funzionale all'analisi: permette di definire in modo sintetico gli IPH dettati da motivi di gelosia e/o vendicativi. Il secondo movente più diffuso, con una percentuale del 29%, ricade invece sotto la dicitura "familiari". Tali motivi sono da intendersi condizionatamente al fenomeno di interesse: tra autore e vittima non può sussistere alcuna relazione di parentela pertanto "familiari" individua IPH che coinvolgono coppie sposate, facenti parte - per l'appunto - dello stesso nucleo familiare. Infine, il rimanente 26% della totalità si presenta così ripartito: lite (15%), follia (8%), altro (2%) e rancori personali (1%). Detto ciò è necessario fare un appunto in merito a questi risultati (sintetizzati in Appendice, Fig. 4.9). Le informazioni contenute all'interno dei *dataset Autori* e *Vittime* (vedi § .3) fanno riferimento alla fase in cui le forze dell'ordine registrano ufficialmente la denuncia di omicidio. Questa tipologia di raccolta determina una diversa affidabilità del dato a seconda della variabile in esame, poichè s'intuisce fin da subito che non tutto possa essere certo alla deposizione della denuncia formale. E nello specifico, proprio un'informazione come il movente dell'autore potrebbe mutare a seguito di ulteriori indagini. Di conseguenza l'interesse primario non si ripone tanto sulle frequenze relative quanto sulla coerenza delle modalità in sè: il risultato più significativo sta nel verificare che quest'ultime siano in linea con la definizione di IPH e più precisamente con il contesto di relazione intima affettiva caratterizzante il tipo di omicidi. L'analisi condotta permette di trarre una serie di conclusioni anche in merito all'arma del delitto. Innanzitutto si ricorda che i mezzi specifici sono stati raggruppati in macro-categorie secondo la classificazione "Marzio": arma

da fuoco (arma da fuoco corta, pistola d'ordinanza, arma da fuoco lunga o fucile subacqueo), arma da taglio (arma da taglio o ascia), arma impropria (arma impropria, corpocontudente, avvelenamento gas da cucina o liquido infiammabile), asfissia (asfissia, soffocamento, strangolamento, annegamento, incaprettamento o sottorramento), lesioni (lesioni, investimento, percosse, carbonizzato, ustionato, precipitazione, impiccagione, decapitato, sezionamento o sevizie) e non rilevata. A partire dalle percentuali di diffusione è possibile constatare una netta prevalenza di armi da fuoco e da taglio (nell'ordine, 37 e 32%) tra gli IPH. Mentre le morti per asfissia (12%), arma impropria (10%), lesioni (7%) o arma non rilevata (2%) risultano sensibilmente meno frequenti<sup>9</sup>. Che la categoria più diffusa sia quella delle armi da fuoco potrebbe apparire poco in linea con la dinamica caratterizzante gli IPH oggetto di studio: «Bartolomeo (2017) parla degli omicidi delle donne come omicidi con modalità "primitive", ammazzamenti a seguito di colluttazioni corpo-a-corpo in cui l'uomo sfoga una rabbia inaudita» (Minello e Dalla Zuanna, 2020). E se all'arma da fuoco si associa una maggior componente premeditativa, allora lo scenario appena descritto sembrerebbe più compatibile con le altre macro-categorie della classificazione Marzio. Tale intuizione trova conferma nello studio di Minello e Dalla Zuanna (2020), dove ulteriori analisi in merito producono il seguente risultato. «[...] le altre forme di aggressione prevalgono se contrapposte insieme alle armi da fuoco. Si tratta di asfissia, lesioni, arma impropria, tutte modalità che rientrano nella definizione di "primitive" proposta da Bartolomeo». Da qui l'idea che le macro-categorie individuate non bastino a descrivere in modo esauriente il ruolo delle armi all'interno degli IPH. Sebbene le percentuali ottenute con la classificazione Marzio costituiscano un buon punto di partenza, si impone la necessità di operare una suddivisione diversa, che tenga conto della dinamica caratterizzante il fenomeno in esame: arma con forte componente premeditativa (arma da fuoco), arma con forte componente "primitiva" (arma da taglio, arma impropria, asfissia, lesioni) e non rilevata. In questo modo, le frequenze relative che ne derivano<sup>10</sup> permettono di verificare quanto già preannunciato: nella maggior parte dei casi l'aspetto istintivo proprio degli IPH trova conferma nelle armi utilizzate dagli aggressori.

---

<sup>9</sup>Per una rappresentazione sintetica di tali risultati si rimanda a Fig. 4.10 in Appendice.

<sup>10</sup>Arma con forte componente premeditativa: 37%; arma con forte componente "primitiva": 61%; non rilevata: 2%.

Infine si ricorda la presenza della variabile dicotomica `n_armi`, creata in fase di risoluzione del problema relativo ai duplicati (per un approfondimento in merito si rimanda a § 2.3.3). Quest'ultima fornisce un altro dato interessante: solo per il 3% degli omicidi oggetto di studio è stata registrata più di un'arma del delitto.

Con riferimento agli autori si segnala la disponibilità di un'informazione aggiuntiva, che esula dall'evento IPH. In particolare, l'89% degli aggressori ha commesso altri crimini prima di consumare quello oggetto di studio. Tuttavia non è possibile accedere alle specifiche di tali reati pregressi; perciò ci si limita ad affermare che gli autori di IPH tendono ad essere soggetti pregiudicati, ossia con precedenti penali di qualche tipo.

Per concludere si riportano alcuni risultati in merito al numero di persone coinvolte. L'IPH, come da definizione, coinvolge due sole figure: autore e vittima, legati in una relazione intima affettiva presente o pregressa. Ma questo tipo di reato potrebbe consumarsi all'interno di una dinamica più ampia, ossia di un omicidio che coinvolge più di due persone. Informazioni a riguardo (numero totale di autori coinvolti e numero totale di vittime coinvolte) sono presenti all'interno del *dataset* e permettono di condurre le seguenti riflessioni. Nella maggior parte dei casi - 1547 su 1643 - il fenomeno oggetto di studio si configura come da definizione formale. Al contrario, il restante 6% si distribuisce tra diverse combinazioni possibili: la più frequente prevede un autore e due vittime (3%), seguita da quella che prevede invece due autori e una sola vittima (1%)<sup>11</sup>. Non avendo ulteriori dettagli in merito, si procede a descrivere alcuni scenari ipotetici tenendo a mente il contesto proprio degli IPH. Nel caso di due aggressori appare piuttosto sensato pensare che l'uomo in relazione (presente o pregressa) con la vittima abbia avuto bisogno di un complice per riuscire a commettere l'omicidio. Molto più complesso lo scenario con due vittime poichè le possibili identità della terza persona coinvolta diventano molteplici. La dinamica d'ira istintiva caratterizzante l'IPH è in effetti compatibile con l'uccisione di un soggetto "esterno", presente al momento dell'aggressione o accorso in aiuto a seguito di un qualche segnale d'allarme; tuttavia sembra ragionevole ipotizzare che figure vicine

---

<sup>11</sup>Per visualizzare le frequenze di tutte le combinazioni osservate si rimanda a Fig. 4.11 in Appendice.



alla vittima, come figli, familiari o *partners*/amanti, abbiano più probabilità di rimanere coinvolte nell'omicidio.



## Capitolo 4

# Il ruolo dei Centri Antiviolenza nell'evoluzione degli IPH italiani

Questo capitolo è interamente dedicata a studiare - tramite l'impiego di modelli statistici parametrici - il ruolo dei centri antiviolenza nell'evoluzione degli *Intimate Partner Homicides* italiani. In dettaglio, si compone delle seguenti sottosezioni:

1. Costruzione del *dataset*;
2. Scelta dei modelli statistici;
3. Procedura d'analisi e discussione dei risultati.

### 4.1 Costruzione del *dataset*

L'intera procedura inizia con la seguente constatazione: nessuno dei *dataset* disponibili (si veda § 2.2 e § 2.3) contiene tutte le informazioni potenzialmente rilevanti per l'analisi di interesse. Pertanto emerge fin da subito la necessità di costruire un nuovo insieme di dati, che sia funzionale all'obiettivo prefissato. Di seguito si riporta la descrizione di una tale procedura.

Voler indagare il ruolo dei centri antiviolenza (CAV) nell'evoluzione degli IPH italiani identifica automaticamente i *dataset* di principale interesse tra quelli disponibili:

- *IPH* (costruito a partire dai *dataset Autori e Vittime* (§ 2.3)): insieme di dati che contiene informazioni in merito agli IPH italiani con riferimento al periodo 1992-2015;
- *CAV* (messo a disposizione per l'analisi (§ 2.2)): insieme di dati che contiene informazioni in merito ai CAV con riferimento al periodo 1965-2017.

Appare evidente che il primo aggiustamento debba essere quello relativo all'intervallo temporale. Nello specifico, si seleziona il sottoinsieme di dati riferiti al periodo 1992-2015 anche all'interno del *dataset CAV*.

Si ricorda che in *CAV* è stato registrato il numero di centri antiviolenza per ogni possibile combinazione di anno e provincia; mentre in *IPH* sono state rilevate una serie di variabili - tra cui l'anno e la provincia - per ogni IPH denunciato. Ed è proprio questa differenza in merito alle unità statistiche dei due insiemi di dati che suggerisce la procedura più sensata: per ogni combinazione anno-provincia, si estrae il numero totale di IPH a partire dal *dataset* che li contiene<sup>1</sup>.

L'obiettivo teorico sembrerebbe tradursi in uno studio dell'eventuale legame tra numero di CAV e numero di IPH registrati, con la combinazione anno-provincia che ne descrive il livello di dettaglio. Tuttavia, il numero di centri antiviolenza è solo una delle tante versioni impiegabili per valutare il loro ruolo. Da qui l'idea di elaborarne altre, delegando all'analisi tramite modelli la scelta dell'opzione migliore. Una prima ipotesi consiste nel ritenere sufficientemente informativa la semplice presenza/assenza di CAV. Un'altra teoria si basa invece sul dubbio che misure di incidenza (nello specifico, il tasso di incidenza e una relativa versione categoriale) possano risultare più esplicative. Il nuovo *dataset* conterrà dunque quattro versioni della stessa variabile (si rimanda all'elenco delle variabili presenti all'interno del *dataset*).

Ma se l'obiettivo è uno studio non distorto della relazione in esame, non si può ignorare la possibilità di influenza anche da parte di altre variabili. Date le quantità assolute (numero totale di IPH e numero totale di CAV), sem-

---

<sup>1</sup>Si noti che una tale aggregazione del *dataset IPH* comporta la perdita di tutte le informazioni specifiche rilevate sui singoli omicidi. Assume ancora più rilevanza, dunque, l'analisi condotta in § 3 con l'obiettivo di indagare le caratteristiche proprie degli IPH italiani e la loro evoluzione nel corso del periodo considerato.

bra ragionevole controllare per il totale della popolazione a rischio di IPH. Nello specifico, il dato sul numero delle donne residenti in Italia di età non inferiore ai 14 anni<sup>2</sup>, condizionato alle diverse combinazioni anno-provincia, si estrae dalla banca dati ISTAT<sup>3</sup>. La natura del problema in esame suggerisce anche un'altra riflessione: se l'interesse viene posto sul numero di IPH, allora potrebbe avere senso controllare per una quantità che tenga conto del numero degli altri omicidi. Più precisamente, si procede a creare un indice di criminalità con riferimento alla specifica combinazione anno-provincia, rapportando i delitti non di tipo IPH al totale della popolazione in esame. Si noti che il numero di omicidi diversi da quelli di tipo IPH (numeratore del rapporto) deve essere ricercato all'interno dei *dataset Autori e Vittime* (si veda § 2.2). A completamento è d'obbligo una precisazione: le variabili di controllo sopra citate corrispondono a quelle ottenibili a partire dai dati a disposizione; pertanto non si può escludere con certezza la possibilità di influenza da parte altre quantità che invece andrebbero rilevate o ricercate altrove.

La procedura appena descritta costruisce un nuovo insieme di dati, sul quale è ora possibile condurre l'analisi di interesse. Nello specifico, con riferimento alla combinazione anno-provincia, si dispone di 2514 osservazioni e 9 variabili:

- IPH: numero totale di IPH (variabile quantitativa)
- CAV: numero totale di CAV (variabile quantitativa)
- CAV\_presenza: presenza di CAV (variabile dicotomica: "sì" o "no")
- CAV\_tasso: tasso di incidenza dei CAV sulla popolazione a rischio di sperimentare l'evento IPH (variabile quantitativa)
- CAV\_tasso\_cat: tasso di incidenza dei CAV sulla popolazione a rischio di sperimentare l'evento IPH (variabile qualitativa con 3 modalità: "basso", "medio" o "alto")

---

<sup>2</sup>Si ricorda che un tale limite è stato scelto individuando - all'interno del *dataset IPH* - l'età minima registrata per le vittime di IPH.

<sup>3</sup>*Ricostruzione intercensuaria della popolazione per età, sesso e territorio al 1° gennaio* (<https://demo.istat.it/?l=it>)

- **POPOLAZIONE**: numerosità della popolazione a rischio di sperimentare l'evento IPH (variabile quantitativa)
- **index\_criminalita\_prov**: indice di criminalità (variabile quantitativa)
- **ANNO**: anno di accadimento degli IPH (variabile quantitativa con supporto {1992, ..., 2015})
- **PROVINCIA**: provincia in cui sono avvenuti gli IPH (variabile qualitativa con modalità corrispondenti alle province italiane nel periodo di rilevazione)

Un'ultima nota: come anticipato in § 3, nel 2002 l'istituzione di 4 nuove province (Barletta-Andria-Trani, Monza e Brianza, Fermo e Sud Sardegna) modifica quella che era la precedente suddivisione territoriale.

## 4.2 Scelta dei modelli statistici

Questa sezione si occupa di individuare i modelli statistici più idonei all'analisi, in base ai dati disponibili e all'obiettivo fondante.

Come noto, l'analisi viene condotta allo scopo di indagare il ruolo dei centri antiviolenza nell'evoluzione degli IPH italiani. In pratica, ossia alla luce dei dati disponibili, questo obiettivo teorico si declina nel voler indagare se e in che modo la presenza (in senso proprio, in termini quantitativi e in termini di incidenza) dei CAV - centri antiviolenza - ha un effetto sul numero di IPH - *Intimate Partner Homicides* - commessi annualmente nelle diverse province italiane. Esistono svariati modelli statistici per un tale problema di regressione. Il primo elemento da esaminare al fine di individuarli correttamente è la natura della variabile risposta. Nello specifico, si noti che una quantità come il numero di IPH corrisponde al risultato di un conteggio, il cui totale non può essere prefissato. Da Figura 4.1 se ne visualizza anche la distribuzione marginale con riferimento all'intero periodo 1992-2015<sup>4</sup>. E si ricordi anche quanto osservato nella sezione precedente (§ 3) in merito alla rarità del fenomeno oggetto di studio. Sembra dunque più che ragionevole procedere alla stima di modelli basati sulla distribuzione di Poisson; poichè

---

<sup>4</sup>La stessa distribuzione viene indagata anche con riferimento al periodo "ridotto" 2004-2015 - insieme di dati sul quale si condurrà un'analisi parallela (come anticipato in § 3) - rilevando un andamento pressochè identico (si rimanda a Figura 4.12 in Appendice).

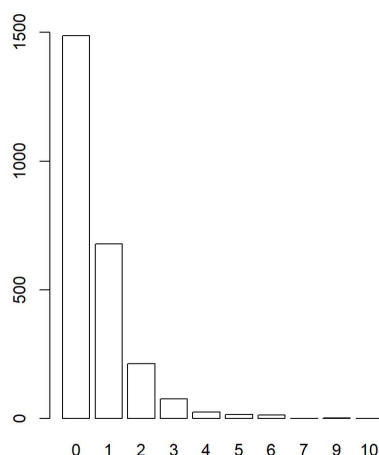


Figura 4.1: Distribuzione marginale della variabile risposta - Italia, 1992-2015.

quest'ultima costituisce un modello adatto a descrivere proprio il numero di occorrenze - in un certo intervallo di tempo o regione dello spazio - di eventi in un qualche senso rari.

Per completezza espositiva, segue un elenco di richiami teorico-formali in merito ai modelli scelti per l'analisi. I modelli lineari generalizzati per risposte univariate rappresentano un'estensione del modello di regressione lineare normale per trattare risposte con (a.) distribuzione che può essere diversa dalla normale e (b.) media che è funzione del predittore lineare (non necessariamente la funzione identità). Modelli di questo tipo sono pertanto definiti da tre elementi.

- *Componente aleatoria*: è la distribuzione di probabilità di  $Y_i$ . Si assume che le v.c.  $Y_1, \dots, Y_n$  siano indipendenti con densità appartenente a una famiglia di dispersione esponenziale<sup>5</sup>. Formalmente,  $Y_i \sim DE_1(\mu_i, a_i(\phi)v(\mu_i))$ , con  $Y_1, \dots, Y_n$  indipendenti.
- *Predittore lineare*: per un vettore di coefficienti di regressione  $\beta = (\beta_1, \dots, \beta_p)^T$  e una matrice del modello  $X$ ,  $n \times p$ , il predittore lineare è  $\eta = X\beta$  con componenti  $\eta_i = \mathbf{x}_i\beta$ .

<sup>5</sup>La densità di  $Y_i$ ,  $i = 1, \dots, n$ , appartiene a una famiglia di dispersione esponenziale univariata se è esprimibile nella forma  $p(y_i; \theta_i, \phi) = \exp\left\{\frac{\theta_i y_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right\}$  con  $y_i \in S \subseteq \mathbb{R}$ ,  $\theta_i \in \Theta \subseteq \mathbb{R}$ ,  $a_i(\phi) > 0$ . Il parametro  $\theta_i$  è detto parametro naturale, mentre  $\phi$  è detto parametro di dispersione.

- *Funzione di legame*: è la funzione  $g(\cdot)$  che collega  $\mu_i$  a  $\eta_i$ , assunta di forma nota, derivabile con continuità e invertibile

$$g(E(Y_i)) = g(\mu_i) = \eta_i = \mathbf{x}_i\beta.$$

Tra le possibili, viene privilegiata quella cosiddetta canonica, ossia la funzione  $g(\mu_i) = \theta_i(\mu_i)$  adottando la quale il parametro naturale  $\theta_i$  (si veda nota a piè pagina <sup>24</sup>) risulta combinazione lineare delle variabili esplicative con coefficienti  $\beta$ ,  $\theta_i = \mathbf{x}_i\beta$ .

Per dati di conteggio con totale non prefissato - come nel caso del numero di IPH - il modello statistico (parametrico) più elementare corrisponde al **modello lineare generalizzato di Poisson**. Quest'ultimo assume che  $y_1, \dots, y_n$  siano realizzazioni di v.c. indipendenti con distribuzione  $Y_i \sim DE_1(\mu_i, \mu_i)$ ,  $\mu_i > 0$ , sul supporto  $\mathbb{N}$  (ossia,  $Y_i \sim Po(\mu_i)$ ) e  $g(\mu_i) = \eta_i = \mathbf{x}_i\beta$ . Di conseguenza,  $Var(Y_i) = E(Y_i) = \mu_i = g^{-1}(\mathbf{x}_i\beta)$ . Con la funzione di legame canonica  $g(\mu_i) = \log(\mu_i) = \eta_i = \mathbf{x}_i\beta$ , si ha  $Y_i \sim P(e^{\eta_i})$ , ossia  $Y_i \sim P(e^{\mathbf{x}_i\beta})$ . Altre funzioni di legame utilizzabili per la regressione di Poisson sono la radice quadrata,  $g(\mu_i) = \sqrt{\mu_i} = \eta_i = \mathbf{x}_i\beta$ , e l'identità<sup>6</sup>,  $g(\mu_i) = \mu_i = \eta_i = \mathbf{x}_i\beta$ .

Tuttavia, il valore atteso di una risposta che rappresenta un conteggio potrebbe essere proporzionale alla dimensione entro cui si osservano gli eventi. Per esempio, condizionatamente al problema in esame, sembra ragionevole supporre che il numero medio di IPH sia proporzionale alla totalità di donne a rischio di sperimentarlo (come già anticipato in § 4.1). In casi come questo diventa più opportuno adattare ai dati un **modello lineare generalizzato di Poisson con un termine offset**. Si indichi in generale con  $t_i$  un indice di dimensione. Un modello lineare generalizzato di Poisson che include un termine *offset* assume che  $y_1, \dots, y_n$  siano realizzazioni di v.c. indipendenti con distribuzione  $Y_i \sim Po(\mu_i)$  con  $\mu_i = t_i\lambda_i$  e  $\log(\lambda_i) = (\eta_i) = \mathbf{x}_i\beta$ . Si noti che il parametro  $\lambda_i$  esprime l'intensità, in quanto rappresenta il numero medio di eventi per unità di dimensione. Se la funzione di legame è quella canonica, si ha  $g(\mu_i) = \log(\mu_i) = \log(t_i) + \mathbf{x}_i\beta$ . Il termine  $\log(t_i)$  - che compare nel predittore lineare come un'ulteriore variabile esplicativa con coefficiente noto

<sup>6</sup>Si noti, tuttavia, che l'identità richiede opportuni vincoli sul vettore dei coefficienti di regressione  $\beta$  per garantire stime  $\hat{\mu}_i > 0$ .



e pari a 1 - è detto *offset*.

Spesso i dati di conteggio presentano sovradisersione, ossia una variabilità campionaria ampiamente superiore a quella prevista dal modello<sup>7</sup>. Vanno quindi ricercate estensioni dei modelli di Poisson che tengano conto di un tale allontanamento dagli assunti di base. Nello specifico, considerare la distribuzione binomiale negativa - anziché quella di Poisson/al posto della Poisson - per la variabile risposta fornisce un semplice modello statistico (parametrico) per dati di conteggio con sovradisersione. Un **modello lineare generalizzato con risposta binomiale negativa** assume che  $y_1, \dots, y_n$  siano realizzazioni di v.c. indipendenti con funzione di densità

$$p(y_i; \mu, \kappa) = \frac{\Gamma(y_i + \kappa)}{\Gamma(y_i + 1)\Gamma(\kappa)} \left(\frac{\mu/\kappa}{1 + \mu/\kappa}\right)^{y_i} \left(\frac{1}{1 + \mu/\kappa}\right)^\kappa,$$

con parametri  $\mu_i = g^{-1}(\mathbf{x}_i\beta)$  e  $\kappa = 1/\tau$ .

Nel caso in esame, i valori di media e varianza campionaria (0.65 e 1.10, rispettivamente) non escludono a priori l'adattamento di un tale modello.

Altrettanto spesso la frequenza di conteggi pari a zero può essere superiore a quella attesa sotto un modello di Poisson (o binomiale negativo). Nello specifico, si dice che distribuzioni empiriche con moda pari a zero e media relativamente grande mostrano un'inflazione di zeri rispetto al modello di Poisson (o binomiale negativo). Tale fenomeno, in genere, si manifesta quando alcune unità risultano del tutto "inattive" rispetto alla caratteristica che si vuole contare e dunque hanno valore della risposta identicamente pari a zero. Il modello naturale per una distribuzione con inflazione di zeri è un modello mistura tra una distribuzione usuale per dati di conteggio (come la Poisson o la binomiale negativa) e una distribuzione genere in 0 (**modello lineare generalizzato per dati di conteggio con inflazione di zeri**). A titolo esemplificativo, un modello di Poisson con inflazione di zeri assume che  $y_1, \dots, y_n$  siano realizzazioni di v.c. indipendenti

$$Y_i \sim \begin{cases} 0 & \text{con probabilità } 1 - \phi_i \\ P(\lambda_i) & \text{con probabilità } \phi_i \end{cases}.$$

---

<sup>7</sup>Si ricorda che il modello lineare generalizzato di Poisson assume l'identità tra media e varianza della risposta.

I parametri  $\phi_i$  e  $\lambda_i$  possono essere espressi in funzione di variabili esplicative,  $\mathbf{x}_{Ai}$  per  $\phi_i$  e  $\mathbf{x}_{Bi}$  per  $\lambda_i$ .<sup>8</sup>

Si noti che se da un lato il caso in esame non coinvolge formalmente unità del tutto "inattive", dall'altro i valori di moda e media della risposta (0 e 0.65, rispettivamente) non escludono a priori l'adattamento di un tale modello.

### 4.3 Procedura d'analisi e discussione dei risultati

L'obiettivo di sottosezione è duplice: illustrare nel dettaglio la procedura d'analisi per poi discuterne i risultati.

L'intera analisi si sviluppa in funzione dell'obiettivo fondante: indagare se e in che modo la presenza (in senso proprio, in termini quantitativi e in termini di incidenza) dei CAV - centri antiviolenza - ha un effetto sul numero di IPH - *Intimate Partner Homicides* - commessi annualmente nelle diverse province italiane, controllando per altri fattori potenzialmente rilevanti. Come anticipato in § 4.2, data la natura della variabile risposta, sembra più che ragionevole procedere alla stima di modelli statistici basati sulla distribuzione di Poisson. Inoltre, si noti fin da subito il fine esplicativo piuttosto che previsivo dell'indagine: descrivere la relazione osservata tra variabile risposta ed esplicative, con particolare *focus* su uno dei fattori in esame. Sembra dunque altrettanto ragionevole valutare i modelli stimati in base a uno o più criteri di informazione, ossia ricercando il miglior compromesso tra i principi di adattamento e parsimonia compatibilmente con l'interesse primo dell'intero studio.

Nello specifico, dunque, la procedura consiste delle seguenti fasi. Per ciascuna classe di modelli individuata al § 4.2, si effettua una regressione su tutti i sottoinsiemi delle variabili esplicative<sup>9</sup>, ordinando poi i risultati secondo il criterio di informazione scelto. A questo punto è l'obiettivo d'analisi a configurare tre diversi scenari possibili.

1. Il sottoinsieme ottimo (ovvero, quel sottoinsieme di predittori che minimizza il criterio di informazione scelto) include la variabile relativa

---

<sup>8</sup>Naturalmente,  $\mathbf{x}_{Ai}$  e  $\mathbf{x}_{Bi}$  possono anche risultare coincidenti.

<sup>9</sup>Si noti che una tale procedura di selezione delle variabili risulta possibile grazie al numero esiguo di predittori.

ai centri antiviolenza. Si può quindi indagare la relazione di interesse esaminando le stime ottenute per il modello migliore in termini di compromesso tra adattamento e parsimonia.

2. Nonostante il sottoinsieme ottimo non includa la variabile relativa ai centri antiviolenza, esiste almeno un modello per cui l'effetto di questo predittore risulta significativo. Dunque, all'indagine del sottoinsieme ottimo si aggiunge inevitabilmente un'analisi in merito alla relazione di interesse, che deve essere condotta valutando i risultati ottenuti per il modello migliore (in termini del criterio di informazione prescelto) tra quelli che includono un effetto significativo della variabile relativa ai centri antiviolenza.
3. L'effetto della variabile relativa ai centri antiviolenza non risulta significativo per nessuno dei modelli in esame. Ci si può quindi limitare allo studio del miglior modello in termini di compromesso tra adattamento e parsimonia, registrando l'assenza di un tale effetto per l'intera classe di modelli in esame.

S'intuisce fin da subito che lo scenario 2. richiede estrema cautela nell'interpretazione dei risultati. Infatti, sebbene l'idea di esaminare il miglior modello tra quelli che includono un effetto significativo dei CAV trovi sostegno nell'obiettivo d'analisi, questo tipo di procedura aumenta sensibilmente il rischio di individuare una relazione spuria tra IPH e CAV. Al fine di ridurre al minimo un tale rischio, diventa fondamentale capire perchè il modello in questione non sia stato individuato - dal criterio di informazione - come migliore all'interno di quella specifica classe. Più precisamente si tratta di distinguere tra due casi:

- (a.) pur non rappresentando la miglior soluzione di compromesso tra adattamento e parsimonia, il modello si dimostra comunque valido (ossia, dimostra un buon adattamento ai dati e una complessità ragionevole) e, al tempo stesso, sensato da un punto di vista delle stime ottenute (ad esempio, i segni dei coefficienti stimati sono in linea con quanto ci si aspetterebbe a livello teorico);
- (b.) il modello stima un effetto significativo della variabile CAV solo perchè il sottoinsieme di predittori in esame non include le variabili rilevanti

(ovvero, quei predittori che hanno un impatto realmente significativo sulla variabile risposta).

Da un punto di vista pratico, si rende dunque necessario un controllo diagnostico (tramite analisi dei residui) del modello in questione, ed un confronto dei risultati - quali, valore del criterio di informazione, sottoinsieme delle esplicative e stime dei parametri - con quelli che si sono ottenuti per il miglior modello della classe indagata.

Infine, l'intera procedura giunge a un termine solo dopo aver ordinato l'insieme dei migliori modelli per ogni classe sulla base del criterio di informazione scelto.

In realtà, da un punto di vista pratico, il procedimento appena illustrato viene ripetuto diverse volte, al fine di condurre un'analisi più completa. Si ricorda quanto deciso in merito al periodo di rilevazione (§ 3) e alla variabile relativa ai centri antiviolenza (§ 4.1).

Sebbene l'assenza di chiare anomalie nella dinamica evolutiva degli IPH permetta di escludere fin da subito un "effetto procedura" importante, i valori registrati per alcuni anni - 1992, 1994, 1999, 2003 - inducono a pensare che possa esserci comunque del condizionamento alle operazioni eseguite durante la fase di imputazione delle "relazioni mancanti". Ne deriva la decisione di condurre due analisi parallele, utilizzando insiemi di dati relativi a periodi temporali differenti (nello specifico, 1992-2015 e 2004-2015); poichè un confronto dei risultati ottenuti potrebbe fornire qualche indicazione rilevante in merito all'entità di un tale effetto.

Alla luce delle variabili inizialmente disponibili, l'obiettivo teorico sembrava tradursi in uno studio dell'eventuale legame tra numero di IPH e numero di CAV. Da qui l'idea di elaborare altre versioni della variabile esplicativa di interesse, cosicchè l'analisi trovasse realizzazione nell'indagine di diverse forme d'effetto agite dai CAV sul totale degli IPH. Nello specifico, l'obiettivo teorico si declina in uno studio dell'eventuale legame tra numero di IPH e presenza - in senso proprio (presenza/assenza), in termini quantitativi o in termini di incidenza - dei CAV. Ne derivano dunque quattro analisi distinte, a partire dalle quali è possibile individuare la miglior versione, in termini esplicativi, della variabile di interesse.

Infine si ricorda che il confronto tra modelli stimati avviene sulla base di uno o più criteri di informazione. Nello specifico, sono stati scelti AIC (*Akaike In-*

formation Criterion) e BIC (*Bayesian Information Criterion*), che risultano così definiti:

$$AIC = -2 \hat{l}_{M_p} + 2p;$$

$$BIC = -2 \hat{l}_{M_p} + p \log(n);$$

dove  $\hat{l}_{M_p}$  denota la log-verosimiglianza massimizzata per il modello  $M_p$  con  $p$  termini. Un'ulteriore osservazione: il BIC soddisfa la proprietà di consistenza<sup>10</sup>. Ma per il caso in esame si rende noto che i due criteri individuano lo stesso ordinamento dei modelli stimati; pertanto non è necessario condurre due analisi parallele (o scegliere di indagare solo i risultati prodotti da una delle due).

Riassumendo: la procedura di cui sopra viene ripetuta per un totale di otto volte, al duplice fine di valutare l'entità dell' "effetto procedura" e studiare versioni differenti della variabile relativa ai centri antiviolenza - individuando, eventualmente, la migliore in termini esplicativi.

A completamento, segue una sintesi di informazioni utili, accompagnate da precisazioni tecnico-pratiche.

Si riporta l'elenco delle variabili disponibili condizionatamente al loro ruolo nell'analisi.

- Variabile risposta: numero totale di IPH (IPH).
- Variabile esplicativa di primario interesse: una tra
  - numero totale di CAV (CAV);
  - presenza/assenza di CAV (CAV\_presenza);
  - tasso di incidenza dei CAV (CAV\_tasso);
  - entità del tasso di incidenza dei CAV (CAV\_tasso\_cat).
- Altre variabili disponibili: indice di criminalità (index\_criminalita\_prov), anno (ANNO), provincia (PROVINCIA), numerosità della popolazione a rischio (POPOLAZIONE).

Una nota in merito alla variabile ANNO: l'andamento oscillatorio che si osserva per il numero di IPH nel corso del periodo 1992-2015, suggerisce la neces-

---

<sup>10</sup>Nel senso che se il modello corretto è tra quelli considerati, BIC lo individua con probabilità 1 al crescere della numerosità campionaria. Si noti che la presenza di  $n$  nella penalizzazione serve proprio a garantire la consistenza del criterio.

sità di considerare una versione più parsimoniosa della variabile temporale. Ciò che interessa, infatti, è una stima della tendenza di fondo relativamente all'evoluzione del fenomeno oggetto di studio. Da un punto di vista pratico, le *splines* di regressione sembrano lo strumento più adatto a modellare un tale andamento, senza tuttavia rincorrere fluttuazioni casuali presenti in corrispondenza di valori specifici del predittore. In particolare, si stimano i modelli delle classi individuate al § 4.2, inserendo l'anno tramite *splines* cubiche (ossia, con grado pari a 3) con diversi valori per il numero di nodi; lasciando così ai dati la scelta della specificazione migliore. Un esame approfondito delle quantità rilevanti (stime dei coefficienti di regressione, livelli di significatività delle stime e valore dei criteri di informazione) ha tuttavia ritenuto che fosse sufficientemente informativo considerare la trasformazione del predittore:  $f(x; \beta) = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3$ , dove con  $x$  si indica la variabile temporale. D'ora in avanti, quindi, con ANNO si fa riferimento alla sua trasformazione di cui sopra.

Segue un elenco in merito a tutte le classi di modelli parametrici scelte per condurre l'analisi. Nell'ordine,

1. **modelli lineari generalizzati di Poisson**  
(con diverse funzioni di legame: canonica, radice quadrata e identità)
2. **modelli lineari generalizzati con risposta binomiale negativa**
3. **modelli lineari generalizzati di Poisson con inflazione di zeri**
4. **modelli lineari generalizzati di Poisson con un termine *offset***  
(termine *offset*:  $\log(\text{POPOLAZIONE})$ )
5. **modelli lineari generalizzati di Poisson ad effetti casuali**  
(effetto casuale: intercetta casuale per PROVINCIA)

Si ricordi che la regressione su tutti i sottoinsiemi - effettuata per ciascuna classe di cui sopra - ha come insieme completo delle variabili esplicative {una versione di CAV, `index_criminalita_prov`, ANNO}.

Un'ultima nota: se non viene specificato altrimenti, la funzione di legame utilizzata è quella canonica.

Dopo aver illustrato nel dettaglio l'intera procedura, ci si concentra sulla discussione dei risultati ottenuti.

Facendo riferimento all'analisi che utilizza l'insieme di dati relativo al periodo 1992-2015, i modelli migliori in termini di AIC (o, equivalentemente BIC) sono:

1. il modello lineare generalizzato di Poisson che include un'intercetta casuale per PROVINCIA,  $\log(\text{POPOLAZIONE})$  come termine *offset* e  $\{\text{ANNO}, \text{index\_criminalita\_prov}\}$  come sottoinsieme ottimo delle variabili esplicative;<sup>11</sup>
2. il modello lineare generalizzato di Poisson che include  $\log(\text{POPOLAZIONE})$  come termine *offset* e  $\{\text{ANNO}, \text{index\_criminalita\_prov}\}$  come sottoinsieme ottimo delle variabili esplicative.<sup>12</sup>

Mentre invece, facendo riferimento all'analisi che utilizza l'insieme di dati relativo al periodo 2004-2015, i modelli migliori in termini di AIC (o, equivalentemente BIC) sono:

1. il modello lineare generalizzato di Poisson che include un'intercetta casuale per PROVINCIA,  $\log(\text{POPOLAZIONE})$  come termine *offset* e  $\{\text{index\_criminalita\_prov}\}$  come sottoinsieme ottimo delle variabili esplicative;<sup>13</sup>
2. il modello lineare generalizzato di Poisson che include  $\log(\text{POPOLAZIONE})$  come termine *offset* e  $\{\text{index\_criminalita\_prov}\}$  come sottoinsieme ottimo delle variabili esplicative.<sup>14</sup>

Coerentemente con l'obiettivo di studio, la prima cosa che si rileva per tutti i modelli sopra elencati è l'assenza di una qualsiasi versione della variabile relativa ai CAV. In realtà, un'indagine dettagliata in merito ai risultati prodotti dalla regressione su tutti i sottoinsiemi, estende questa assenza di effetto significativo da parte del predittore di interesse a tutti i sottoinsiemi che lo coinvolgono. In altre parole, anche quando l'insieme delle esplicative in esame include una versione della variabile relativa ai CAV, il coefficiente associato a quest'ultima non risulta mai significativamente diverso da 0.

---

<sup>11</sup>Per la sintesi delle quantità rilevanti che descrivono un tale modello si rimanda a Figura 4.15 in Appendice.

<sup>12</sup>Per la sintesi delle quantità rilevanti che descrivono un tale modello si rimanda a Figura 4.16 in Appendice.

<sup>13</sup>Per la sintesi delle quantità rilevanti che descrivono un tale modello si rimanda a Figura 4.13 in Appendice.

<sup>14</sup>Per la sintesi delle quantità rilevanti che descrivono un tale modello si rimanda a Figura 4.14 in Appendice.

Si può dunque affermare che, per entrambe le classi di modelli parametrici (modelli lineari generalizzati di Poisson ad effetti casuali e modelli lineari generalizzati di Poisson con un termine *offset*) e per entrambi i periodi di rilevazione (1992-2015 e 2004-2015), nessuna delle variabili relative ai centri antiviolenza sembra avere un effetto significativo sul numero di IPH. Questa conclusione assume ancora più importanza se si pensa che vale anche - non solo - per i modelli migliori in termini di AIC (o, equivalentemente BIC).

I risultati ottenuti per le analisi relative a periodi differenti si rivelano molto simili. In particolare, i due modelli migliori coincidono - non solo per classe di appartenenza ma anche per tipologia e modalità di inclusione delle variabili - a meno del predittore relativo all'anno; quest'ultimo, infatti, risulta significativo solo considerando l'intero insieme di dati a disposizione (1992-2015). Tuttavia, un esame più approfondito in merito al miglior modello per ogni classe parametrica indagata, con distinzione di periodo (1992-2015, 2004-2015), ha permesso di rilevare la stessa differenza di significatività per tutti i modelli selezionati da AIC (o, equivalentemente BIC). S'intuisce che le ragioni di un tale risultato vadano ricercate nella relazione tra il numero di IPH (variabile risposta) e l'anno di rilevazione (variabile esplicativa), al netto degli altri predittori inclusi.

**Caso a.** Se l'andamento osservato ed in parte ricostruito (tramite imputazione delle "relazioni mancanti") per il numero di IPH al passare del tempo, riflette adeguatamente quella che è la dinamica evolutiva reale. Allora un risultato distinto in merito alla significatività della variabile temporale a seconda del periodo considerato può dirsi ragionevole. Poichè questo sarebbe compatibile con l'andamento specifico che caratterizza il fenomeno in esame: la totalità degli IPH registra una leggera crescita prima di stabilizzarsi attorno al valore medio di 67 omicidi l'anno (si rimanda a § 3, più precisamente alla discussione in merito a Figura 4.2). In altre parole, da un'analisi basata sull'intero periodo (1992-2015) - caratterizzato da un'andamento più vario - è ragionevole aspettarsi un effetto significativo della variabile temporale; mentre da un'analisi basata sul periodo ridotto 2004-2015 - caratterizzato da un andamento più stabile - è ragionevole attendersi invece un effetto non significativo della variabile temporale.

**Caso b.** Se l'andamento osservato ed in parte ricostruito (tramite impu-



tazione delle "relazioni mancanti") per il numero di IPH al passare del tempo, non riflette adeguatamente quella che è la dinamica evolutiva reale. Relativamente al primo periodo, dunque, la procedura di imputazione adottata non è stata in grado di ricostruire un andamento sufficientemente attendibile per il totale degli IPH. Ma allora nessun risultato in merito alla significatività della variabile temporale può ritenersi valido; poichè sicuramente influenzato da un "effetto procedura".

Si noti che in realtà, non conoscendo la vera dinamica evolutiva per il numero di IPH italiani nel periodo 1992-2004, risulta difficile stabilire con certezza in quale situazione ci si trovi. Tuttavia, se da un lato non si riesce a valutare l'attendibilità del risultato in merito all'effetto della variabile temporale; dall'altro, ciò che invece rassicura è la concordanza di tutti gli altri risultati. E nello specifico, di quello relativo ai centri antiviolenza: come già anticipato, nessuna delle versioni di CAV sembra avere un impatto significativo sul numero di IPH, a prescindere dal periodo temporale (1992-2015 o 2004-2015) su cui si basa l'analisi. Da qui la scelta di concentrare il *focus* sulla discussione dei risultati ottenuti per il periodo 2004-2015, in quanto più affidabili; solo dopo aver ribadito, ancora una volta, che le conclusioni per il periodo più esteso (1992-2015) coincidono - in termini di AIC (o BIC), variabili incluse, segni e livelli di significatività delle stime - a meno del predittore anno.

Un'altra informazione riscontrabile dall'esame dei modelli migliori individuati tramite AIC (o, equivalentemente BIC), è la presenza costante dell'indice di criminalità specifico per combinazione anno-provincia. In linea con quanto ci si aspetterebbe a livello teorico, un aumento del tasso di criminalità favorisce una crescita anche degli omicidi di tipo IPH. Nello specifico, si ricorda che per modelli lineari generalizzati di Poisson con legame canonico, l'interpretazione dei coefficienti di regressione è la seguente. Un incremento unitario della variabile esplicativa - indice di criminalità - comporta un incremento moltiplicativo del valore atteso della variabile risposta - numero medio di IPH - per un fattore  $e^\beta$ , fermo restando il valore delle altre variabili esplicative nel modello. Dove con  $\beta$  si indica il coefficiente associato allo specifico predittore in esame. Non devono dunque preoccupare i valori elevati delle stime per  $\beta$  che si osservano a livello empirico (si vedano, ad esempio, Figure 4.13 e 4.14), poichè l'entità dell'effetto sul numero medio di IPH è funzione del prodotto tra questi valori - nell'ordine di  $10^3$  - e quelli dell'indice

di criminalità - nell'ordine di  $10^{-5}$ . Un'osservazione interessante: l'impatto di un tale predittore perde un po' di significatività (pur rimanendo sempre significativamente diverso da zero) nel momento in cui viene incluso un termine *offset* all'interno del modello. Questo risultato trova spiegazione nelle riflessioni che seguono. 1. Il problema d'indagine richiede un termine *offset*, come il totale di donne a rischio di IPH per combinazione anno-provincia, affinché i dati di conteggio siano rapportati all'indice di dimensione entro cui si rilevano<sup>15</sup>. 2. Il numero medio di IPH (valore atteso della variabile risposta) si rivela proporzionale all'indice di criminalità. Dunque, in assenza di un termine *offset*, il modello valuta il tasso di criminalità come un buon sostituto dell'indice di dimensione richiesto dal problema; da qui deriva la sovraenfaticazione dell'effetto relativo al predittore in esame.

In realtà una situazione affine si riscontra anche per le variabili relative ai centri antiviolenza. In particolar modo, gli effetti di *CAV* (numero di CAV) e *CAV\_tasso* (tasso di incidenza dei CAV) vengono stimati come significativamente diversi da zero, solo finché non si include nel modello la (log-)popolazione a rischio come termine *offset*. Ciò che non convince dal principio sono i valori positivi delle stime: secondo i modelli adattati, un incremento del numero/tasso di CAV comporta un incremento del numero medio di IPH. Si noti che in questo caso il modello senza *offset* non si limita a valutare le variabili come buoni sostituti dell'indice di dimensione richiesto dal problema, sovraenfaticandone l'effetto; bensì produce vere e proprie anomalie nei risultati. Tuttavia, considerati l'obiettivo d'analisi e i risultati ottenuti, si vuole fare un'ulteriore verifica. L'idea è quella di creare una variabile indicatrice del tasso di affollamento registrato per ciascuna provincia<sup>16</sup>; per valutare se, includendo un termine di interazione tra quest'ultima e una versione della variabile CAV, i risultati migliorino. Ciò non accade, perciò quanto affermato sopra trova ulteriore conferma.

Con riferimento all'AIC (o, equivalentemente al BIC), si noti che il modello migliore differisce dal secondo migliore per l'inclusione di un termine ag-

---

<sup>15</sup>Per un'evidenza empirica a sostegno di tale affermazione è sufficiente indagare l'elenco dei modelli migliori in termini di AIC (o, equivalentemente BIC), verificando che tutti includono il logaritmo della popolazione a rischio come termine *offset*.

<sup>16</sup>La variabile dicotomica *metropoli* è stata creata a partire dalla variabile POPOLAZIONE: si distinguono le province con più di 400000 abitanti (modalità "sì") da quelle con una popolazione a rischio inferiore al valore di soglia (modalità "no"). Le province più densamente popolate sono Bergamo, Bologna, Brescia, Catania, Firenze, Genova, Milano, Napoli, Palermo, Roma, Salerno, Torino, Padova, Verona e Caserta.

giuntivo: un'intercetta casuale per la variabile provincia. Perciò è possibile dedurre che quest'ultima apporti un contributo significativo alla *performance* del modello (come si può verificare anche dal valore della devianza spiegata in Figura 4.13). In altre parole, il numero di IPH sembra essere influenzato in maniera significativa dalla suddivisione territoriale in province. Si rende necessaria un'ultima osservazione: il modello lineare generalizzato ad effetti casuali è l'unico, tra quelli scelti per l'analisi, che permette di valutare - in modo parsimonioso - il contributo della variabile provincia. Ed è quindi grazie a questo modello che la scelta di condurre un'analisi con le province come livello di dettaglio territoriale trova sostegno.

A completamento, si segnala che le estensioni del modello lineare generalizzato di Poisson considerate - modello lineare generalizzato con risposta binomiale negativa e modello lineare generalizzato con inflazione di zeri - non apportano miglioramenti significativi in termini di AIC (o, equivalentemente BIC) rispetto alla formulazione *standard*. È dunque possibile asserire che i dati di conteggio in esame non presentano una sovradisersione e/o un'inflazione di zeri tali da richiedere l'adattamento di modelli più complessi.

Considerati l'obiettivo dell'analisi e i risultati ottenuti, si procede ad indagare la dinamica evolutiva del numero di centri antiviolenza (Figura 4.17) rispetto a quella osservata per il totale degli IPH (Figura 4.2), ricercando indicazioni utili a spiegare la non significatività del predittore di interesse. Mentre il totale degli IPH registra solo una leggera crescita prima di stabilizzarsi attorno al valore medio di 67 omicidi l'anno; il totale dei CAV registra una crescita sostenuta per l'intero periodo 1992-2015. Nel dettaglio, il numero di IPH passa da 44 a 59, con valore minimo registrato pari a 42 (1994) e valore massimo registrato pari a 93 (2003). Mentre il numero di CAV passa da 29 a 209, con una variazione percentuale del +621%. Si può dunque affermare - facendo riferimento ai dati disponibili - che ad una crescita importante del numero di centri antiviolenza non corrisponde una diminuzione altrettanto evidente del numero di *Intimate Partner Homicides*. Tuttavia, la dinamica evolutiva osservata per il numero di CAV nel corso del periodo considerato, fornisce un'altra indicazione utile: potrebbe aver senso includere un termine di interazione tra il predittore relativo ai centri antiviolenza e la variabile temporale. Si procede, dunque, ad una ri-stima del

modello migliore in termini di AIC (o, equivalentemente BIC)<sup>17</sup> per valutare la significatività di un tale termine di interazione. I risultati ottenuti vanno distinti per versione della variabile CAV:

1. nel caso delle variabili *CAV*, *CAV\_presenza* e *CAV\_tasso\_cat*, si ottengono stime significativamente diverse da 0 (con un livello di significatività pari allo 0.1 o 0.05) al più per uno o due dei nuovi coefficienti associati al termine di interazione;
2. nel caso della variabile *CAV\_tasso*, si ottengono stime significativamente diverse da 0 per tutti i nuovi coefficienti inclusi nel modello.

Nel caso 1. non si ottengono miglioramenti - in termini di significatività delle stime e/o interpretabilità della relazione tra variabile di interesse e risposta - tali da giustificare l'inclusione del termine in esame. Mentre invece per il caso 2. il risultato è sorprendente. Nello specifico, includere un termine di interazione tra il tasso di incidenza dei centri antiviolenza e la variabile temporale ha un triplice effetto:

- apporta un contributo significativo in termini esplicativi;
- rende fortemente significativo il coefficiente associato al tasso di incidenza dei CAV (coefficiente che se aggiunto solo come effetto principale del predittore di interesse non risulta affatto significativo);
- cambia segno al coefficiente associato al tasso di incidenza dei CAV (coefficiente che se aggiunto solo come effetto principale del predittore di interesse viene stimato con segno positivo).

Per la sintesi delle quantità rilevanti che descrivono un tale modello si rimanda a Figura 4.18. Si noti che - in termini di AIC (o, equivalentemente BIC) - il nuovo modello non risulta migliore rispetto al precedente<sup>18</sup>. Tuttavia, i miglioramenti - in termini di significatività delle stime ed interpretabilità della relazione tra predittore di interesse e variabile risposta - sono tali da

<sup>17</sup>Si ricorda che, facendo riferimento all'analisi che utilizza l'insieme di dati relativo al periodo 1992-2015, il modello migliore in termini di AIC (o, equivalentemente BIC) è il modello lineare generalizzato di Poisson che include un'intercetta casuale per *PROVINCIA*,  $\log(\text{POPOLAZIONE})$  come termine *offset* e  $\{\text{ANNO}, \text{index\_criminalita\_prov}\}$  come sottoinsieme ottimo delle variabili esplicative.

<sup>18</sup>Questo risultato appare ragionevole se si pensa che il nuovo modello prevede la stima di 4 parametri in più rispetto alla sua versione precedente; e AIC o BIC individuano invece la miglior soluzione di compromesso tra adattamento e parsimonia.

giustificare l'inclusione del termine di interazione (e dell'effetto principale relativo al predittore di interesse).

In ultimo si riportano le conclusioni finali, con riferimento specifico all'obiettivo d'analisi. L'intera procedura è stata condotta al fine di indagare se e in che modo la presenza (in senso proprio, in termini quantitativi e in termini di incidenza) dei CAV - centri antiviolenza - ha un effetto sul numero di IPH - *Intimate Partner Homicides* - commessi annualmente nelle diverse province italiane, controllando per altri fattori potenzialmente rilevanti. Alla luce dei risultati ottenuti, è possibile affermare che l'analisi relativa al periodo 1992-2015 ha evidenziato un effetto significativo del tasso di incidenza dei CAV sul numero di IPH commessi ogni anno nelle diverse province italiane. In particolare, la stima di un tale effetto si ottiene adattando ai dati relativi al periodo 1992-2015, un modello lineare generalizzato di Poisson che include un'intercetta casuale per provincia, il logaritmo della popolazione a rischio come termine *offset* e {indice di criminalità, anno di rilevazione, tasso di incidenza dei CAV, un termine di interazione tra l'anno e il tasso di incidenza dei CAV} come sottoinsieme ottimo delle variabili esplicative. E, nello specifico, la presenza del termine di interazione tra la variabile temporale e il predittore di interesse si rivela fondamentale per una stima significativa dell'effetto in esame: controllando per l'evoluzione temporale del tasso di incidenza dei CAV, emerge che un incremento del tasso di incidenza dei CAV comporta una riduzione significativa del numero medio di IPH.

A completamento, è necessario ribadire che le variabili esplicative esaminate corrispondono a quelle ottenibili a partire dai dati a disposizione. Dunque non si può escludere con certezza la possibilità di influenza da parte altre quantità rilevanti; quali, ad esempio, informazioni in merito al numero di donne che si rivolgono ai centri antiviolenza oppure in merito alla percentuale di vittime che aveva chiesto un qualche tipo di supporto prima dell'omicidio.



# Appendice



Figura 4.2: Numero totale di IPH - Italia, 1992-2015.

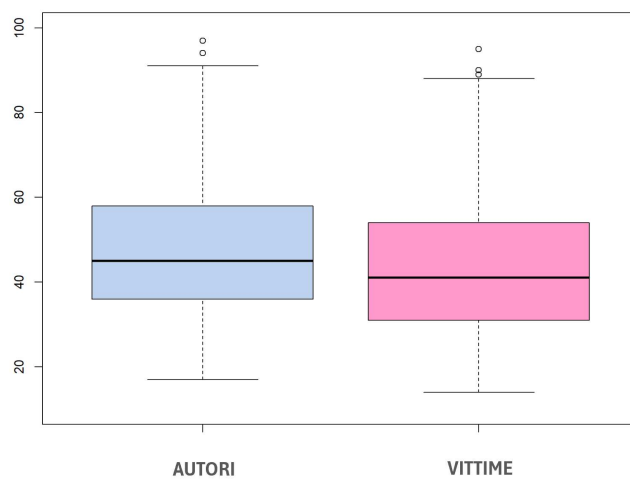


Figura 4.3: Boxplot relativo alla variabile età di autori (sinistra) e vittime (destra) - Italia, 1992-2015.



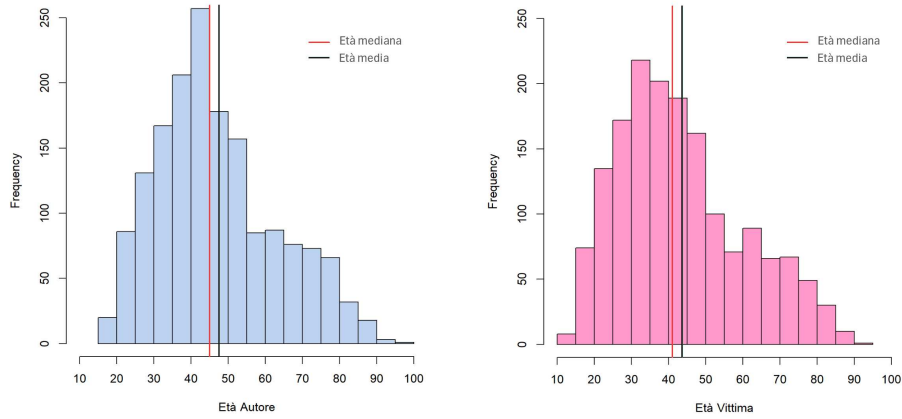


Figura 4.4: Distribuzione di frequenza relativa alla variabile età di autori (sinistra) e vittime (destra) - Italia, 1992-2015.

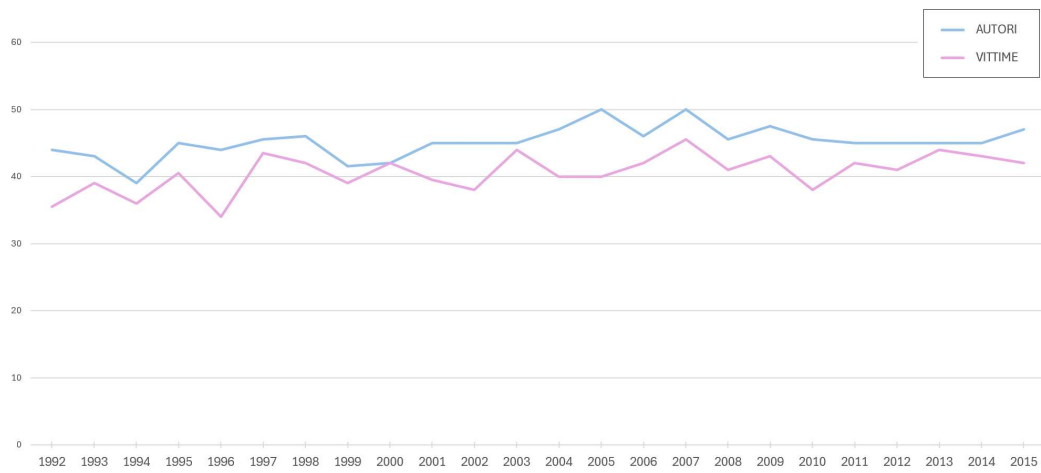


Figura 4.5: Età mediana all'evento IPH - Italia, 1992-2015.

		VITTIME		
		Italiana	Straniera	NR
AUTORI	Italiano	74,2%	7,4%	-
	Straniero	3,5%	10,8%	-
	NR	-	-	4,1%

Figura 4.6: Tabella di frequenza congiunta (percentuale) relativa alle nazionalità di vittime e autori - Italia, 1992-2015. Legenda: NR = "Non Rilevata".

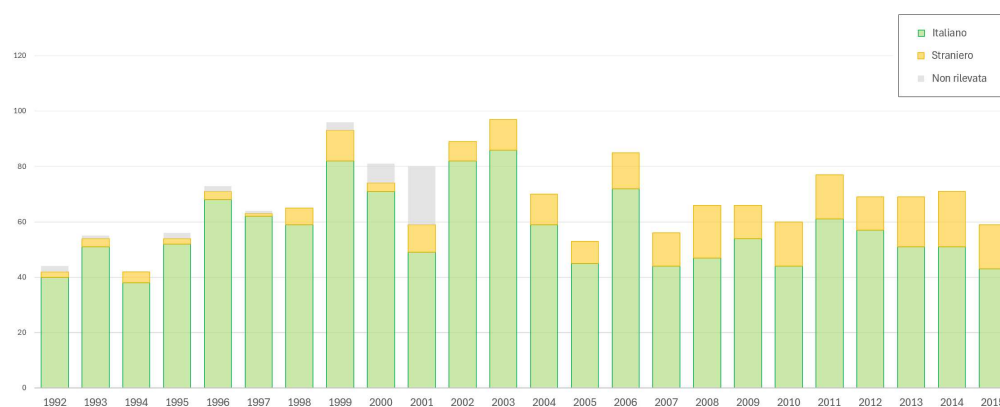


Figura 4.7: Numero totale di IPH, stratificato per nazionalità dell'autore - Italia, 1992-2015.

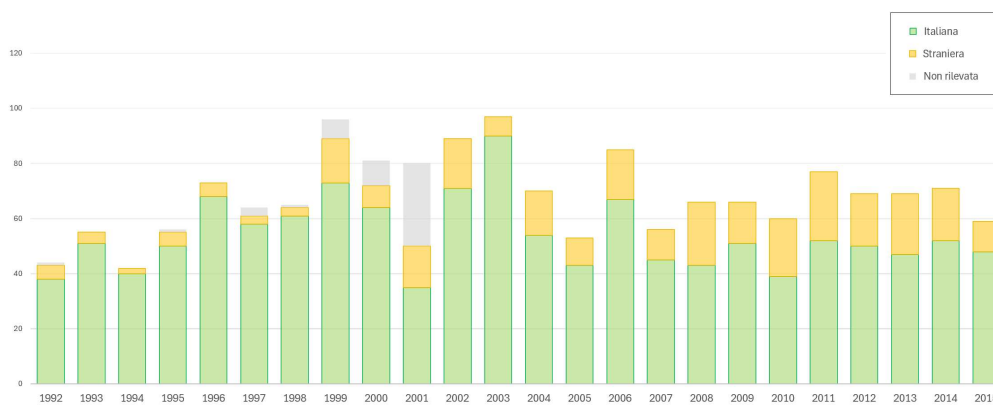


Figura 4.8: Numero totale di IPH, stratificato per nazionalità della vittima - Italia, 1992-2015.

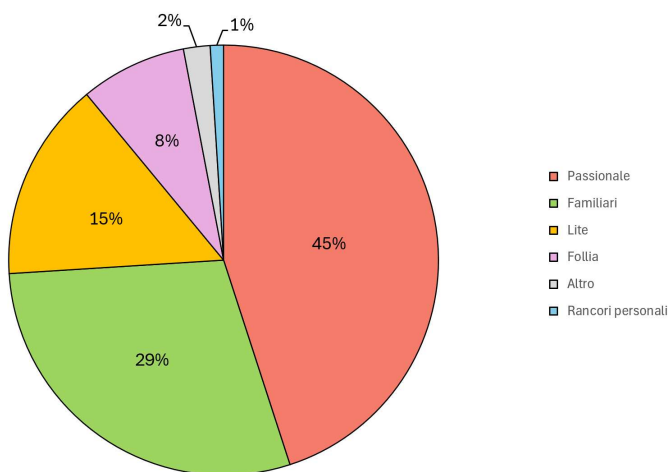


Figura 4.9: IPH per movente dell'autore - Italia, 1992-2015.

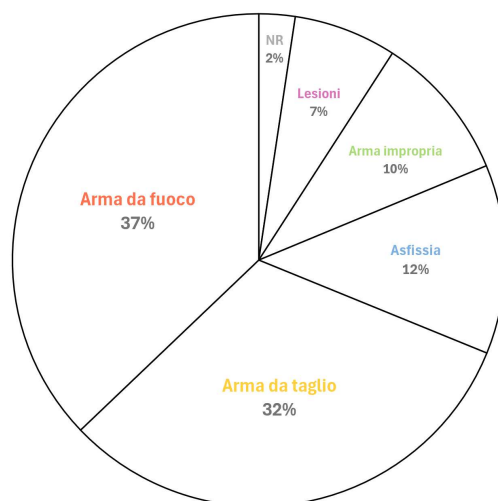


Figura 4.10: IPH per arma del delitto - Italia, 1992-2015.  
Legenda: NR = "Non Rilevata".

		VITTIME			
		1	2	3	4
AUTORI	1	1547	50	6	1
	2	25	4	-	-
	3	3	-	-	-
	4	3	3	-	-
	7	1	-	-	-

Figura 4.11: Tabella di frequenza congiunta relativa al numero di vittime e autori coinvolti - Italia, 1992-2015.

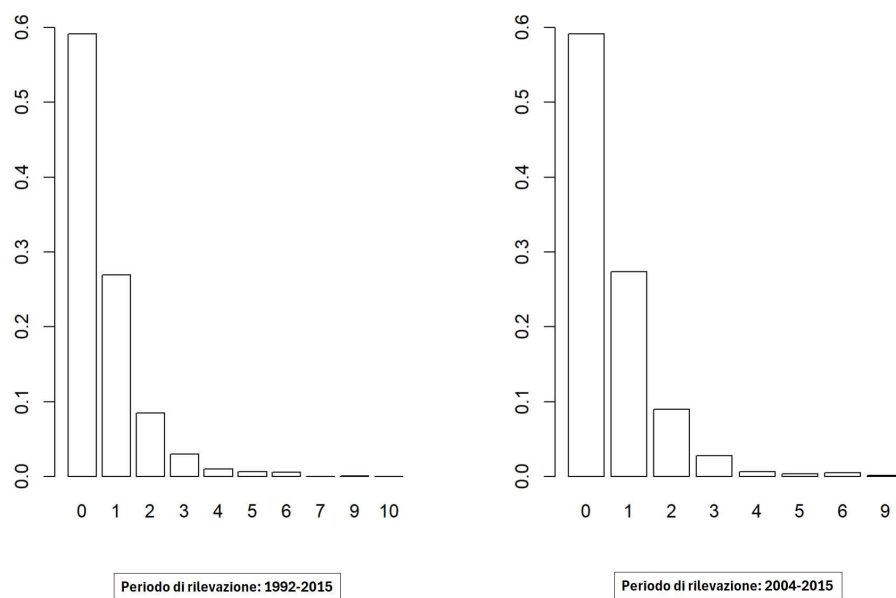


Figura 4.12: Distribuzione marginale della variabile risposta relativa al periodo 1992-2015 (sinistra) e 2004-2015 (destra) - Italia.

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: poisson ( log )
Formula: dati$IPH ~ (1 | dati$PROVINCIA) + dati$index_criminalita_prov
Offset: log(dati$POPOLAZIONE)

      AIC      BIC   logLik deviance df.resid
 2408.9  2424.3 -1201.4  2402.9    1269

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.6129 -0.6510 -0.4754  0.6512  3.9630

Random effects:
 Groups          Name      Variance Std.Dev.
 dati$PROVINCIA (Intercept) 0.03157  0.1777
Number of obs: 1272, groups: dati$PROVINCIA, 106

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -12.95933    0.04187 -309.55 <2e-16 ***
dati$index_criminalita_prov 3813.36261    51.21125   74.46 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
              (Intr)
dt$ndx_cr_  -0.021

```

Figura 4.13: Sintesi delle quantità rilevanti per il miglior modello lineare generalizzato di Poisson ad effetti casuali - 2004-2015.

```

Call:
glm(formula = dati$IPH ~ dati$index_criminalita_prov, family = poisson,
     offset = log(dati$POPOLAZIONE))

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -12.94983    0.04248 -304.851 <2e-16 ***
dati$index_criminalita_prov 3589.79967 1512.06303    2.374  0.0176 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 1233.9 on 1271 degrees of freedom
Residual deviance: 1228.9 on 1270 degrees of freedom
AIC: 2411.8

Number of Fisher Scoring iterations: 5

```

Figura 4.14: Sintesi delle quantità rilevanti per il miglior modello lineare generalizzato di Poisson con un termine *offset* - 2004-2015.

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: poisson ( log )
Formula: dati$IPH ~ -1 + (1 | dati$PROVINCIA) + bs(dati$ANNO, intercept = T) + dati$index_criminalita_prov
Offset: log(dati$POPOLAZIONE)

      AIC      BIC   logLik deviance df.resid
 4914.4  4949.4 -2451.2  4902.4    2508

Scaled residuals:
   Min       1Q   Median       3Q      Max
-1.8375 -0.6611 -0.4808  0.6203  6.3053

Random effects:
 Groups              Name          Variance Std.Dev.
 dati$PROVINCIA (Intercept) 0.04299  0.2073
Number of obs: 2514, groups:  dati$PROVINCIA, 106

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
bs(dati$ANNO, intercept = T)1 -13.42729    0.10525 -127.58 <2e-16 ***
bs(dati$ANNO, intercept = T)2 -12.01056    0.19719  -60.91 <2e-16 ***
bs(dati$ANNO, intercept = T)3 -13.20543    0.19142  -68.99 <2e-16 ***
bs(dati$ANNO, intercept = T)4 -12.91968    0.08921 -144.81 <2e-16 ***
dati$index_criminalita_prov  2315.22461   23.56927   98.23 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
              b($ANNO,i=T)1 b($ANNO,i=T)2 b($ANNO,i=T)3 b($ANNO,i=T)4
b($ANNO,i=T)2 -0.649
b($ANNO,i=T)3  0.457      -0.791
b($ANNO,i=T)4 -0.121      0.427      -0.591
dt$ndx_crm_   -0.007      -0.002      -0.003      -0.004

```

Figura 4.15: Sintesi delle quantità rilevanti per il miglior modello lineare generalizzato di Poisson ad effetti casuali - 1992-2015.

```

Call:
glm(formula = dati$IPH ~ -1 + bs(dati$ANNO, intercept = T) + dati$index_criminalita_prov,
     family = poisson, offset = log(dati$POPOLAZIONE))

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
bs(dati$ANNO, intercept = T)1 -13.42345    0.10528 -127.504 <2e-16 ***
bs(dati$ANNO, intercept = T)2 -12.00965    0.19665  -61.070 <2e-16 ***
bs(dati$ANNO, intercept = T)3 -13.20755    0.19099  -69.154 <2e-16 ***
bs(dati$ANNO, intercept = T)4 -12.92282    0.08682 -148.851 <2e-16 ***
dati$index_criminalita_prov  2028.36820   872.82941    2.324  0.0201 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 1.2525e+09 on 2514 degrees of freedom
Residual deviance: 2.5736e+03 on 2509 degrees of freedom
AIC: 4931.6

Number of Fisher scoring iterations: 6

```

Figura 4.16: Sintesi delle quantità rilevanti per il miglior modello lineare generalizzato di Poisson con un termine *offset* - 1992-2015.

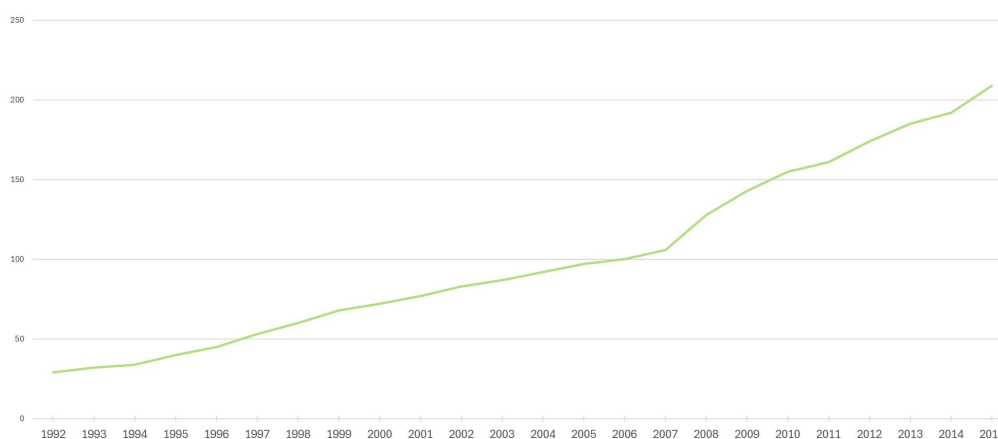


Figura 4.17: Numero totale di CAV - Italia, 1992-2015.



```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: poisson ( log )
Formula: dati$IPH ~ (1 | dati$PROVINCIA) + -1 + bs(dati$ANNO, intercept = T) +
  dati$index_criminalita_prov + dati$CAV_tasso + (bs(dati$ANNO, intercept = T):dati$CAV_tasso)
offset: log(dati$POPOLAZIONE)

      AIC      BIC   logLik deviance df.resid
4917.9  4976.2 -2448.9  4897.9    2504

Scaled residuals:
    Min       1Q   Median       3Q      Max
-1.8441 -0.6586 -0.4765  0.6123  6.2872

Random effects:
 Groups          Name          Variance Std.Dev.
dati$PROVINCIA (Intercept) 0.04169  0.2042
Number of obs: 2514, groups:  dati$PROVINCIA, 106

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
bs(dati$ANNO, intercept = T)1 -1.350e+01  1.060e-01 -127.33 <2e-16 ***
bs(dati$ANNO, intercept = T)2 -1.177e+01  1.979e-01  -59.48 <2e-16 ***
bs(dati$ANNO, intercept = T)3 -1.359e+01  1.920e-01  -70.77 <2e-16 ***
bs(dati$ANNO, intercept = T)4 -1.272e+01  8.995e-02 -141.41 <2e-16 ***
dati$index_criminalita_prov    2.229e+03  2.096e+01  106.33 <2e-16 ***
dati$CAV_tasso                 -3.186e+04  2.571e+01 -1239.53 <2e-16 ***
bs(dati$ANNO, intercept = T)1:dati$CAV_tasso  6.006e+04  2.114e+01  2841.56 <2e-16 ***
bs(dati$ANNO, intercept = T)2:dati$CAV_tasso -2.451e+04  2.163e+01 -1133.34 <2e-16 ***
bs(dati$ANNO, intercept = T)3:dati$CAV_tasso  1.070e+05  2.436e+01  4392.72 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figura 4.18: Sintesi delle quantità rilevanti per il modello lineare generalizzato di Poisson ad effetti casuali con termine di interazione tra la variabile anno e il tasso di incidenza dei CAV - 1992-2015.



# Bibliografia

Barbagli, M. e Colombo A. (2011), *Rapporto sulla Criminalità e Sicurezza in Italia*. Roma, Fondazione Icsa.

Bartolomeo, F. (2017). Inchiesta con analisi statistica sul femminicidio in Italia. *Ministero della giustizia – Direzione generale di statistica e analisi organizzativa*.

Boari, M. (2007). *La coercizione privata nella Magna glossa: tracce fra diritto e violenza*. Milano, Giuffrè editore.

Brownmiller, S. (1975). *Against Our Will: Men, Women, and Rape*. New York, Fawcett Books.

Cavina, M. (2011). *Nozze di sangue. Storia della violenza coniugale*. Bari, Editori Laterza.

Ciaravolo, V. (2021). *Violenza di genere dalla A alla Z*. Roma, Armando Editori.

Dean, T. (2007). *Crime and justice in late medieval Italy*. Cambridge, Cambridge University Press.

Grzyb, M., Naudi, M. and Marcuello-Servòs, C., *Femicide definitions*, in Weil, S., Corradi, C. and Naudi, M. (2018). *Femicide across Europe: Theory, research and prevention*. Bristol, Bristol University Press.

Minello, A. e Dalla Zuanna, G. (2020). *Morire in Italia: omicidi di don-*

ne, omicidi in famiglia. *Popolazione e storia*, **20(2)**, 21-48.

Russell, D.E.H. e Harmes, R.A. (2001). *Femicide in global perspective*. New York, NY, Teacher's College Press.

Schettini, L. (2023). Le radici sociali, culturali e giuridiche della violenza di genere. *Il Bo live*.

# Sitografia e Filmografia

<https://www.infodata.ilsole24ore.com/2018/04/19/si-sposano-gli-italiani-differenze-deta-senso-della-coppia/>

<https://www.istat.it/it/violenza-sulle-donne/la-fuoriuscita-dalla-violenza/centri-antiviolenza>

<https://www.istat.it/it/violenza-sulle-donne/il-fenomeno/violenza-dentro-e-fuori-la-famiglia/numero-delle-vittime-e-forme-di-violenza>

<https://www.istat.it/it/violenza-sulle-donne/il-fenomeno/violenza-dentro-e-fuori-la-famiglia/consapevolezza-e-uscita-dalla-violenza>

<https://www.treccani.it/vocabolario/socioculturale/>

[https://www.unodc.org/documents/AnnualReport2015/Annual\\_Report\\_2016\\_WEB.pdf](https://www.unodc.org/documents/AnnualReport2015/Annual_Report_2016_WEB.pdf)  
(UNODC 2016, Annual Report 2015)

[http://apps.who.int/iris/bitstream/handle/10665/77432/WHO\\_RHR\\_12.36\\_eng.pdf;jsessionid=731120F825649B0BDC8C54FE42C7F922?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/77432/WHO_RHR_12.36_eng.pdf;jsessionid=731120F825649B0BDC8C54FE42C7F922?sequence=1)  
(WHO World Health Organization (2012). Understanding and addressing violence against women.)

*Un altro domani*, regia di Soldini, S. (2023).