



**Master I livello Data Analytics**

**L'approvazione del terrorismo come strumento di  
affermazione politica, ideologica, religiosa**

**Determinanti sociali, culturali e psicologiche in un campione armeno**

*Relatrice: prof.ssa Italia De Feis, prof.ssa Paola Celio*

*Candidato: Giammaria Trimarco*

*Anno Accademico: 2023-2024*



## **ABSTRACT**

La crescente radicalizzazione ideologica del dibattito politico e l'aumento nella frequenza degli eventi riconducibili al terrorismo, pongono seri rischi alla sicurezza e alla stabilità delle società.

Il presente lavoro intende considerare le determinanti della radicalizzazione terroristica teorizzate dalla letteratura sociale e psicologica, e rilevare quali di esse abbiano un'effettivo potere esplicativo sull'approvazione del terrorismo come strumento di lotta politica, ideologica o religiosa.

A questo scopo è stato utilizzato un dataset pubblico di alta qualità (World Values Survey) e selezionato un campione di partecipanti di nazionalità armena, paese storicamente interessato dal fenomeno.

I risultati delle analisi indicano che: i dati non rispettano gli assunti per l'uso dello stimatore ai minimi quadrati, il modello testato è sottospecificato, e infine che lo studio delle teorie sulla radicalizzazione terroristica, dal punto di vista psicologico-sociale, richiede uno spazio delle variabili elevato e metodi appropriati al confronto con tale complessità.

Studi futuri su dataset di natura simile a quello usato dovranno considerare pertanto l'uso di modelli statistici adatti a trattare dati ad alta dimensionalità, flessibilità nel metodo usato (classificazione vs regressione), e adottare modelli teorici che permettano una maggiore precisione nell'operazionalizzazione delle variabili.

## **QUADRO TEORICO E MOTIVAZIONI**

### **INQUADRAMENTO DEL FENOMENO**

Il terrorismo è una forma di lotta politica (Kalyvas, 2019) che fa uso della violenza al fine di provocare terrore nella popolazione generale, in un gruppo di persone o in particolari soggetti, in vista della realizzazione di scopi giustificati da considerazioni di natura politica, filosofica, ideologica, razziale, etnica o religiosa (UNGA resolution 49/60, 194, para.3).

La cosiddetta 'guerra globale al terrorismo', iniziata con la risposta degli Stati Uniti agli attacchi dell'11 Settembre 2001, è ancora in corso. In seguito a quell'evento, si è osservata un'espansione nella frequenza e nella diffusione geografica del fenomeno, con attacchi da parte di gruppi o di singoli individui (e.g., Moskalenko & McCauley, 2011) avvenuti in diverse aree del mondo, anche in epoca molto recente: Londra, Madrid, Jakarta, Parigi, Brussel, Kabul, solo per citarne alcuni.

Nella sola Europa, incluso il Regno Unito, negli anni dal 2010 al 2021 gli eventi riconducibili ad attentati terroristici riusciti, falliti o sventati, attuati da organizzazioni di diversa matrice ideologica (islamica, etno-separatista, politica) sono stati in totale 1821, con una media di 155 attacchi ogni

anno e 1560 arresti solo tra il 2019 e il 2021 (Europol, 2022).

Al 2021 il budget istituito per il fondo di sicurezza interna (ISF) per la lotta alle minacce transfrontaliere come terrorismo, criminalità informatica e criminalità organizzata, ammontava a circa 2 miliardi di euro (Parlamento Europeo, 2021).

Considerando il costo sociale ed economico del terrorismo, la conoscenza delle sue determinanti può essere utile a informare le decisioni di policy e orientare interventi per la prevenzione e il contrasto del fenomeno.

## LA RICERCA SUL TERRORISMO NELLE SCIENZE SOCIALI

Dal punto di vista delle scienze sociali il fenomeno del terrorismo può essere inquadrato come una conseguenza violenta della radicalizzazione ideologica, politica o religiosa. Il comportamento che ne segue eccede quanto è socialmente e normativamente atteso, nei termini sia dei mezzi utilizzati che dei loro effetti (Molinario, Jasko, Webber, & Kruglansky, 2022; Kruglansky, Bélanger, & Gunaratna, 2019).

Le determinanti della radicalizzazione sono oggetto di studio di diverse discipline nell'ambito delle scienze sociali, ciascuna dalla propria angolatura e in base al proprio livello di analisi. Le diverse proposte includono, ad esempio, il fallimento del processo di costituzione di uno stato (Piazza, 2008), l'occupazione straniera, la soppressione del dissenso politico (Pape, 2006), povertà (Berrebi, 2007), anomia – la percezione che la società sia incapace di fornire una guida morale al singolo - e stress sociale (Gupta, 2004).

Tuttavia, è irrealistico pensare che il fenomeno possa essere adeguatamente spiegato sulla base di poche determinanti, a maggior ragione considerando che le spiegazioni multicausali e multilivello sono la norma, più che l'eccezione, nella ricerca sociale.

Una spiegazione adeguata della radicalizzazione terroristica e delle sue conseguenze, dovrebbe tener conto di diversi livelli di analisi, considerando fattori di ordine: intraindividuale (e.g., valori, convinzioni personali), interpersonale (e.g., rapporti tra pari), di gruppo (e.g., reti sociali, gruppi, nazioni), sociale (e.g., status socio-economico, struttura sociale), storico (e.g., conflitto domestico), economico (e.g., PIL), culturale (e.g., norme e usanze culturalmente stabilite).

Una rassegna recente (Kruglansky et al. 2019) tenta di ordinare la letteratura sulla radicalizzazione violenta in cinque approcci. Di seguito se ne fornisce una sintesi e in Tabella 1 le determinanti considerate in ognuno di essi:

### *Approccio delle cause principali*

La logica di fondo che governa tale approccio è che la deprivazione sociale ed economica, presente a un certo livello in ogni società, dà origine al dissenso collettivo e a un ambiente sociale che, interagendo con i precipitanti, motiverebbero la nascita di gruppi terroristici e la legittimazione dell'azione violenta (e.g., Horgan, 2014).

### *Modelli psicologici*

Il livello di analisi dell'approccio è quello intraindividuale e coinvolge le strutture di conoscenza, gli atteggiamenti, la personalità, le motivazioni di un individuo, e le dinamiche interne che li coinvolgono. I quadri teorici che ne fanno parte provengono dalla psicologia clinica, perlopiù di riferimento psicoanalitico (e.g., teoria del narcisismo; Crayton, 1983), dai modelli di profilazione psicologica basata sui tratti di personalità in interazione con variabili socio-economiche, e dalla ricerca psicologico-sociale.

Le prime due famiglie di modelli si sono dimostrate deboli alla prova empirica (cfr. Lankford, 2013). L'ultima famiglia di modelli considera la radicalizzazione violenta come un processo che si dispiega nel tempo, facilitato e mantenuto da fattori di natura sociale e psicologica, e da processi interpersonali e di gruppo derivanti dall'identificazione sociale (Horgan, 2008; McCauley & Moskalenko, 2008).

### *Modello dell'attore razionale*

In questo approccio l'azione terroristica è inquadrata come azione strategica la cui razionalità risiede nella ricerca della massima utilità attesa conseguibile dal successo operativo, in relazione alle risorse disponibili e al rischio previsto (Enders & Su, 2007).

L'attore razionale indifferente al rischio è anche un 'attore devoto' (Atran, 2016; Atran & Sheikh, 2015), disposto a sacrificare sé stesso in nome della propria ideologia religiosa (e.g., Islam, Cristianesimo) e/o secolare (e.g., Diritti umani) e indifferente a incentivi di tipo materiale.

### *Teoria delle reti sociali*

La teoria delle reti sociali si avvale degli strumenti concettuali e analitici forniti dalla teoria dei grafi per studiare le dinamiche delle reti sociali, in questo caso costituite dalle relazioni che si instaurano tra le organizzazioni terroristiche, i loro membri e i simpatizzanti (e.g., McHallister, 2004), anche tramite reti sociali virtuali (e.g., Facebook).

### *Teoria dei movimenti sociali*

I movimenti sociali sono reti costituite da organizzazioni che condividono un'identità comune, e scopi in conflitto con le condizioni sociali esistenti, per conseguire i quali fanno uso di mezzi non convenzionali come marce, occupazioni, proteste violente e simili (della Porta, 2013) .

Il modello ipotizza l'esistenza di fattori che ne spiegano l'origine, la sopravvivenza e l'espansione, tra i quali rientrano ad esempio la percezione di ingiustizia sociale, la limitatezza delle risorse disponibili alla collettività , una rete di contatti ideologicamente affini e così via.

### *Significance Quest Theory e Modello delle 3N*

Per la Significance Quest Theory (Kruglanski, Molinario, Jasko, Webber, Pontus Leander, & Pierro, 2009) e il modello delle 3N (Webber & Kruglanski, 2018) l'estremismo violento è la conseguenza di un cammino individuale di ricerca di significato, che segue dalla percezione della mancanza di significato (Need), e sostenuta dall'adozione di un'ideologia radicale (Narrative) e dall'affiliazione a una rete sociale radicalizzata (Network).

Il bisogno di significato è assunto come una motivazione fondamentale della condotta umana, ed è inteso come il desiderio di percepirsi come persona di valore, di meritare rispetto, di 'sentirsi qualcuno'. Quando ciò non avviene, l'individuo sperimenta uno stato di perdita di significato (significance loss) e si attiva alla ricerca di opportunità e di mezzi per ripristinarlo. In questo senso l'ideologia radicale, l'affiliazione a una rete sociale e l'azione violenta sono espressione concreta di questa ricerca (Kruglanski, Szumowska, & Kopetz, 2023).

In Tabella1 si fornisce un quadro sinottico delle determinanti considerate dai diversi approcci descritti (cfr. Kruglanski et al., 2022).

## **OBIETTIVO GENERALE**

Il presente lavoro ha l'obiettivo di stimare l'impatto delle determinanti considerate, sull'approvazione del terrorismo come mezzo di affermazione politica, religiosa o ideologica, in un campione armeno.

A questo scopo si è scelto di utilizzare un dataset ampio e di elevata qualità, noto da tempo alla ricerca sociale, sul quale sono state applicate tecniche di apprendimento supervisionato.

## **MOTIVAZIONI E RILEVANZA PRATICA**

I risultati potrebbero fornire spunti per integrare le informazioni rilevanti per decisioni di policy e per orientare interventi di tipo psicosociale per la prevenzione e il contrasto del fenomeno, come anche fornire un riferimento empirico per raffinare i modelli teorici esistenti.

**Tabella 1. Determinanti della radicalizzazione in base agli approcci presentati**

<b>Cause principali (deprivazione relativa)</b>	Assenza di democrazia e/o di libertà civili, rule of law Fallimento o debolezza dello Stato Modernizzazione Ideologia estremista secolare o religiosa Antecedenti storici (violenza politica, guerre civili, rivoluzioni, dittatura, occupazione) Ineguaglianza di potere, egemonia Governo illegittimo o corrotto Repressione da parte di una potenza straniera (occupazione, colonialismo) Discriminazione su base etnica o religiosa Fallimento nell'integrazione di gruppi dissidenti o classi sociali emergenti Ingiustizia sociale Leader ideologici carismatici Eventi attivanti Povertà e bassi livelli di istruzione
<b>Modelli psicologici (profili di personalità, tratti, percezioni e credenze soggettive)</b>	Teorie psicologiche: crisi d'identità, narcisismo/ego threat, paranoia Profili psicologici: personalità autoritaria-estremista, psicopatologia, tratti di personalità in interazione con fattori socio-economici Processi socio-psicologici: vulnerabilità emotiva, insoddisfazione verso società e governo, identificazione con la vittima, legittimità nell'uso della violenza, azione politica convenzionale considerata inefficace, legittimazione della morte per giusta causa, legami con vittime di ingiustizia sociale o individui già radicalizzati, vittimizzazione, auto-persuasione, polarizzazione delle opinioni nel gruppo di riferimento, isolamento e minaccia del gruppo, competizione tra gruppi, aderenza a fazioni estremiste (fissione e condensazione del gruppo), azione provocatoria – reazione legittimante, odio, cultura del martirio
<b>Modello dell'attore razionale (utilità attesa)</b>	Massimizzazione dell'utilità attesa, indifferenza al rischio personale, valori sacri (secolari o religiosi), norme sociali e valori di gruppo
<b>Teoria delle reti sociali (relazioni tra nodi in una rete)</b>	Relazioni tra partecipanti e organizzazioni radicalizzate (amicizia, parentela), esposizione all'ideologia radicale (riviste, forum, social media etc.)
<b>Teoria dei movimenti sociali (convergenza dell'azione di gruppi diversi)</b>	Network tra individui e organizzazioni con identità comuni Fattori di iniziazione (politica di escalation, escalation competitiva, rete militante attiva) e di mantenimento (compartimentazione organizzativa, azione militare, incapsulamento ideologico, inclusione militante)
<b>Modello 3N (Needs, Narrative,</b>	Soddisfazione/recupero/ evitamento della perdita del bisogno di significato (stima, successo, competenza, controllo, rispetto

<b>Network)</b>	<p>etc.)</p> <p>Reti sociali (famiglia, amici, colleghi, altri significativi, piattaforme di social media, riviste)</p> <p>Narrativa (ideologia religiosa o secolare promotrice di violenza, narrativa di umiliazione e/o sopruso, valori sacri, doveri verso il gruppo, legittimazione della violenza come mezzo)</p>
-----------------	--



## DATASET UTILIZZATO

Il World Values Survey è un programma di ricerca internazionale per lo studio dei valori sociali, politici, economici, religiosi e culturali, a livello mondiale, il cui obiettivo complessivo è rilevare il loro impatto sullo sviluppo sociale, politico ed economico dei Paesi coinvolti.

Il suo principale strumento di ricerca è un sondaggio basato su campioni rappresentativi, condotto a livello globale ogni 5 anni, la cui prima versione risale al quinquennio 1981-1986.

La raccolta dati è effettuata mediante questionari somministrati in presenza o a distanza, e attuata da ricercatori singoli o da gruppi di ricerca sotto la guida di un responsabile del progetto a livello nazionale, che a sua volta riferisce all'infrastruttura principale.

Il campionamento della popolazione avviene su base casuale tra la popolazione adulta.

Al termine dei cinque anni il prodotto finale è un dataset che viene reso pubblicamente disponibile sul sito dell'organizzazione (<https://www.worldvaluessurvey.org/>)

Attualmente il dataset è alla sua settima versione e contiene dati raccolti tra il 2017 e il 2022, in 66 paesi (Haerpfer et al., 2022) .

Il progetto copre un ampio spettro di tematiche di interesse per diverse discipline delle scienze sociali, come sociologia, scienze politiche, relazioni internazionali e psicologia sociale, per citarne alcune. Il questionario è strutturato attorno a 14 sezioni tematiche che includono:

- Valori sociali, atteggiamenti e stereotipi
- Benessere sociale
- Capitale sociale, fiducia e appartenenza organizzativa
- Valori economici
- Corruzione
- Immigrazione
- Indici di post-materialismo
- Scienza e tecnologia
- Valori religiosi
- Sicurezza
- Norme e valori etici
- Interesse e partecipazione politica
- Cultura e regime politico
- Variabili demografiche

Ogni sezione tematica contiene un numero variabile di domande, con risposta su scala nominale oppure su scala ordinale da 4 a 10 livelli. Alcuni paesi presentano domande aggiuntive in base agli accordi presi dai responsabili nazionali con altri enti, in accordo con l'infrastruttura centrale.

Ogni item prevede inoltre la possibilità di 'non risposte' (-1.Don't know, -2.No answer, -3.Not applicable (Not first choice), -4. Not asked, -5. Missing; Unknown).

Il dataset si distingue per l'elevata qualità dei dati e il suo utilizzo diffuso. Una ricerca per parola chiave "World Valu\* surve\*" tra gli abstract della banca dati Scopus, restituisce ad oggi (15/12/24) 2233 risultati.

### SCELTA DEL CAMPIONE

Il campione è stato scelto per ragioni teoriche e di convenienza. L'Armenia è un paese che ha fatto esperienza di alcuni degli antecedenti considerati dall'approccio esplicativo basato sulle cause principali: antecedenti storici (violenza politica), discriminazione su base etnica o religiosa, repressione da parte di forze occupanti straniere.

La Repubblica dell'Armenia è uno Stato dell'Asia occidentale, da un punto di vista storico-culturale considerabile europeo (Lombaerde, Baert, & Felício, 2012) e appartenente al consiglio d'Europa.

Il terrorismo armeno è nato in seno a sentimenti di stampo nazionalista.

La sua prima manifestazione risale al 26 Agosto 1896, con l'occupazione della Banca Centrale Ottomana di Istanbul e la presa in ostaggio di 156 civili da parte della Federazione Rivoluzionaria Armena, che esitò nella conquista di una rappresentazione politica al parlamento dell'Impero Ottomano.

Gli inizi del XX secolo ne hanno sancito la crescita in ampiezza e intensità, con episodi di omicidio a sfondo politico contro ufficiali turchi, azeri e russi, entro e al di fuori dei propri confini nazionali, come anche in incursioni in territorio straniero, come ad esempio avvenuto in Azerbaigian nel Marzo del 1918, dove si stima siano morte più di 50000 persone, tra civili e soldati azeri.

Alla metà del XX secolo, al terrorismo armeno si attribuiscono più di 200 attacchi attuati in quattro diversi continenti, inclusi Europa, Asia, America e Australia. La seconda metà del XX secolo vede le azioni terroristiche legate principalmente a una questione territoriale che coinvolge Armenia e Azerbaigian, per il controllo dei territori nel Nagorno-Karabakh. L'ultimo episodio di terrorismo risale al 2015, ma frequenti sono le scaramucce nei territori al confine con quella regione, che nel corso del tempo sono spesso esitate in morti da entrambe le parti.

## PREPROCESSAMENTO E ANALISI DEI DATI

### LIBRERIE USATE

Il preprocessing e l'analisi dei dati sono stati effettuati utilizzando l'interfaccia JupyterLab (versione 4.0.11), all'interno dell'ambiente Anaconda (versione 24.11.1), basato su Python (versione 3.12.4). Sono state inoltre usate le seguenti librerie:

- Pandas (versione 2.2.2) per la manipolazione e la pulizia dei dati;
- Seaborn (versione 0.13.2) per la visualizzazione dei grafici;
- Statsmodels (versione 0.14.2) per le analisi dei modelli statistici e il calcolo degli outliers;
- Scikit-learn (versione 1.4.2) per la trasformazione dei dati e le analisi di regressione con coefficiente di penalizzazione;
- SciPy (versione 1.11.4) per il calcolo numerico

### SELEZIONE DEI PREDITTORI

Gli item sono stati selezionati sulla base dell'aderenza concettuale alle determinanti considerate nella tabella sinottica riportata alle pagine precedenti (Tabella 1).

Laddove gli item presentavano formulazione e contenuto simile, oppure si aveva una ricodifica dello stesso item su scale di misura diverse, si sono scartati gli item meno comprensivi dal punto di vista del contenuto – per favorire la generalizzazione dei risultati - e quelli con più livelli di risposta – per ridurre la dimensionalità dei dati nel momento in cui si fossero codificate le variabili dummy.

La procedura è stata ripetuta dall'autore tre volte per garantire un certo grado di coerenza interna nella scelta delle variabili.

In totale sono stati selezionati 93 predittori.

### PREPARAZIONE E PULIZIA DEI DATI

Il dataset al momento dell'importazione presenta 97220 osservazioni e 613 colonne.

In seguito alla selezione delle colonne e al filtraggio dei dati per il campione armeno, il dataset contiene 1223 osservazioni e 93 colonne.

Per rendere più agevole l'interpretazione dei risultati si è scelto di eliminare le righe che presentano valori 'missing' (-1,-2,-3,-4,-5), portando il dataset a 332 osservazioni. All'inizio si era scelto di optare per un'imputazione dei dati basata sulla mediana, tuttavia l'elevata numerosità delle sostituzioni da fare avrebbe trasformato in maniera sostanziale le distribuzioni dei predittori.

Le colonne selezionate sono state successivamente rinominate.

In psicologia sociale – disciplina di elezione dello scrivente – è pratica comune trattare variabili misurate su scala ordinale con 5 o più livelli di risposta, come se fossero continue (e.g., Robitzsch 2021). Si è pertanto scelto di aderire a questa pratica, e si sono trattate come continue tutte le variabili con scala di risposta a 5 o più livelli, e per le rimanenti è stata usata una codifica dummy, portando il numero totale dei predittori a 147.

È stata infine calcolata l'asimmetria delle distribuzioni delle variabili considerate continue, e si è proceduto alla loro trasformazione per valori di asimmetria inferiori a -1 e superiori a +1.

#### OBIETTIVI DEL PRESENTE LAVORO:

- stimare la bontà di adattamento di un modello esplicativo delle determinanti dell'approvazione del terrorismo come strumento di affermazione politica, ideologica e religiosa, in un campione armeno;
- ridurre la dimensionalità del modello allo scopo, di aumentarne la precisione, la robustezza, e migliorarne l'interpretabilità.

L'elenco dei predittori considerati, inclusi i livelli di misura prima della trasformazione dei predittori categoriali, è visionabile nel notebook.

La variabile responso usata è Justifiable - Terrorism as a political, ideological or religious mean (1.Never justifiable - 10.Always justifiable)

#### METODOLOGIA ADOTTATA PER LE ANALISI

Inizialmente, considerando la natura ordinale dei dati, si era deciso di usare un modello di regressione logistica ordinale. Tuttavia la matrice Hessiana è risultata non invertibile, compromettendo la possibilità di stimare i parametri del modello.

La non invertibilità della Hessiana si verifica quando la matrice non è definita positiva o è mal condizionata, spesso a causa di multicollinearità o di una matrice dei dati scarsamente informativa.

Nel caso del modello testato si presentano entrambi i casi.

Per la multicollinearità, quando una variabile categoriale con risposta su  $n$  livelli è trasformata in dummy, il risultato sono  $n-1$  variabili altamente correlate tra loro. Il dataset trasformato presenta 54 predittori dummy, altamente correlati tra loro.

Per quanto riguarda l'informazione nella matrice dati, essa dipende, tra le altre cose, anche dal rapporto tra predittori e osservazioni. In questo caso si ha un rapporto elevato (147/332), che associato alla numerosità delle dummy introduce ridondanza, rendendo l'Hessiana non invertibile.

Per risolvere il problema della non invertibilità, si è ridotta la dimensione dei dati, eliminando i predittori con valore di VIF maggiore di 5, tuttavia senza risolvere il problema.

Si è quindi optato per un modello di regressione multipla con coefficienti di primo grado, considerando i predittori con cinque o più livelli ordinati come variabili continue.

#### *MODELLO OLS (OLS)*

Il modello OLS presenta 332 osservazioni e 147 variabili.

Il modello si adatta meglio ai dati rispetto al modello nullo ( $F = 2.403$ ,  $p < .05$ ), presenta un valore di R-quadro aggiustato di 0.384 e un errore quadratico medio (MSE) di 0.342, valori di skewness accettabili (skewness = 0.703), frequenti valori estremi (curtosi = 4.757), e residui non autocorrelati (Durbin Watson < 2.5).

La distribuzione dei residui è asimmetrica positiva, non normale ( $\text{Prob}(\text{Omnibus}) < 0.05$  e  $\text{Prob}(\text{Jaque-Bera}) < 0.05$ ), e al confronto con i valori predetti presenta dei pattern.

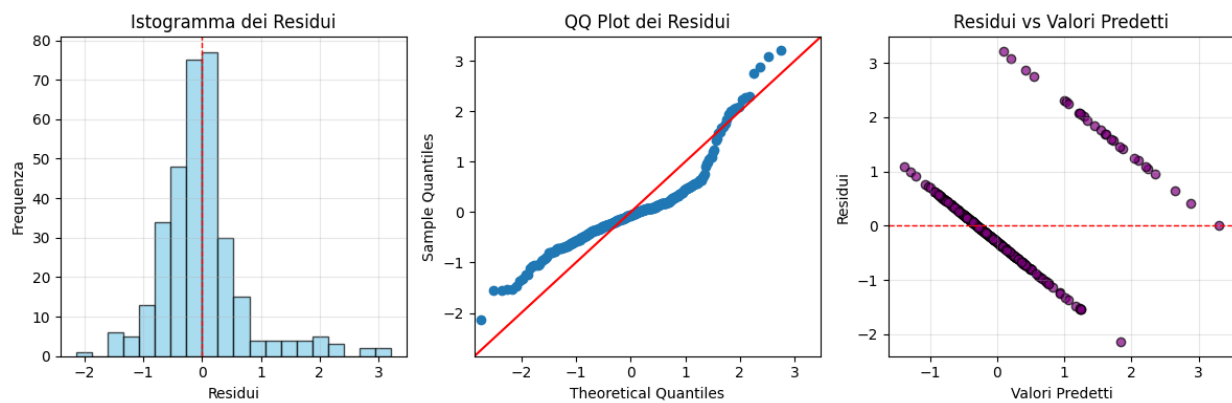
Il modello risulta infine essere particolarmente sensibile all'introduzione di nuove osservazioni (Numero di condizionamento = 605).

Il quadro teorico che si vuole testare non esplicita una soluzione polinomiale, pertanto allo scopo di migliorare le stime del modello e rendere la soluzione più stabile, generalizzabile e interpretabile, si è deciso di applicare un metodo di riduzione della dimensionalità noto come Sure Independence Screening (SIR; Fan & Lv, 2008), che seleziona i predittori ordinando in modo decrescente il valore assoluto dei coefficienti univariati dei predittori standardizzati, stimati senza intercetta, e scegliendo tra di essi i primi  $n/\log(n)$  predittori, dove  $n$  è la numerosità campionaria.

Al termine della selezione delle variabili il dataset è costituito da 332 osservazioni e 57 predittori.

#### *MODELLO OLS RIDOTTO (OLS\_rid)*

Il modello OLS ridotto si adatta meglio ai dati rispetto al modello nullo ( $F = 4.334$ ,  $p < .05$ ), ma con indici R-quadro aggiustato (0.365) e MSE (0.526) peggiori del precedente. Rimangono accettabili i valori di skewness (skewness = 1.363), aumenta la curtosi (curtosi = 6.978), e i residui rimangono non autocorrelati (Durbin Watson < 2.5).



La distribuzione dei residui è asimmetrica positiva, non normale ( $\text{Prob}(\text{Omnibus}) < 0.05$  e  $\text{Prob}(\text{Jaque-Bera}) < 0.05$ ), e al confronto con i valori predetti presenta pattern.

Il modello risulta meno sensibile del precedente all'introduzione di nuove osservazioni, ma comunque il numero di condizionamento rimane elevato (Numero di condizionamento = 68.5).

Il QQ-plot e la distribuzione dei residui rispetto ai valori predetti suggeriscono che un modello di regressione polinomiale e/o l'introduzione di interazioni tra le variabili, potrebbero fornire un modello meglio adattato. Il RESET test di Ramsey conferma questa ipotesi ( $p < 0.05$ ).

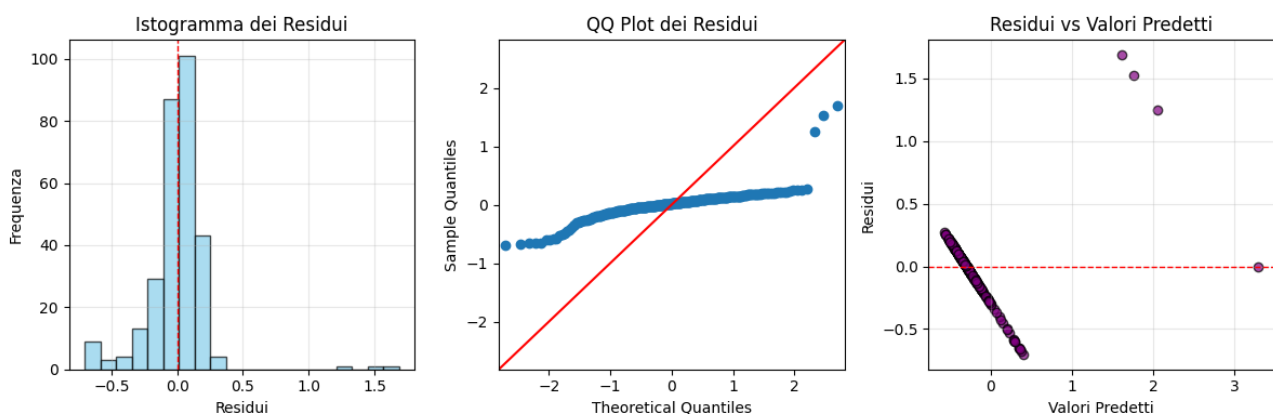
Per migliorare la performance del modello sono successivamente stati eliminati gli outliers, usando come riferimento le distanze di Cook e i residui studentizzati.

Si è poi provveduto a rifittare il modello.

Con l'eliminazione degli outliers il dataset presenta 296 osservazioni e 57 variabili.

### *MODELLO OLS RIDOTTO SENZA OUTLIERS (OLS\_rid\_no\_auth)*

Il modello si adatta meglio ai dati rispetto al modello nullo ( $F = 9.133$ ,  $p < .05$ ). Migliorano rispetto al precedente R-quadro aggiustato (0.611) e MSE (0.054). Aumentano l'asimmetria positiva (skewness = 1.932) e la curtosi (curtosi = 20.633). I residui rimangono non autocorrelati (Durbin Watson  $< 2.5$ ).



Il modello diviene più sensibile all'introduzione di nuove osservazioni (Numero di

condizionamento = 138).

La distribuzione dei residui è asimmetrica positiva, non normale ( $\text{Prob}(\text{Omnibus}) < 0.05$  e  $\text{Prob}(\text{Jaque-Bera}) < 0.05$ ) e al confronto con i valori predetti presenta regolarità.

Il QQ-plot, la distribuzione dei residui e il RESET test di Ramsey ( $p < 0.05$ ) confermano che il modello è mal specificato. Il test di Breusch-Pagan fornisce ulteriori evidenze di eteroschedasticità ( $p < 0.05$ ).

Rifittare il modello con i soli predittori significativi (OLS\_rid\_sign) non migliora sensibilmente le stime (si veda la tabella2 al termine della presentazione dei modelli).

Considerando le problematiche riscontrate nella distribuzione dei residui, come l'eteroschedasticità o la presenza di pattern, per migliorare l'affidabilità delle stime si è deciso di utilizzare un modello di regressione con stimatore robusto.

#### *OLS RIDOTTO CON STIME BASATE SU ERRORI ROBUSTI (OLS\_rid\_rob)*

Per la stima dei parametri si è usato il metodo IRLS (Iteratively Reweighted Least Squares), un metodo iterativo che assegna pesi adattivi a ciascun dato (per una descrizione analitica delle formule si veda il notebook).

La funzione di perdita utilizzata è quella di Huber, che applica funzioni di perdita distinte a seconda dell'entità dei residui rispetto a un parametro di regolarizzazione: lineare per residui piccoli, quadratica per residui grandi.

Per stimare la deviazione standard dei residui si è usata la deviazione assoluta dalla mediana.

Per il controllo dell'eteroschedasticità, nel calcolo della matrice di covarianza tra gli stimatori è stato usato il metodo H1.

Dai risultati si ha un peggioramento del valore di MSE rispetto ai modelli precedenti ( $\text{MSE} = 0.131$ ).

Come ultima analisi, si è pensato opportuno tentare di ridurre ulteriormente la dimensionalità del modello, utilizzando una regressione LASSO.

Il LASSO (Least Absolute Shrinkage and Selection Operator) è una tecnica di regressione basata sulla norma L1, che seleziona i coefficienti del modello, penalizzando alcuni di essi in base a un parametro di regolarizzazione, allo scopo di minimizzare l'errore quadratico medio (per le specifiche tecniche si veda il notebook).

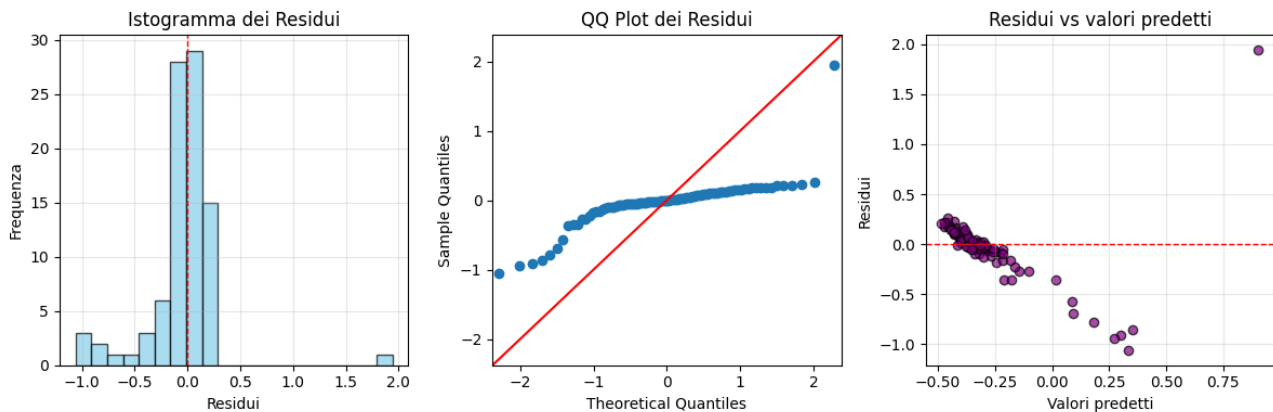
## LASSO (LASSO\_rid)

In primis si è suddiviso il dataset in dati di training e dati di test.

Successivamente, poiché i metodi di penalizzazione non sono invarianti di scala, a differenza del metodo dei minimi quadrati, i predittori sono stati scalati e centrati per renderli a varianza unitaria.

Applicando l'algoritmo di ricerca su 1000 valori di alpha, 10000 iterazioni, e 5-folds cross-validation, è trovata una soluzione a 40 predittori, con MSE pari a 0.115.

I predittori selezionati sono visionabili sul notebook.



I grafici dei residui confermano quanto detto in precedenza per gli altri modelli: I residui non si distribuiscono normalmente, la loro distribuzione è asimmetrica, il modello presenta code grasse ed è mal specificato.

Di seguito si presenta una tabella sintetica dei risultati ottenuti, che include anche i modelli con i soli predittori significativi.

Modello	P	F_value	Adj_r_squared	Log-likelihood	AIC	BIC	MSE
OLS	147	2.40	0.38	-293.21	882.42	1445.58	0.34
OLS ridotto	58	4.33	0.36	-364.40	844.79	1065.49	0.53
OLS rid_no_out	57	9.13	0.61	11.44	93.12	307.16	0.05
OLS rid_sign	9	56.61	0.63	-8.72	37.45	74.35	0.06
OLS robusto_rid	57	--	--	--	--	--	0.13
OLS robusto_rid_sign	13	--	--	--	--	--	0.13
LASSO_rid	40	--	0.41	--	--	--	0.12
LASSO_rid_sign	40	--	0.41	--	--	--	0.13

Guardando esclusivamente all'errore quadratico medio dei modelli, il migliore risulta essere il modello OLS ridotto e senza osservazioni estreme (OLS rid\_no\_out; MSE= 0.05). Confrontandolo con il modello dove i valori estremi non sono stati eliminati (Ols\_ridotto; MSE = 0.53), la differenza nei valori di MSE è molto elevata. In altre parole, i residui e i, valori estremi hanno un impatto elevato sulle stime del modello.



È interessante inoltre notare che fittare un modello ridotto con i soli coefficienti significativi e senza outlier (OLS rid\_sign) non cambia significativamente le stime rispetto al modello precedente (OLS rid\_no\_out), né in termini di MSE (0.05 vs 0.06) né in termini di R-quadro aggiustato (0.61 vs 0.63). R-quadro aggiustato è un parametro che risente della numerosità dei predittori, e penalizza le stime in presenza di predittori non significativi. Pertanto i valori di R-quadro e di MSE simili nei due modelli potrebbero indicare overfitting nel modello ridotto pieno (OLS rid\_no\_out).

Alla LASSO (LASSO\_rid) , con divisione del dataset in dati di training e dati di test, i risultati dimostrano che a un numero di predittori inferiori (40) corrisponde una perdita significativa di capacità esplicativa del modello (R-quadro aggiustato = 0.41) e una perdita di precisione delle stime (MSE = 0.12) rispetto ai modelli pieni.

Il modello LASSO è stato sottoposto a validazione, pertanto è un modello più robusto, da ciò segue che molto probabilmente il modello pieno è sovradattato ai dati.

Tuttavia, è interessante notare come utilizzare un modello robusto (OLS robusto\_rid) non validato, che penalizza residui elevati, peggiori la bontà di adattamento del modello pieno (MSE = 0.13). Se il modello pieno fosse sovradattato, le stime del modello robusto non dovrebbero peggiorare di molto, dato che esso non presenta outliers (si veda il notebook).

Da ciò segue che probabilmente sono i valori di curtosi elevata a distorcere le stime.

Se oltre a ciò si considerano le distribuzioni dei residui e i risultati al test di Ramsey, emerge chiaramente come il modello non sia ben specificato.

Da quanto visto sulle distribuzioni dei residui (eteroschedasticità, non normalità, presenza di pattern) e in base alle considerazioni appena fatte, sembra pertanto lo stimatore ai minimi quadrati non sia adeguato allo studio del modello.

Pertanto, sembra che ai dati attuali sia necessario applicare metodi di analisi diversi da quelli utilizzati.

## CONCLUSIONI

Dalle analisi effettuate emergono alcune conclusioni.

La prima di esse riguarda la pratica, diffusa nelle scienze sociali e soprattutto in psicologia sociale, di considerare come continue variabili di tipo categoriale, allo scopo di applicare metodi di stima dei parametri basati sui minimi quadrati. Per ogni modello considerato, le distribuzioni dei residui non rispettano le prime due condizioni di Gauss-Markov (rumore a media nulla, omoschedasticità, non correlazione tra errori). Ciò implica la possibilità che metodi basati su altri stimatori possano ottenere risultati superiori in termini di tradeoff bias-varianza e di precisione delle stime.

Un secondo aspetto interessante riguarda i valori elevati di curtosi nei modelli visti.

Valori elevati di curtosi indicano che la distribuzione ha code "più pesanti" (ossia valori estremi più frequenti) rispetto alle code di una distribuzione normale. Valori elevati di curtosi rendono le stime degli intervalli di confidenza e i valori medi meno affidabili per fare previsioni.

Nei termini del modello in relazione alla variabile responso, ciò vuol dire che la propensione ad approvare il terrorismo come strumento politico, ideologico o religioso, non è ben predetta dal modello stimato.

Quanto discusso chiama in causa un terzo aspetto.

Un modello ben specificato è un modello che spiega adeguatamente la relazione tra predittori e responso. Se si guardano la distribuzione degli errori allo scatterplot, la presenza di pattern rispetto ai valori predetti, la loro disposizione al QQ-plot e il test di Ramsey (specificazione) e quello di Breusch-Pagan (eteroschedasticità), tutto suggerisce che il modello sia mal specificato e che sia necessario, se si vuole usare uno stimatore OLS, inserire altri predittori e/o usare termini polinomiali o di interazione.

Infine, il peggioramento della performance del modello usando la regressione LASSO e usando una modellazione degli errori robusta indicano che il metodo di stima basato sui minimi quadrati non sia adeguato a spiegare la relazione tra le variabili e il responso, a causa della natura dei dati e dalla presenza di 'code grasse'.

Quanto visto sin'ora dimostra che per lo studio del modello ipotizzato usando i dati presenti nel World Values Survey, sarebbe più opportuno utilizzare metodi di classificazione che, se possibile, tengano anche adeguatamente conto dell'elevata frequenza di valori estremi, indicati dai valori di curtosi elevata in ogni modello testato. I vantaggi sarebbero molteplici.

Il primo sarebbe sicuramente di non dipendere dai vincoli imposti dagli assunti dei modelli lineari e le trasformazioni che spesso, a causa di ciò, sono necessarie per avere stime precise.

I metodi di classificazione, inoltre, rendono più semplice il trattamento dei predittori categoriali e, in generale, più semplice la rilevazione di pattern complessi nei dati.

Tali metodi, inoltre, rendono più semplice l'interpretazione dei risultati delle analisi, poiché basata su metriche di performance più intuitive (accuratezza, precisione, richiamo, F1-score, area sotto la curva ROC).

Rispetto al problema dell'elevata dimensionalità del dataset, l'uso di metodi di regressione ordinale con parametro di regolarizzazione (Wurm, Rathouz, & Hanlon, 2021) permetterebbe di aggirare il problema dell'interpretabilità del modello a fronte dell'elevato numero di predittori.

Infine, considerati gli elevati valori di curtosi, non si esclude che risultati interessanti potrebbero essere ottenuti modellando direttamente i valori estremi, utilizzando gli strumenti offerti dalla statistica degli eventi estremi (EVT).

I tentativi, fatti inizialmente, di applicare una regressione logistica ordinale considerando tutti i predittori come variabili categoriali, ha generato problemi nella stima dei parametri del modello (non invertibilità dell'Hessiana). Ciò suggerisce che in futuro possa esser utile un'attenzione maggiore alla scelta delle variabili e al pretrattamento dei dati.

Ciò fornisce l'introduzione ai limiti del presente lavoro.

Non è stata data la necessaria attenzione alle distribuzioni delle variabili: un'analisi descrittiva delle variabili potrebbe fornire indicazioni utili per analisi successive.

Il secondo limite riguarda la scelta di eliminare i valori considerati 'missing' (-1, -2, -3, -4, -5). L'eliminazione di osservazioni implica la perdita di informazione. Nel caso di dimensionalità elevata l'Hessiana, per questa ragione, potrebbe non essere invertibile. Studi futuri potrebbero utilizzare metodi di imputazione per aumentare il numero di osservazioni utili per le analisi.

Infine, il modello testato è stato ispirato dalla ricerca teorica ed empirica sul fenomeno del terrorismo ma, nella maggior parte delle teorie considerate, le variabili non sono operazionalizzate in modo diretto. Ciò ha implicato un elevato grado di arbitrarietà nella scelta degli item da considerare come predittori.

La loro selezione, poi, è stata effettuata da parte di un solo *rater*. Studi futuri potrebbero beneficiare di una maggiore precisione nella selezione delle variabili, scegliendo queste ultime esclusivamente in base alla letteratura empirica quantitativa e affidando il processo di selezione a più *raters* indipendenti, così da garantire un modello più affidabile (Saal, Downey, & Lahey, 1980).

## BIBLIOGRAFIA

- Atran, S. (2016). The devoted actor: Unconditional commitment and intractable conflict across cultures. *Current Anthropology*, 57(S13), S192-S203.
- Atran, S., & Sheikh, H. (2015). Dangerous terrorists as devoted actors. In V. Zeigler-Hill, L. L. M. Welling, & T. K. Shackelford (Eds.), *Evolutionary perspectives on social psychology* (pp. 401–416). Springer International Publishing/Springer Nature. [https://doi.org/10.1007/978-3-319-12697-5\\_31](https://doi.org/10.1007/978-3-319-12697-5_31)
- Crayton, J. W. 1983. Terrorism and the psychology of the self. In *Perspectives on terrorism*, L. Z. Freedman & Y. Alexander,(Eds.) 33-41. Wilmington, DE: Scholarly Resources, Inc.
- Della Porta, D. (2013). *Clandestine political violence*. Cambridge University Press.
- Europol (2022), *European Union Terrorism Situation and Trend Report*, Publications Office of the European Union, Luxembourg.
- Enders, W., & Su, X. (2007). Rational terrorists and optimal network structure. *Journal of Conflict Resolution*, 51(1), 33-57.
- Fan, J., & Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 70(5), 849-911.
- Gupta, S. (2004). The Changing Dimensions of International Terrorism And The Role Of The United States: A Comprehensive And Multilateral Approach To Combat Global Terrorism. *The Indian Journal of Political Science*, 556-587.
- Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano J., M. Lagos, P. Norris, E. Ponarin & B. Puranen (eds.). 2022. *World Values Survey: Round Seven - Country-Pooled Datafile Version 5.0*. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. doi:10.14281/18241.24
- Horgan, J. (2008). From profiles to pathways and roots to routes: Perspectives from psychology on radicalization into terrorism. *The ANNALS of the American Academy of Political and Social Science*, 618(1), 80-94.
- Horgan, J. (2014). *The psychology of terrorism* (2nd ed.). London, UK: Routledge.
- Horgan, J. G. (2017). Psychology of terrorism: Introduction to the special issue. *American Psychologist*, 72(3), 199.
- Kalyvas, S.N. (2019)The landscape of political violence. In E. Chenoweth, R. English, A. Gofas, & S.N. Kalyvas (Eds.), *The Oxford Handbook of Terrorism* (pp. 10-33). Oxford University Press.
- Kruglanski, A. W., Bélanger, J. J., & Gunaratna, R. (2019). *The three pillars of radicalization: Needs, narratives, and networks*. Oxford University Press.

- Kruglanski, A. W., Molinario, E., Jasko, K., Webber, D., Leander, N. P., & Pierro, A. (2022). Significance-quest theory. *Perspectives on Psychological Science*, 17(4), 1050-1071.
- Lankford, A. (2013). A comparative analysis of suicide terrorists and rampage, workplace, and school shooters in the United States from 1990 to 2010. *Homicide studies*, 17(3), 255-274.
- Lombaerde, P., Baert, F., & Felício, T. (Eds.).(2012). *The United Nations and the regions: Third world report on regional integration* (Vol. 3). Springer Science & Business Media.
- McCauley, C., & Moskaleiko, S. (2008). Mechanisms of political radicalization: Pathways toward terrorism. *Terrorism and political violence*, 20(3), 415-433.
- Moskaleiko, S., & McCauley, C. (2011). The psychology of lone-wolf terrorism. *Counselling psychology quarterly*, 24(2), 115-126.
- Pape, R. A. (2006). *Suicide terrorism and democracy: What we've learned since 9/11*. Cato Institute.
- Robitzsch, A. (2022). On the bias in confirmatory factor analysis when treating discrete variables as ordinal instead of continuous. *Axioms*, 11(4), 162.
- Saal, F. E., Downey, R. G., & Lahey, M. A. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin*, 88(2), 413–428. <https://doi.org/10.1037/0033-2909.88.2.413>
- Szumowska, E., Kruglanski, A. W., & Kopetz, C. (2023). Extremism and the extreme personality. Goal Systems Theory. *Psychological Processes and Applications*, 239-256.
- UNGA resolution 49/60 (1994), annexed Declaration on Measures to Eliminate International Terrorism.
- Webber, D., & Kruglanski, A. W. (2018). The social psychological makings of a terrorist. *Current opinion in psychology*, 19, 131-134.
- Whittaker, J. (2022). *Online Radicalisation. What we know*. Luxembourg: Publications Office of the European Union.
- Wurm, M. J., Rathouz, P. J., & Hanlon, B. M. (2021). Regularized Ordinal Regression and the ordinalNet R Package. *Journal of Statistical Software*, 99(6), 1–42.  
<https://doi.org/10.18637/jss.v099.i06>

## SITOGRAFIA

- Parlamento Europeo. (2021). *Come fermare il terrorismo: quali sono le misure adottate dall'UE?*.  
<https://www.europarl.europa.eu/topics/it/article/20180316STO99922/come-fermare-il-terrorismo-quali-sono-le-misure-adottate-dall-ue#:~:text=Il%20programma%20sopracitato%2C%20operer%C3%A0%20in,1%2C9%20miliardi%20di%20euro> consultato il 17 dicembre 2024