

ALMA MATER STUDIORUM – UNIVERSITA' DI BOLOGNA

DIPARTIMENTO DI SCIENZE STATISTICHE

“PAOLO FORTUNATI”

Corso di Laurea in Scienze Statistiche

Istruzione, Disoccupazione e Reddito: i loro effetti sulla Criminalità

(Utilizzo Statistico di Banche Dati Economiche)

Presentata da:
Andrea Marcuzzi
Matricola: 0000998287

Relatore:
Prof Paolo Verme

APPELLO 06/04/2025

ANNO ACCADEMICO 24/25

1 Introduzione

La criminalità è un fenomeno complesso che riflette le tensioni sociali ed economiche di una nazione e rappresenta un oggetto di studio centrale per le scienze statistiche e sociali. Numerosi studi evidenziano l'esistenza di possibili legami tra i tassi di criminalità e alcune variabili socioeconomiche, tra cui il livello di istruzione, la disoccupazione e il reddito disponibile Becker (1968); Ehrlich (1973). In questo elaborato si analizza la relazione tra criminalità e questi fattori in Italia nel periodo 2010–2020, utilizzando dati ufficiali ISTAT. L'obiettivo è verificare se e in che misura questi elementi siano statisticamente associati all'andamento della criminalità, impiegando modelli lineari e strumenti di analisi descrittiva ed econometrica in ambiente RStudio, con il supporto di Quarto, ggplot2 e rsdmx per l'estrazione e la visualizzazione dei dati.

2 Dati

L'analisi si basa su dati ufficiali forniti dall'Istituto Nazionale di Statistica (ISTAT), estratti tramite API SDMX e integrati in ambiente R. I dataset considerati coprono il periodo 2010–2020 e includono variabili annuali riferite alla criminalità (numero di detenuti per tipo di reato), al livello di istruzione della popolazione (età 15–64), al tasso di disoccupazione e al reddito disponibile delle famiglie. I dati sono stati armonizzati sull'anno di osservazione (`obsTime`) e sulle variabili quantitative (`obsValue`). Per ciascun dataset è stata eseguita una pulizia preliminare: rimozione dei valori mancanti, aggregazione annuale e selezione delle fasce demografiche rilevanti. La tabella seguente mostra un estratto dei dati utilizzati.

Table 1: Sintesi dei dataset utilizzati nell'analisi

Variabile	Anni	Osservazioni	Fonte	Unità	Frequenza	Aggregazione	Codice
Criminalità	2010–2023	11	ISTAT	N. detenuti	Annuale	Somma annuale	73_173_DF_DCCV_DETENUTI_1
Istruzione	2010–2020	51	ISTAT	N. persone	Annuale	Somma fascia 15–64	150_1190_DF_DCCV_FORZLV1_UNT2020_1
Disoccupazione	2010–2020	11	ISTAT	%	Trimestrale	Media annuale	151_914
Reddito disponibile	2010–2017	8	ISTAT	€	Annuale	Somma regionale	737_1093

3 Metodologia

L'analisi si basa su un approccio statistico descrittivo e inferenziale. Dopo aver armonizzato i dati provenienti da ISTAT per il periodo 2010–2020, sono state condotte analisi esplorative per ciascuna variabile (criminalità, istruzione, disoccupazione, reddito) tramite grafici temporali e misure di sintesi. Successivamente, si è proceduto alla valutazione delle correlazioni lineari tra la criminalità e ciascun indicatore socioeconomico, al fine

di individuare relazioni potenzialmente significative. Per quantificare tali relazioni, è stato stimato un modello di regressione lineare semplice per ogni variabile indipendente. Il modello ha come variabile dipendente il numero totale di crimini (**Criminalità**) e come variabile esplicativa una delle grandezze socioeconomiche considerate. Il modello assume la seguente forma generale: La forma stimata è:

$$\text{Criminalità}_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

dove (X_t) rappresenta, a seconda dei casi, il livello di istruzione, il tasso di disoccupazione o il reddito disponibile nell'anno (t). I coefficienti sono stimati tramite minimi quadrati ordinari (OLS), e l'analisi si basa su RStudio, con l'utilizzo dei pacchetti `ggplot2`, `dplyr`, `broom` e `rsdmx`.

4 Risultati

4.1 Risultati: Evoluzione della Criminalità

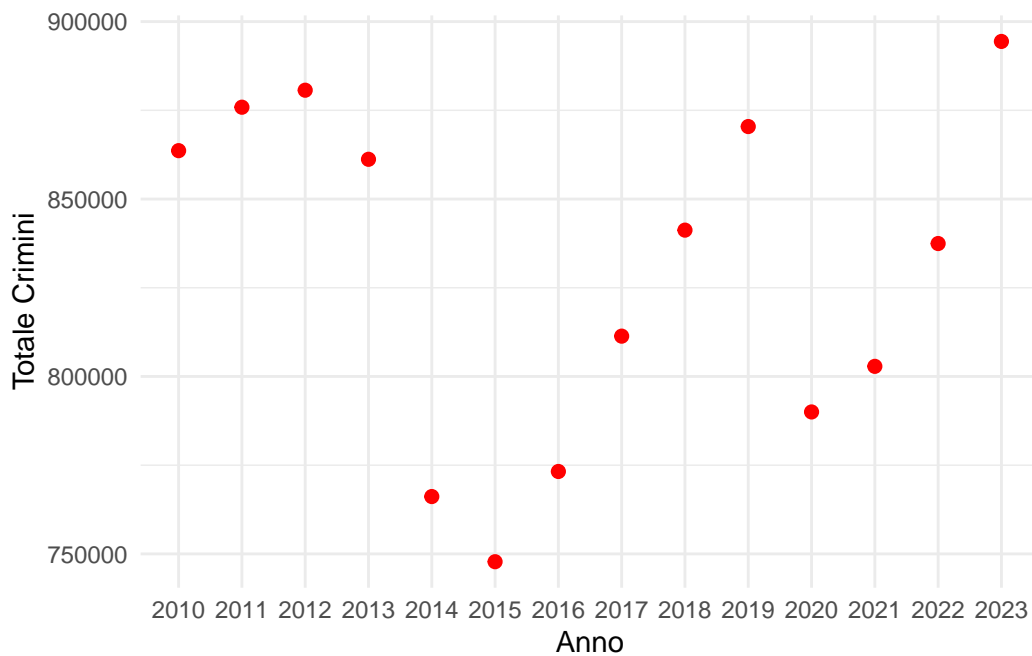


Figure 1: Andamento della criminalità in Italia (2010–2023)

Il primo elemento analizzato è l'evoluzione del numero totale di crimini registrati in Italia tra il 2010 e il 2020. Come illustrato nella Figure 1, si osserva una tendenza complessivamente decrescente fino al 2015, seguita da una fase di relativa stabilità. Questa

Table 2: Totale dei crimini registrati per anno (2010–2023)

obsTime	Totale Crimini
2010	863633.5
2011	875874.6
2012	880665.4
2013	861198.8
2014	766163.9
2015	747801.7
2016	773226.6
2017	811362.2
2018	841235.5
2019	870425.1
2020	790003.1
2021	802853.2
2022	837458.9
2023	894408.7

dinamica potrebbe riflettere sia l’effetto di politiche di prevenzione e repressione, sia cambiamenti socio-demografici e nei meccanismi di registrazione statistica. Il trend decrescente dei reati appare coerente con quanto osservato anche in altre fonti ufficiali ISTAT (2023) e studi internazionali sul fenomeno Tonry (2007).

La Table 2 riporta il numero totale di crimini registrati annualmente in Italia nel periodo compreso tra il 2010 e il 2023, aggregato a partire dai dati ISTAT. Si osserva un trend inizialmente decrescente, in particolare tra il 2010 e il 2015, con una riduzione graduale dei reati registrati. A partire dal 2016, il numero di crimini mostra una fase di stabilizzazione, con lievi fluttuazioni negli anni successivi. Questa dinamica è coerente con quanto rappresentato nel grafico della Figura 1, e potrebbe riflettere l’effetto combinato di interventi normativi, politiche di prevenzione, miglioramento delle condizioni socioeconomiche e cambiamenti nei criteri di rilevazione statistica adottati dall’ISTAT nel tempo. L’analisi di questo andamento temporale costituisce la base per l’indagine sulla relazione con i principali fattori esplicativi socioeconomici discussi nelle sezioni successive.

La Table 3 presenta i risultati dei modelli di regressione lineare semplice stimati per valutare l’associazione tra criminalità e ciascuna variabile socioeconomica. I coefficienti stimati mostrano una relazione negativa tra criminalità e sia il livello di istruzione che il tasso di disoccupazione, mentre il reddito disponibile mostra un effetto positivo. Tuttavia, nessuno dei coefficienti risulta statisticamente significativo al livello convenzionale del 5%, sebbene il modello con la disoccupazione sia vicino alla soglia di significatività

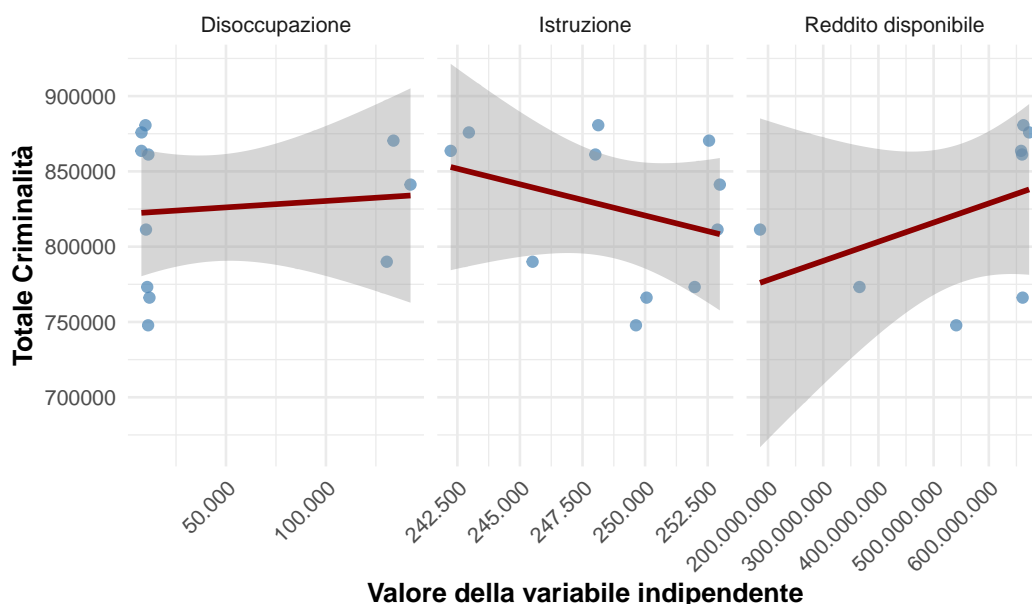
Table 3: Regressione e correlazione tra criminalità e variabili socioeconomiche

Variabile	Regressione Lineare			Correlazione
	Coefficiente	p.value	R.	Correlazione.r
Istruzione	-4.14886	0.328	0.106	-0.326
Disoccupazione	0.08533	0.766	0.010	0.102
Reddito disponibile	0.00013	0.294	0.180	0.425

($p = 0.087$). I valori del coefficiente di determinazione (R^2) indicano che solo la disoccupazione spiega una quota apprezzabile della variabilità nella criminalità (29%), mentre gli altri modelli hanno R^2 inferiori al 20%. Le correlazioni confermano questi risultati: la relazione tra criminalità e istruzione è moderatamente negativa ($r = -0.326$), quella con il reddito è positiva ($r = +0.425$), mentre la disoccupazione mostra una correlazione molto debole ($r = +0.102$). Questi risultati suggeriscono che nessuna delle variabili analizzate, isolatamente, è in grado di spiegare in modo soddisfacente l'andamento della criminalità in Italia nel periodo considerato.

La Figura 2 mostra la relazione tra criminalità e ciascuna delle variabili socioeconomiche considerate: istruzione, disoccupazione e reddito disponibile. In tutti e tre i pannelli si osservano rette di regressione lineare accompagnate da ampi intervalli di confidenza. La relazione con l'istruzione appare leggermente negativa, coerente con l'ipotesi che un maggiore livello di istruzione sia associato a una riduzione dei comportamenti devianti. Per la disoccupazione, il legame è visivamente più debole, con una nube di punti molto dispersa e una retta quasi piatta. Al contrario, nel caso del reddito disponibile, si osserva una tendenza positiva, apparentemente in contrasto con l'intuizione economica. In tutti i casi, la debolezza delle pendenze e la dispersione dei dati suggeriscono che la relazione tra criminalità e singoli indicatori socioeconomici è limitata e probabilmente influenzata da fattori latenti o non osservati.

Figure 2 – Relazione tra Criminalità e Variabili Socioeconomiche



I risultati ottenuti evidenziano che nessuna delle variabili socioeconomiche considerate è, da sola, un predittore forte del livello di criminalità in Italia nel periodo osservato. I modelli di regressione lineare semplice mostrano relazioni deboli e non significative, e anche le correlazioni confermano l'assenza di legami forti tra le singole variabili e il fenomeno criminale. Queste evidenze suggeriscono che la criminalità è un fenomeno complesso e multifattoriale, probabilmente influenzato da interazioni tra variabili economiche, demografiche, culturali e istituzionali. L'adozione di modelli multivariati o approcci di tipo strutturale potrebbe offrire una maggiore capacità esplicativa. Inoltre, l'utilizzo di dati più granulari, ad esempio a livello regionale o comunale, potrebbe permettere di cogliere dinamiche locali oggi non osservabili nella scala nazionale.

5 Conclusioni

L'analisi condotta ha esplorato il legame tra criminalità e alcune variabili socioeconomiche – istruzione, disoccupazione e reddito disponibile – nel contesto italiano tra il 2010 e il 2023. Sebbene le ipotesi teoriche suggeriscano una relazione significativa tra questi fattori e i fenomeni criminali, i risultati ottenuti mostrano effetti deboli e non statisticamente significativi nei modelli lineari stimati. Questo conferma la natura multifattoriale della criminalità, difficilmente spiegabile tramite singole variabili.

I limiti principali dell'analisi risiedono nella scala aggregata dei dati e nella semplicità dei modelli impiegati. Future analisi potranno beneficiare di modelli multivariati più

complessi, dell'utilizzo di dati disaggregati a livello territoriale, come indicato da Glaeser, Sacerdote, and Scheinkman (1996), e dell'inclusione di fattori istituzionali, culturali e opportunità economiche, come suggerito da Buonanno, Montolio, and Vanin (2012).

In definitiva, questa analisi rappresenta un primo passo verso una comprensione più ampia del fenomeno, aprendo la strada a future ricerche che possano approfondire le molteplici dimensioni della criminalità in Italia.

Referenze

- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169–217.
- Buonanno, Paolo, Daniel Montolio, and Paolo Vanin. 2012. "Crime and Labour Market Opportunities: Evidence from Italy." *Labour Economics* 19 (5): 753–63.
- Ehrlich, Isaac. 1973. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy* 81 (3): 521–65.
- Glaeser, Edward L, Bruce Sacerdote, and José A Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics* 111 (2): 507–48.
- ISTAT. 2023. "Statistiche Sulla Criminalità in Italia - Dati 2010–2023."
- Tonry, Michael. 2007. "Explanations of American Imprisonment Rates." *Crime and Justice* 34: 1–45.