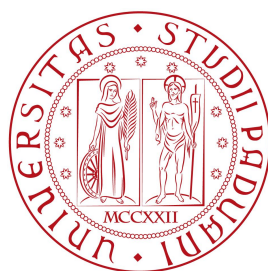


Università degli Studi di Padova  
Dipartimento di Scienze Statistiche  
Corso di Laurea Triennale in

Statistica per l'Economia e l'Impresa



**LA SOLITUDINE DEGLI ANZIANI IN EUROPA:  
ANALISI SULLE DIFFERENZE PRIMA E DURANTE LA  
PRIMA ONDATA DEL COVID-19**

**Relatore:** Prof. Omar Paccagnella

Dipartimento di Scienze Statistiche

**Laureando:** Veronica Cassarà

Matricola n. 1226856

Anno Accademico 2022/2023



# Indice

<b>Elenco delle figure</b>	<b>2</b>
<b>Elenco delle tabelle</b>	<b>5</b>
<b>Introduzione</b>	<b>7</b>
<b>1 La solitudine e i suoi significati</b>	<b>9</b>
1.1 Solitudine e anziani: l'impatto del COVID-19 . . . . .	12
1.2 Il ruolo della tecnologia durante la crisi COVID-19: il divario digitale . . . . .	14
<b>2 Descrizione del <i>Dataset</i></b>	<b>17</b>
2.1 Il campione dell'indagine . . . . .	19
2.1.1 Data Wrangling and Cleaning . . . . .	21
2.2 Esplorazione preliminare dei dati: analisi sul campione .	25
2.2.1 Analisi descrittiva per età . . . . .	25
2.2.2 Analisi descrittiva per genere . . . . .	27
2.2.3 Analisi descrittiva per titolo di studio . . . . .	27
2.2.4 Analisi descrittiva per Paese . . . . .	31
2.2.5 Analisi descrittiva per la presenza del partner . .	37
2.2.6 Analisi descrittiva per numero di figli . . . . .	39

<b>3</b>	<b>Analisi dei fattori associati alla solitudine degli anziani durante la pandemia di COVID-19</b>	<b>41</b>
3.1	Una misura di solitudine: introduzione alla variabile risposta e alle variabili esplicative . . . . .	41
3.2	Analisi visiva della correlazione tra solitudine e fattori di contesto . . . . .	44
3.2.1	Solitudine e genere . . . . .	45
3.2.2	Solitudine e Paese . . . . .	46
3.2.3	Solitudine ed età . . . . .	48
3.2.4	Solitudine e presenza del partner e/o figli . . . . .	51
3.2.5	Solitudine, titolo di studio e gestione economica . . . . .	58
<b>4</b>	<b>Analisi multivariata: Modelli di regressione lineare multipla</b>	<b>65</b>
4.0.1	Le variabili esplicative presenti nel modello di regressione . . . . .	67
4.0.2	Stima del modello di regressione lineare multipla completo . . . . .	68
4.0.3	Selezione del modello tramite criterio dello step AIC	73
4.0.4	La creazione di altre due variabili dipendenti per un'analisi più approfondita . . . . .	76
<b>5</b>	<b>Conclusioni</b>	<b>83</b>
	<b>Bibliografia</b>	<b>87</b>

# Elenco delle figure

2.1	Istogramma della distribuzione del campione per classe di anno di nascita. . . . .	26
2.2	Grafico a torta che illustra la distribuzione del campione in base al sesso. . . . .	27
2.3	Grafico a torta: ripartizione della popolazione in base al livello di istruzione. . . . .	28
2.4	Istogramma sulla distribuzione del giudizio sulla capacità di far quadrare i conti in base al titolo di studio e al lavoro. . . . .	30
2.5	Istogramma sulla distribuzione della popolazione in base al Paese di provenienza. . . . .	32
2.6	Istogramma sulla distribuzione del titolo di studio della popolazione, per Paese. . . . .	33
2.7	Istogramma sulla distribuzione percentuale del titolo di studio per Paese. . . . .	34
2.8	Istogramma sul confronto tra la presenza e l'assenza di partner nella popolazione. . . . .	37
2.9	Istogramma per la distribuzione della presenza di un partner in base al sesso nel campione di studio. . . . .	38

2.10	Istogramma per la distribuzione percentuale del numero di figli nella popolazione, suddivisi in 4 categorie: 1 figlio, 2 figli, 3 figli, e 4 o più figli. . . . .	39
3.1	Istogramma della variazione % di solitudine in base al sesso. . . . .	45
3.2	Istogramma sulla media della differenza di solitudine tra prima e durante il Covid-19 per ogni Paese europeo. . .	47
3.3	Istogramma della variazione % di solitudine in base all'anno di nascita. . . . .	49
3.4	Istogramma della variazione % di solitudine in base alla presenza del partner. . . . .	52
3.5	Istogramma della variazione % di solitudine in base al numero di figli. . . . .	54
3.6	Tabella trivariata - Variazione % della solitudine in base al numero dei figli, senza la presenza del partner. . . . .	56
3.7	Tabella trivariata - Variazione % della solitudine in base al numero dei figli e con la presenza di un partner. . . .	56
3.8	Istogramma della variazione % di solitudine in base al numero di figli e alla presenza del partner ("Si"=presenza del partner, "No"= assenza del partner). . . . .	56
3.9	Istogramma della variazione % di solitudine in base al titolo di studio. . . . .	59
3.10	Istogramma della variazione % di solitudine in base al giudizio sulla solvibilità. . . . .	61

3.11	Tabella trivariata - Variazione % della solitudine in base ad un livello di istruzione basso e alla valutazione sulla solvibilità. . . . .	62
3.12	Tabella trivariata - Variazione % della solitudine in base ad un livello di istruzione medio e alla valutazione sulla solvibilità. . . . .	62
3.13	Tabella trivariata - Variazione % della solitudine in base ad un livello di istruzione alto e alla valutazione sulla solvibilità. . . . .	62
3.14	Istogramma della variazione % di solitudine in base al titolo di studio e alla valutazione sulla solvibilità. . . .	63
4.1	Output del modello di regressione lineare multipla. . . .	68
4.2	Output del modello di regressione lineare multipla con selezione delle variabili utilizzando il criterio dello step AIC. . . . .	75

# Elenco delle tabelle

1.1	Tabella sulla differenza tra <i>Solitude</i> e <i>Loneliness</i> nella letteratura inglese. . . . .	10
2.1	Tabella per codici ISO dei Paesi europei più Israele. . . . .	22
2.2	Tabella per le codifiche sul titolo di studio. . . . .	23
2.3	Età per ogni anno di nascita presente. . . . .	24
2.4	Tabella sulle percentuali delle modalità di titolo di studio, per Paese. . . . .	34
3.1	Tabella delle modalità della variabile risposta Y. . . . .	43
3.2	Tabella di frequenza relativa - Variazione % della solitudine per sesso. . . . .	45
3.3	Tabella sulla media della Y per ogni Paese. . . . .	47
3.4	Tabella di frequenza relativa - Variazione % della solitudine per anno di nascita. . . . .	49
3.5	Tabella di frequenza relativa - Variazione % della solitudine in base alla presenza del partner. . . . .	52
3.6	Tabella di frequenza relativa - Variazione % della solitudine in base al numero di figli. . . . .	54



3.7	Tabella di frequenza relativa - Variazione % della solitudine in base al titolo di studio. . . . .	59
3.8	Tabella di frequenza relativa - Variazione % della solitudine in base alla valutazione della solvibilità . . . . .	60
4.1	Variabili esplicative e relative modalità presenti nel <i>Dataset</i> .	67
4.2	Alcune statistiche del modello di regressione lineare multipla per la $Y_1$ e $Y_2$ . . . . .	77
4.3	Tabella delle stime dei coefficienti e p-value della regressione lineare multipla della $Y_1$ e $Y_2$ . . . . .	80

# Introduzione

La solitudine può essere definita come “lo stato spiacevole di riconoscere una discrepanza tra i legami sociali desiderati e quelli realmente raggiunti” (Perlman *et al.*, 1984) ed è inoltre caratterizzata da una percepita mancanza di controllo sulla quantità e qualità della propria attività sociale.

Lo scopo di questa relazione finale è studiare la variazione di solitudine in un campione rappresentativo di anziani di età superiore ai 60 anni, pensionati, prima e durante la prima ondata di COVID-19, mettendo in luce alcuni fattori che possono aver influenzato tale sensazione, tra cui la presenza o meno di partner e/o figli, il genere, la condizione economica percepita, e capire quanto effettivamente e quantitativamente la pandemia abbia peggiorato la situazione di chi già soffriva di solitudine. L'analisi verrà sviluppata a livello europeo, per poter analizzare in maniera sintetica l'eventuale differenza tra i vari Paesi.

Essendo un argomento che potrebbe ampliarsi su molti altri campi, questo testo focalizzerà l'attenzione sulla solitudine e le variabili ad essa strettamente connesse, senza estendere l'argomento sull'approfondimento delle conseguenze naturali, quali la depressione e l'ansia.

I dati utilizzati provengono da SHARE (Survey of Health, Ageing and Retirement in Europe), un'infrastruttura di ricerca che studia gli effetti delle politiche sanitarie, sociali, economiche e ambientali nel corso della vita dei cittadini europei e oltre.

Dal 2004 ad oggi sono state condotte oltre 380.000 interviste approfondite ad oltre 140.000 persone di età pari o superiore a 50 anni provenienti da 28 Paesi europei e Israele. Pertanto, SHARE è il più grande studio paneuropeo di esperti di scienze sociali che fornisce microdati longitudinali comparabili a livello internazionale, grazie ai quali è possibile realizzare approfondimenti nei settori della salute pubblica e delle condizioni di vita socioeconomiche degli individui europei, sia per gli scienziati che per i responsabili politici.

SHARE è diventato sempre più importante come strumento per l'elaborazione di politiche basate su dati concreti a causa dello scoppio della pandemia di COVID-19 ed è la banca dati ideale per studiare le conseguenze socio-economiche e sanitarie non intenzionali delle decisioni di contenimento epidemiologico e gli effetti a lungo termine del virus, grazie al suo corso di vita e all'approccio multidisciplinare che combina salute e dati socio-economici.

Verranno inizialmente proposte delle analisi descrittive del campione e del fenomeno oggetto d'indagine, per sintetizzare alcune informazioni contenute nei *dataset* ed evidenziare le relazioni che sussistono tra le variabili prese in considerazione. Saranno quindi stimati anche dei modelli di regressione multipla, per sintetizzare queste relazioni tra le variabili. Tutte le elaborazioni numeriche sono state eseguite con il software R.

# Capitolo 1

## La solitudine e i suoi significati

La solitudine è un'esperienza soggettiva spiacevole che inizia quando la rete sociale di un individuo subisce una perdita qualitativa o quantitativa e fa sì che le persone si sentano vuote, sole e indesiderate.

Essa è diversa dall'essere soli o dal vivere da soli, poichè una persona può soffrire di questa sensazione anche in presenza di altre persone, o vivere da sola senza sentirsi sola. Infatti, la letteratura inglese suddivide due termini, *solitude* e *loneliness*, attribuendogli significati diversi che generano esperienze opposte. Entrambi fanno riferimento allo stato di solitudine, ma la *loneliness* è uno stato mentale di mancanza, di insoddisfazione, segnato da un sentimento di alienazione dal mondo, una consapevolezza di eccessiva solitudine, una sensazione tumultuosa e per lo più deprimente. Spesso viene percepita come una separazione involontaria, un rifiuto o un abbandono da parte di altre persone.

La *solitude*, invece, viene considerata come un tempo di riflessione, una ricerca interiore, qualcosa che si può volontariamente scegliere di fare. Le persone a cui piace passare del tempo da sole continuano a mantenere relazioni sociali positive, bilanciando periodi di interazione sociale con

quelli di solitudine.

I significati dei due termini, spesso confusi o utilizzati come fossero la stessa cosa, possono essere riassunti in Tabella 1.1<sup>1</sup>:

	<i>Solitude</i>	<i>Loneliness</i>
<b>Definizione</b>	Stato di solitudine pacifica e piacevole.	L'infelicità o la depressione che si prova quando non si ha compagnia o amici.
<b>Stato Mentale</b>	Positivo.	Negativo.
<b>"Crowd"</b>	Essere fisicamente isolati è una condizione necessaria per sentirsi soli.	Ci si può sentire soli anche in mezzo alla folla.
<b>Risultati</b>	Emozioni positive, come la pace interiore e la tranquillità.	Può dare origine a tristezza, depressione e avere effetti negativi sulla salute.

**Tabella 1.1:** Tabella sulla differenza tra *Solitude* e *Loneliness* nella letteratura inglese.

In questa relazione verrà preso in considerazione solamente uno di questi due stati, cioè la *loneliness*, poiché essa è considerata come un grave problema di salute pubblica essendo associata ad un aumento della morbilità e mortalità; elementi essenziali in un periodo storico come quello preso in considerazione.

Gli individui che sono spesso soli, bramano il contatto umano, ma il loro stato d'animo rende più difficile formare connessioni con altre persone. Le persone sviluppano uno standard interno di paragone, idealizzando e giudicando le loro relazioni nella loro vita reale ed esterna. Se le relazioni che hanno instaurato riescono a sorpassare questo standard, l'individuo non sperimenterà sentimenti di solitudine e sarà realmente soddisfatto.

<sup>1</sup>Nella terza riga della tabella è presente il termine "crowd", che può essere inteso come una massa anonima di persone, che non si conoscono e che non hanno alcun tipo di interazione o relazione significativa tra di loro. In altre parole, può rappresentare la mancanza di connessione e di relazione con gli altri, nonostante la presenza fisica di molte persone intorno.

to delle proprie relazioni, qualitativamente e quantitativamente. Altrimenti, nel caso in cui le sue relazioni non soddisfacessero queste sue aspettative, l'individuo sperimenterà insoddisfazione e, di conseguenza, la solitudine, con intensità diversa a seconda dei casi. Le aspettative sociali negative tendono a suscitare comportamenti da parte degli altri che confermano le aspettative delle persone sole, mettendo così in moto una "profezia" che si autoavvera in cui le persone sole si allontanano attivamente dalle potenziali parti sociali anche se credono che la causa della distanza sociale sia attribuibile agli altri e sia al di fuori del loro controllo. Questo ciclo di solitudine auto-rinforzante è accompagnato da sentimenti di ostilità, stress, pessimismo, ansia e bassa autostima. Infatti, se lasciata incurata, essa ha gravi conseguenze per la cognizione, le emozioni, il comportamento e la salute, soprattutto nelle persone più vulnerabili e ad alto rischio, in particolare quelle socialmente, psicologicamente ed economicamente svantaggiate. Weiss (1973), inoltre, suddivide la solitudine in altre due categorie: emotiva e sociale.

La solitudine emotiva, condizione soggettiva, è correlata all'assenza di un legame stretto e affettuoso desiderabile con una persona, l'assenza di qualcuno a cui rivolgersi. La solitudine sociale, condizione oggettiva, comporta la mancanza di contatti, di reti sociali e il senso di appartenenza a una cerchia più piccola o più ampia di persone.

Perciò, riassumendo, essa è riconosciuta generalmente come una complessa eccitazione mentale che viene vissuta sotto forma di una spiacevole sensazione di stress, ed è considerata come un importante indicatore di benessere negli esseri umani. E' raramente osservata nelle persone che

hanno un'interazione sociale e le persone socialmente attive hanno una salute generale migliore rispetto a quegli individui che non interagiscono socialmente con altri.

## **1.1 Solitudine e anziani: l'impatto del COVID-19**

Mentre la solitudine sembra essere in aumento tra la popolazione generale a livello globale, gli anziani potrebbero essere i soggetti a maggior rischio, le cui conseguenze potrebbero portare a problemi di salute fisica e mortalità.

Molti fattori potrebbero contribuire alla solitudine tra gli anziani, come ad esempio la perdita del coniuge, il declino della salute fisica e mentale, la riduzione delle relazioni sociali e l'ospedalizzazione.

Nella primavera del 2020, lo scoppio dell'epidemia di COVID-19 ha colpito tutti i Paesi europei. Essa ha avuto effetti diretti sulla salute delle persone ma anche effetti indiretti, a volte discussi in termini di danni collaterali, inducendo preoccupazione, paura, ansia e sintomi depressivi.

Una delle principali misure di contenimento del virus sono state le limitazioni dei contatti e il distanziamento sociale. Quest'ultimo è stato associato a visite domiciliari interrotte, interruzioni della fornitura di assistenza regolare e concentrazione solo sui bisogni di base, imponendo un grande onere psicologico in particolare agli anziani non autosufficienti che vivevano da soli e/o che ricevevano assistenza domiciliare da parte di familiari, amici, badanti o servizi sociali.

Gli anziani rientravano nella fascia dei "soggetti fragili", la cui definizione fa riferimento a quel soggetto che di fronte a un evento di stress, quale

è la pandemia da Covid-19, non è in grado di rispondere in maniera adeguata e quindi soccombe, con un aumentato rischio di eventi negativi: mortalità, disabilità e aggravamento delle proprie condizioni generali.

Qualsiasi contatto con le persone anziane, perciò, è stato scoraggiato per poter proteggere da gravi progressioni della malattia questo gruppo vulnerabile. La loro vulnerabilità era legata allo stato fisico compromesso legato all'età, all'aumento della prevalenza di condizioni di salute croniche e altre disabilità, al declino delle capacità cognitive e alla potenziale presenza di condizioni psicosociali avverse.

Tali misure, che hanno limitato gli individui nelle attività sociali e nei contatti, hanno sicuramente portato a effetti positivi contro il virus, ma vengono associate anche a importanti ripercussioni a livello psicologico. La crisi del COVID-19 ha presentato molte sfide per la gestione dei sentimenti di solitudine. Gli studi sulla quarantena hanno scoperto che gli individui lottano per adattarsi a uno stile di vita incongruente con la natura sociale degli esseri umani e riportano una serie di reazioni psicologiche negative. Infatti, la solitudine nel contesto pandemico è stata associata ad un aumento della depressione, dell'ansia e del suicidio. Essere costretti a rimanere in casa, oltre al declino cognitivo, ha portato anche a conseguenze serie sul sistema cardiovascolare a causa della mancanza di movimento, determinando, quindi, peggioramenti a livello fisico dell'individuo.

Altri fattori di rischio per la solitudine includono circostanze e caratteristiche personali, salute e transizioni di vita (inteso come un periodo in cui si sperimenta un cambiamento significativo dello stile di vita).



Per esempio, in alcuni studi, un rischio più elevato di solitudine da lockdown è stato maggiormente riportato da alcune categorie:

- Sesso: donne
- Legami familiari: gli individui senza partner e/o figli
- Stato economico: gli individui con meno risorse economiche
- Uso della tecnologia: gli individui digitalmente esclusi

Oltre le precedenti categorie, le perdite personali, le preoccupazioni per la pandemia e un calo della fiducia nelle istituzioni sociali sono stati associati a crescenti problemi di salute mentale.

Tenendo conto del fatto che gli anziani costituiscono una percentuale significativa della popolazione <sup>2</sup>, essi possono continuare a mantenere un ruolo attivo nella società ed essere una popolazione significativa nel campo della ricerca.

## **1.2 Il ruolo della tecnologia durante la crisi COVID-19: il divario digitale**

Un ruolo fondamentale in questo sfondo è stato svolto dalla tecnologia digitale, che ha fornito opportunità virtuali non solo per le imprese e la fornitura di assistenza sanitaria, ma anche per la connessione sociale di persone lontane e il *networking*. Utilizzata per fornire supporto sociale e cognitivo, consentire l'acquisto e la vendita, facilitare il tempo libero e gli hobby, ha sicuramente migliorato lo stile di vita e l'umore durante la

---

<sup>2</sup>Secondo le stime dell'ultimo rapporto dell'Organizzazione Mondiale della Sanità (OMS), la popolazione anziana è in costante aumento: entro il 2050, nel mondo, una persona su cinque avrà più di 60 anni.

pandemia, fornendo strumenti adeguati per mitigare il rischio di un'epidemia di solitudine e mantenere le connessioni nonostante la necessità di rimanere fisicamente separati.

Per affrontare l'isolamento sociale e la solitudine durante la pandemia di COVID-19, l'Organizzazione Mondiale della Sanità ha raccomandato di mantenere i *social network* e rimanere in contatto con familiari, amici, colleghi e membri della comunità tramite mezzi digitali ((World Health Organization, 2020). La tecnologia digitale è diventata in quel periodo vitale, per riuscire ad affrontare la solitudine perché altri mezzi utilizzati normalmente (come la prescrizione sociale, la quale offre la possibilità di migliorare il benessere, soprattutto negli anziani, riducendo la solitudine e indirizzando i clienti bisognosi verso un coordinatore, che nel dialogo trova attività che l'individuo può svolgere all'interno della comunità locale) sono diventati difficili da accedere.

Tuttavia gli anziani, avendo un basso livello di conoscenza tecnologica, hanno sicuramente incontrato maggiori difficoltà e hanno avuto minor fiducia nell'utilizzo degli strumenti della tecnologia digitale online, andando incontro a una sorta di "esclusione digitale". Le percezioni di questi soggetti sull'accessibilità e la facilità d'uso di dispositivi, servizi e piattaforme tecnologiche hanno determinato considerevolmente la loro volontà di utilizzare la tecnologia per connettersi socialmente con gli altri.

Gli anziani hanno ripetutamente espresso il desiderio di istruzioni più chiare (rispetto a quelle tipicamente disponibili) su come utilizzare le applicazioni tecnologiche ed hanno anche descritto la necessità o l'uso di

servizi tecnologici di supporto che facilitassero il loro uso. Inoltre, hanno riscontrato spesso ostacoli sia sociali che materiali, ad esempio gli individui che vivevano in comunità rurali o isolate hanno sperimentato maggiori barriere. Ma son stati rilevati anche facilitatori, che includevano le conoscenze personali, sostegno da parte di familiari e amici e le pressioni sociali imposte dalla pandemia.

L'assenza di normali attività sociali ha colpito gli anziani in molti modi, oltre alla solitudine causata dalla mancanza di interazione sociale. Le interruzioni della routine possono aver causato una perdita di struttura generale e aver influito sui programmi di farmaci, sulla cura di sé, sull'igiene e sul sonno, che esacerbando ulteriormente la solitudine e altri problemi di salute. Sia l'interazione sociale che il ripristino di alcune routine sono stati assistiti dalla tecnologia, utilizzando app di videoconferenza (come FaceTime o Zoom), al posto delle visite di persona ad amici e familiari.

Sono molte le barriere psicologiche legate all'adozione della tecnologia anche se ora viene vista come un metodo che ha aiutato a rispettare le misure di distanziamento fisico, mantenendo i legami sociali e comunitari.

## Capitolo 2

### Descrizione del *Dataset*

Questi dati sono stati raccolti tramite interviste personali con l'utilizzo di un personal computer che gestiva il questionario elettronico (CAPI) nel pre-COVID e tramite interviste telefoniche assistite anch'esse da computer (CATI) durante il periodo COVID, per tutelare la salute di intervistati e intervistatori.

I dati presi in esame per la redazione di questa relazione sono:

- Quelli relativi alla Wave 8 (pre pandemia) le cui interviste sono iniziate a ottobre 2019 e sono state sospese a causa della pandemia a marzo 2020.
- Quelli relativi alla Wave 8 COVID-19 Survey 1 (prima ondata di pandemia), un modulo aggiuntivo speciale, raccolto tra giugno e settembre 2020.

Ai fini di questa analisi, verranno utilizzati entrambi i *dataset* per riuscire ad evidenziare il cambiamento del sentimento di solitudine nel campione di anziani prima e durante la prima ondata di COVID-19.

Il primo *dataset* della WAVE 8, riguardante il periodo precedente la pandemia, risulta molto più denso di informazioni rispetto a quello della

prima ondata, questo perchè è stato sviluppato un questionario specifico che copre gli stessi argomenti del normale questionario SHARE, ma notevolmente abbreviato e mirato alla situazione di vita COVID-19 delle persone di età pari o superiore a 50 anni.

I questionari coprono gli aspetti della vita più importanti della popolazione target e pongono domande specifiche riguardanti la salute fisica, mentale, l'uso del tempo, i social network, la situazione lavorativa, e molte altre informazioni, alcune delle quali non verranno incluse in questa relazione, poichè non interessanti ai fini dell'analisi.

Nel primo *dataset* "Wave 8", sono stati selezionati 7 moduli:

- Dati demografici (DN): informazioni demografiche di base sul genere, età, stato civile, il Paese di nascita e l'istruzione. Utili per selezionare il campione di anziani, e mettere in luce le principali differenze in base al genere e alla presenza o meno di un partner.
- Salute fisica (PH): informazioni su diversi aspetti della salute delle persone
- Salute mentale (MH): informazioni sulla salute mentale ed emotiva degli intervistati. Il dataset più importante tra tutti, poichè contiene la domanda che è il punto di partenza di questa analisi ("Quanto spesso si sente solo/a?").
- Bambini (CH): raccoglie informazioni sui figli dell'intervistato.
- Occupazione e pensioni (EP): utile per selezionare il campione di pensionati.
- Consumi (CO) : breve modulo sulle spese domestiche, che contiene il livello di reddito autovalutato, utile a capire quanto potrebbe

influenzare sulla vita emotiva dell'individuo.

- Attività (AC): contiene domande su una serie di attività e informazioni su come le persone si sentono riguardo a vari aspetti della loro vita.

Nel secondo *dataset* “Wave 8 COVID-19 Survey 1”, sono stati raccolti dati concentrati in un unico *dataset* (CA), il quale riguarda i comportamenti e le condizioni delle persone durante la pandemia. Ai partecipanti è stato chiesto di indicare il loro stato di vita in relazione alla pandemia, per questo motivo i dati selezionati si riferiscono principalmente alla salute fisica ed emotiva e ai contatti avuti durante il periodo preso in esame.

Per poter procedere con le analisi e realizzare i confronti adeguati, sono stati uniti i diversi *dataset* selezionati inizialmente in un unico *dataset*, la cui variabile chiave è *mergeid*, che è un identificatore di persona univoco e immutabile per tutte le wave, utilizzata per combinare i vari moduli a livello individuale.

## 2.1 Il campione dell'indagine

La popolazione target SHARE è costituita da tutte le persone di età pari o superiore a 50 anni al momento del campionamento che hanno il loro domicilio regolare nel rispettivo Paese SHARE. Le persone escluse dai campioni di riferimento sono i ricoverati in ospedale, le persone in carcere, coloro che erano fuori dal Paese durante l'intero periodo d'indagine, chi era incapace di parlare la lingua o le lingue del Paese o chi si era trasferito ad un indirizzo sconosciuto.

Per lo strumento del SHARE Corona Survey, il campione comprendeva:

1. I membri del panel che non erano stati intervistati nella Wave 8 prima della sospensione del lavoro sul campo.
2. I membri del panel che erano già stati intervistati nella Wave 8.

Entrambi i gruppi di intervistati hanno ricevuto lo stesso questionario; l'unica differenza è che i membri del panel che erano stati intervistati non dovevano rispondere di nuovo alle domande sulla composizione del nucleo familiare.

Di conseguenza, il *dataset* relativo alla Wave 8 e quello del Covid risultavano avere due numerosità differenti, 54.817 e 46.733 individui, rispettivamente. Per questo motivo è stato necessario selezionare esclusivamente i soggetti che avevano risposto a entrambi i questionari, diminuendo in questo modo la numerosità congiunta a 35.470 unità.

Successivamente vengono presi in considerazione solo gli individui nati nel 1959, che nella rilevazione del 2020 avevano 60 anni, riducendo lo status occupazionale a quello dei soli pensionati.

La numerosità finale risulta essere, quindi, di 23.490 soggetti.

È utile ricordare che SHARE mette a disposizione dei campioni non distorti, ed è per questo che avere campioni rappresentativi riduce problematiche, ad esempio, facendo riferimento al genere degli intervistati:

1. Bias di selezione: risultano da un errore nella modalità di selezionare i soggetti, portando a stime non corrette. Ad esempio, se la proporzione di partecipanti di uno dei generi è significativamente più alta rispetto all'altro, ci potrebbe essere questo rischio. Ciò potrebbe derivare dal fatto che il reclutamento dei partecipanti non è

stato equilibrato o dal fatto che il campione non rappresenta accuratamente la popolazione di riferimento.

2. Generalizzabilità: se la maggior parte dei partecipanti è di uno dei generi, le conclusioni dell'analisi potrebbero non essere generalizzabili all'intera popolazione.
3. Potenza statistica: una bassa proporzione di partecipanti di un determinato genere potrebbe limitare la potenza statistica. Ciò potrebbe rendere difficile rilevare differenze significative tra i due gruppi.

Con i dati a disposizione, nessuna di queste conseguenze verranno successivamente investigate.

### 2.1.1 Data Wrangling and Cleaning

È stata posta in essere una pulizia approfondita dei dati, eliminando eventuali valori NA (*Not Available* = dati mancanti), la cui presenza potrebbe bloccare l'esecuzione di calcoli sulle variabili numeriche che li contengono. Inoltre, sono state eliminate anche variabili con un'alta percentuale di valori come -1,-2,-3,-4,-5,-7, -9 (o equivalentemente 99), che corrispondevano a valori non plausibili, non codificabili, non applicabili o a domande rifiutate o senza risposta. Ritenendo comunque necessario mantenere le variabili con basse percentuali di questi valori (per non perdere troppa informazione e di conseguenza creare analisi distorte), da quest'ultime sono state tolte solo le osservazioni aventi queste codifiche di non plausibilità. La numerosità finale è quindi pari a 14.769 anziani. Al fine di ottenere risultati e analisi descrittive più chiare, sono state



ricodificate alcune modalità delle variabili, come ad esempio la variabile "Sesso", prima codificata con i numeri 1 e 2 riferiti al genere, ora rinominata con i nomi corrispondenti: Maschi e Femmine.

Un'altra modifica è stata apportata alla variabile Paese, la quale era codificata anch'essa con una serie di numeri, ciascuno dei quali si riferiva a un determinato Paese europeo. La ricodifica è stata fatta utilizzando i codici ISO (*International Organization for Standardization*). Esso è un codice a due (o tre) lettere assegnato a ciascun Paese del mondo per identificarlo in modo univoco. Questo codice viene utilizzato in molte applicazioni, tra cui la classificazione di lingue, valute e giurisdizioni. Questi codici sono stati standardizzati dalla ISO attraverso la sua serie di norme ISO 3166.

All'interno della Tabella 2.1 vengono riportati i codici utilizzati e i Paesi a cui corrispondono, per rendere la lettura della relazione più comprensibile.

<b>CODICE ISO</b>	<b>STATO MEMBRO</b>	<b>CODICE ISO</b>	<b>STATO MEMBRO</b>
DE	GERMANIA	CZ	REPUBBLICA CECA
SE	SVEZIA	PL	POLONIA
NL	OLANDA	LU	LUSSEMBURGO
ES	SPAGNA	HU	UNGHERIA
IT	ITALIA	SI	SLOVENIA
FR	FRANCIA	EE	ESTONIA
DK	DANIMARCA	HR	CROAZIA
EL	GRECIA	LT	LITUANIA
CH	SVIZZERA	BG	BULGARIA
BE	BELGIO	CY	CIPRO
ISR	ISRAELE	FI	FINLANDIA
CZ	REPUBBLICA CECA	LV	LETTONIA
PL	POLONIA	MT	MALTA
LU	LUSSEMBURGO	RO	ROMANIA
HU	UNGHERIA	SK	REPUBBLICA SLOVACCA

**Tabella 2.1:** Tabella per codici ISO dei Paesi europei più Israele.

Per la variabile "Titolo di studio" invece è stato necessario proseguire

con degli accorpamenti, al fine di rendere più chiare le codifiche originarie. Infatti, SHARE utilizza la scala ESeC (European Socio-economic Classification), che è una classificazione standardizzata a livello europeo delle posizioni sociali basata sulle occupazioni e sul livello di istruzione. La scala ESeC comprende sette categorie di istruzione codificate dall'1 al 7, in cui 1 rappresenta il livello di istruzione più basso e 7 il livello più alto<sup>1</sup>. Questa classificazione è stata sviluppata per consentire una comparabilità tra i diversi Paesi europei, in modo da standardizzare la misurazione del livello di istruzione e delle posizioni sociali a livello transnazionale. Pertanto, le codifiche utilizzate da SHARE riflettono la scala ESeC e sono conformi alle raccomandazioni europee per la misurazione delle posizioni sociali e del livello di istruzione.

La Tabella 2.2 sottostante fa emergere le modifiche e gli accorpamenti apportati ai fini di questa analisi.

<b>CODIFICHE SHARE</b>	<b>SIGNIFICATO</b>	<b>CODIFICHE UTILIZZATE</b>
1-2-3	Livello di istruzione basso	"Basso"
4	Livello di istruzione medio	"Medio"
5-6-7	Livello di istruzione alto	"Alto"

**Tabella 2.2:** Tabella per le codifiche sul titolo di studio.

Altre modifiche relativamente rilevanti sono state apportate alle variabili:

- "Figli" = comprende il conteggio dei figli naturali, in affidato, adottati e figliastri, accorpate in 4 modalità (1,2,3,4+).

---

<sup>1</sup>

- Livello di istruzione basso: non si è completata la scuola dell'obbligo o si è completata solo la scuola dell'obbligo.
- Livello di istruzione medio: si è completata la scuola superiore, ma non si ha una laurea universitaria.
- Livello di istruzione alto: si ha una laurea universitaria o un titolo di studio equivalente.

- "Anno"= cioè l'anno di nascita, in cui sono stati creati degli intervalli proporzionali alla numerosità presente, per evitare che ci fossero particolari divari tra le classi e, quindi, rischiare analisi distorte. Le classi <sup>2</sup> (con le relative età<sup>3</sup>) sono state create come riassunto della Tabella 2.3:

Anno di nascita	Età
(1915,1935]	85-105
(1935,1940]	80-85
(1940,1945]	75-80
(1945,1950]	70-75
(1950,1955]	65-70
(1955,1960]	60-65

**Tabella 2.3:** Età per ogni anno di nascita presente.

---

<sup>2</sup>La dicitura: "(1915,1935]" significa che è escluso l'anno 1915 e incluso il 1935.

<sup>3</sup>L'età è stata calcolata prendendo come anno di riferimento il 2020, poichè la relazione si concentra sugli individui che nel 2020 avevano più di 60 anni.

## 2.2 Esplorazione preliminare dei dati: analisi sul campione

Nella presente sezione si esplorano i dati raccolti al momento dello scoppio della pandemia sul campione di anziani di età superiore ai 60 anni e pensionati ( $n=14.769$ )<sup>4</sup>, al fine di fornire una panoramica preliminare delle principali caratteristiche del campione.

In particolare, verranno analizzate le variabili più interessanti per l'oggetto di studio, tra cui l'età, il genere, il titolo di studio, il Paese di provenienza, la presenza del partner, e la presenza di figli. La presentazione dei risultati sarà accompagnata da grafici che permetteranno di visualizzare in maniera chiara e sintetica le distribuzioni e le tendenze presenti nei dati. Tale analisi descrittiva costituisce un passo preliminare fondamentale per una comprensione più approfondita dei fenomeni oggetto di studio e per l'individuazione delle eventuali relazioni tra le diverse variabili.

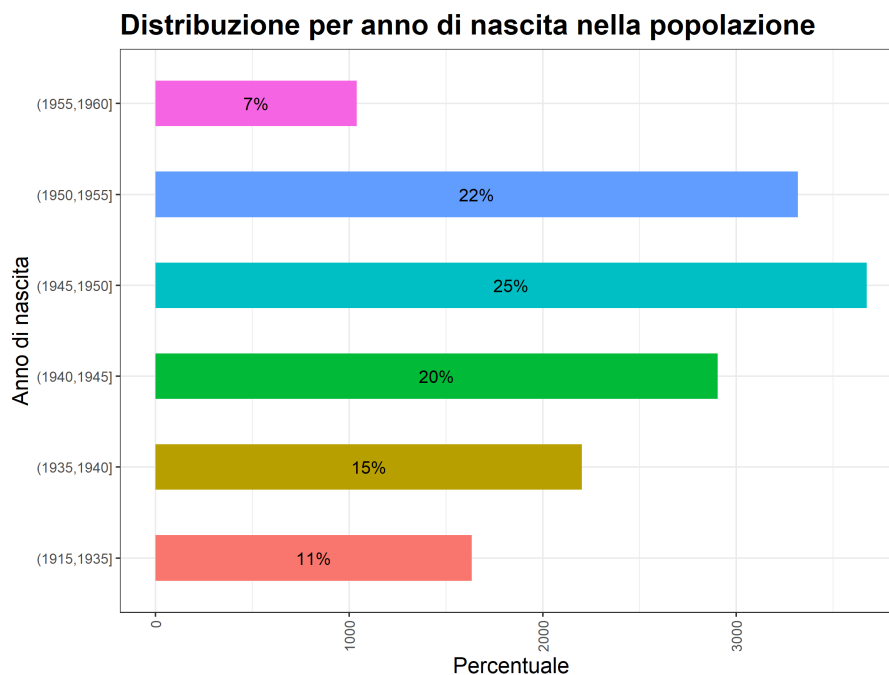
### 2.2.1 Analisi descrittiva per età

Per avere una visione più dettagliata delle caratteristiche del campione, è stata analizzata la distribuzione per anno di nascita dei partecipanti. Nel grafico a barre in Figura 2.1 è rappresentata la frequenza relativa di partecipanti per anno di nascita. L'analisi di questa variabile permetterà di comprendere meglio la distribuzione di età del campione, evidenziando le fasce più rappresentative e quelle meno frequenti.

---

<sup>4</sup>Nota: Il campione copre il periodo dal 1959 al 2020.

L'asse y rappresenta gli anni di nascita, mentre l'asse x rappresenta la percentuale di partecipanti per anno di nascita



**Figura 2.1:** Istogramma della distribuzione del campione per classe di anno di nascita.

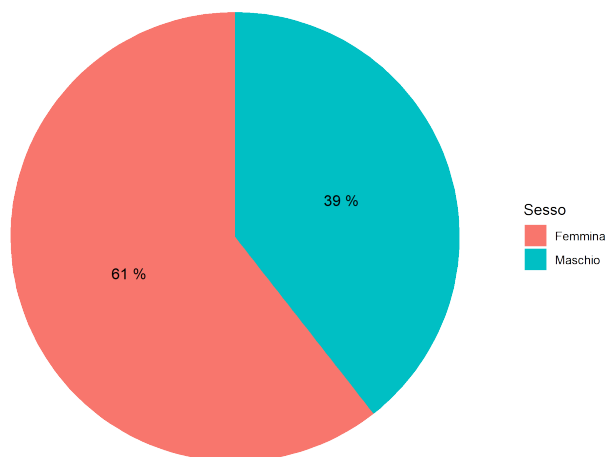
Come si può notare, la maggior parte dei partecipanti è nata tra il 1945 e il 1955, con una frequenza relativa del 24,87% e del 22,47%, rispettivamente. I partecipanti nati negli anni precedenti e successivi hanno frequenze relative più basse, con una netta diminuzione per quelli nati prima del 1935. Si può notare inoltre una riduzione rilevante della frequenza relativa per i partecipanti nati dopo il 1955, con una percentuale del 7% per quelli nati tra il 1960 e il 1965. Questi risultati potrebbero suggerire l'influenza di eventi storici o sociali nella determinazione del tasso di natalità in diversi periodi.<sup>5</sup>

<sup>5</sup>Per esempio, si può notare un'alta concentrazione di individui nati negli anni '40 e '50, corrispondenti alle generazioni del baby boom e della generazione X.

## 2.2.2 Analisi descrittiva per genere

Le informazioni sul genere dei partecipanti vengono rappresentate in un grafico a torta in Figura 2.2.

**Distribuzione di genere nella popolazione**



**Figura 2.2:** Grafico a torta che illustra la distribuzione del campione in base al sesso.

Come si può notare la dimensione del segmento del grafico relativo alle donne è più grande rispetto al segmento relativo agli uomini, indicando una forte maggioranza di partecipanti di sesso femminile.

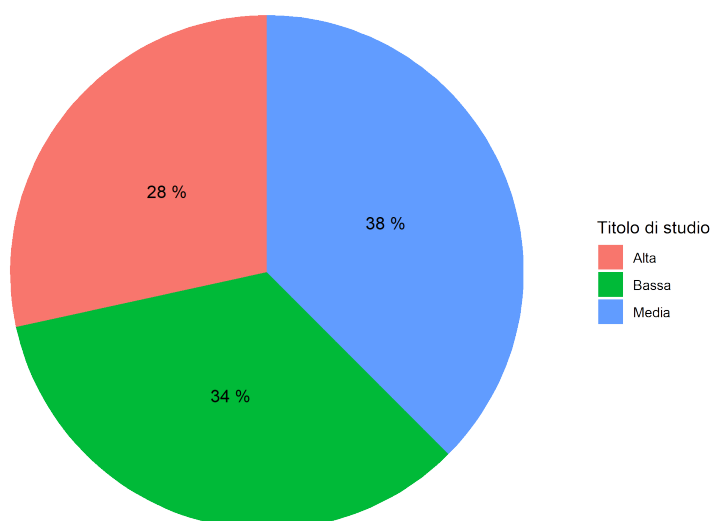
In generale, la differenza di proporzione tra i due gruppi rappresentati nel grafico a torta sarà essere presa in considerazione durante la progettazione dello studio, l'analisi dei dati e l'interpretazione dei risultati.

## 2.2.3 Analisi descrittiva per titolo di studio

Il grafico a torta in Figura 2.3 rappresenta la distribuzione della popolazione in base al titolo di studio, suddivisa in tre categorie: Alto, Medio e Basso. Il titolo di studio è un indicatore importante del livello di istruzione e formazione della popolazione, e può avere un impatto im-

portante sulle opportunità di lavoro e sullo status socio-economico. La visualizzazione della distribuzione della popolazione per titolo di studio aiuta a capire meglio il profilo educativo della popolazione e a identificare eventuali disuguaglianze educative.

### Distribuzione della popolazione per livello di istruzione



**Figura 2.3:** Grafico a torta: ripartizione della popolazione in base al livello di istruzione.

Come è possibile notare, la maggioranza degli anziani nel campione ha un livello di istruzione basso o medio. Ci possono essere diversi fattori che contribuiscono a questo fenomeno: per esempio, in passato, in alcune regioni o Paesi, l'accesso all'istruzione era limitato, soprattutto per le fasce sociali più povere. Inoltre, alcune persone potrebbero aver abbandonato la scuola in giovane età per dedicarsi al lavoro o per motivi familiari. Altri fattori possono includere la mancanza di risorse economiche o di supporto familiare per continuare gli studi. Questi gruppi demografici, quindi, come conseguenza, potrebbero essere particolarmente vulnerabili alla solitudine e all'isolamento durante il periodo

di pandemia causato dal COVID-19, poiché l'accesso a lavori di alta qualità e il supporto educativo possono influire sul benessere mentale e sulla connessione sociale. Ciò comporterebbe un rischio anche a causa di un accesso limitato alle tecnologie e alle risorse online, che potrebbero essere utilizzate per mantenere i contatti con gli amici e la famiglia.

### **Analisi della relazione tra titolo di studio, lavoro e percezione sulla capacità di far quadrare i conti**

Il livello di istruzione è spesso associato al reddito, all'occupazione e alla salute, e rappresenta un importante fattore predittivo del benessere individuale e collettivo. In questo lavoro è stata analizzata la relazione tra il titolo di studio e la percezione riguardo la capacità di far quadrare i conti. In questo caso è stata aggiunta una variabile ancora non approfondita in questa relazione, che rappresenta la domanda nel questionario "Pensando al reddito mensile complessivo della sua famiglia, direbbe che la sua famiglia riesce a far quadrare i conti...", con le seguenti modalità (e relative codifiche poi presenti nel grafico che segue):

1. Con grande difficoltà
2. Con qualche difficoltà
3. Abbastanza facilmente
4. Facilmente

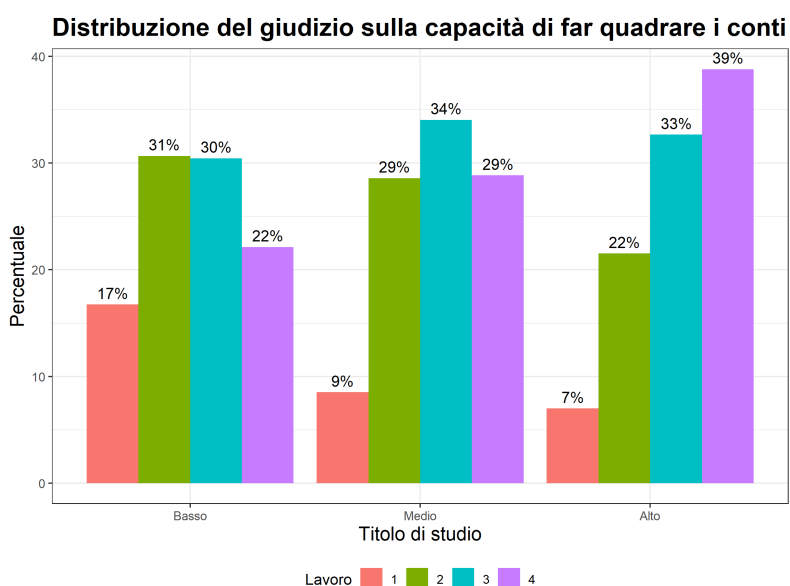
Questa variabile economica presa in considerazione riflette la valutazione soggettiva della capacità di bilanciare il proprio budget e rappresenta un aspetto importante nella valutazione della capacità finanziaria dei



rispondenti, poiché riflette la loro consapevolezza e competenza nell'affrontare le questioni legate alle finanze personali.

I dati mostrano la distribuzione percentuale delle risposte affermative alla domanda sulla capacità di far quadrare i conti per ciascuna categoria di titolo di studio (Figura 2.4). Il grafico mostra come, per ogni livello di istruzione, si distribuiscono le risposte in base alla percezione della difficoltà nel far quadrare i conti.

Le barre rappresentano la percentuale di rispondenti per ogni categoria di giudizio, sommata per ciascun livello di istruzione.



**Figura 2.4:** Istogramma sulla distribuzione del giudizio sulla capacità di far quadrare i conti in base al titolo di studio e al lavoro.

Si può notare come l'aumentare del livello di istruzione sia associato a una diminuzione della percentuale di persone che hanno una percezione negativa (modalità 1 e 2) sulla capacità di far quadrare i conti, e un aumento delle percentuali di coloro che hanno una percezione positiva (modalità 3 e 4).

In particolare, tra coloro che hanno un titolo di studio basso, il 48% ha

una percezione relativamente negativa sulla capacità di far quadrare i conti, mentre solo il 22% ha una percezione del tutto positiva.

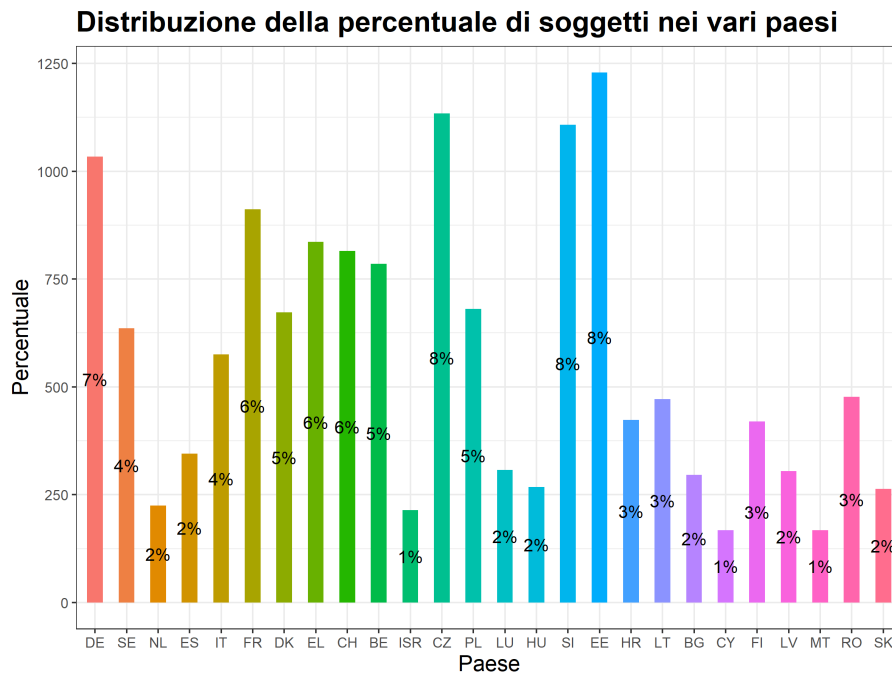
Tra coloro che hanno un titolo di studio medio, la percentuale di persone che hanno una percezione negativa si riduce al 38%, mentre il 63% ha una percezione positiva sulla capacità di far quadrare i conti.

Tra coloro che hanno un titolo di studio alto, solo il 29% ha una percezione negativa, mentre il 72% ha una percezione positiva sulla capacità di far quadrare i conti.

Questi risultati suggeriscono che il livello di istruzione è un fattore importante nella percezione della difficoltà nel far quadrare i conti, e che un titolo di studio più elevato potrebbe essere associato a una maggiore sicurezza finanziaria e ad una migliore comprensione delle finanze personali.

#### **2.2.4 Analisi descrittiva per Paese**

Il grafico a barre presentato in Figura 2.5 rappresenta la distribuzione della popolazione in base al Paese di provenienza. Sono stati analizzati i dati di un campione di individui provenienti da 26 Paesi diversi. Ogni barra rappresenta la percentuale di individui provenienti da un determinato Paese. Il grafico permette di osservare la distribuzione geografica della popolazione analizzata.



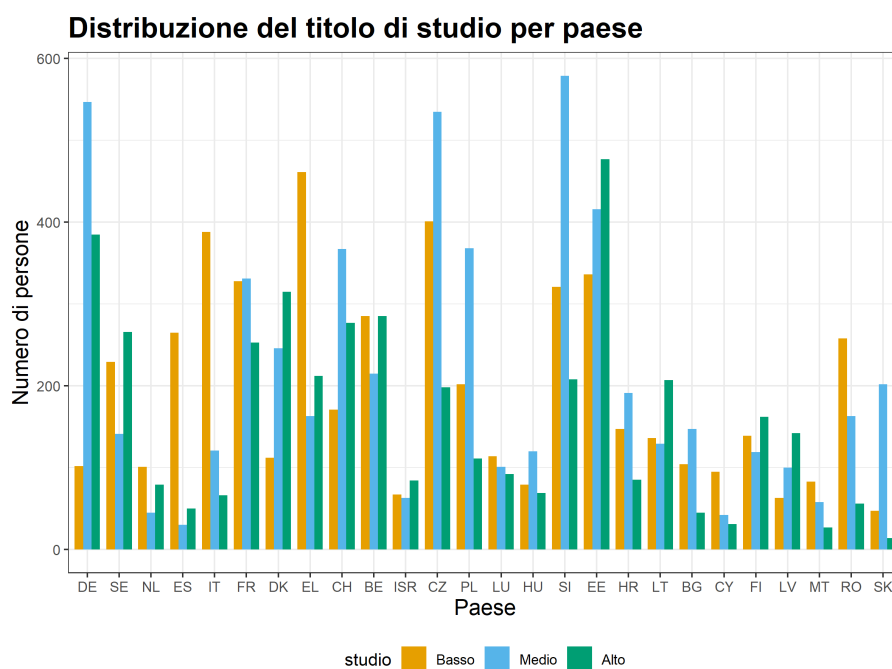
**Figura 2.5:** Istogramma sulla distribuzione della popolazione in base al Paese di provenienza.

L'istogramma evidenzia che la distribuzione della popolazione varia, in minima parte, tra i Paesi presi in considerazione. Ad esempio, si può notare che i Paesi con la maggiore percentuale di soggetti rappresentati nel campione sono: Germania (DE), Repubblica Ceca (CZ), Slovenia (SL) ed Estonia (EE) con una percentuale intorno al 7/8%. I Paesi con la percentuale minore sono Israele (ISR), Cipro (CY) e Malta (MT) con l'1%. In generale, si può notare che non c'è una variazione esageratamente rilevante della distribuzione della popolazione tra i Paesi considerati.

### **Confronto internazionale della formazione degli anziani pensionati: differenze e analogie**

Per capire meglio la situazione del campione di anziani pensionati, è stata esaminata la distribuzione del titolo di studio per Paese. Il grafico di

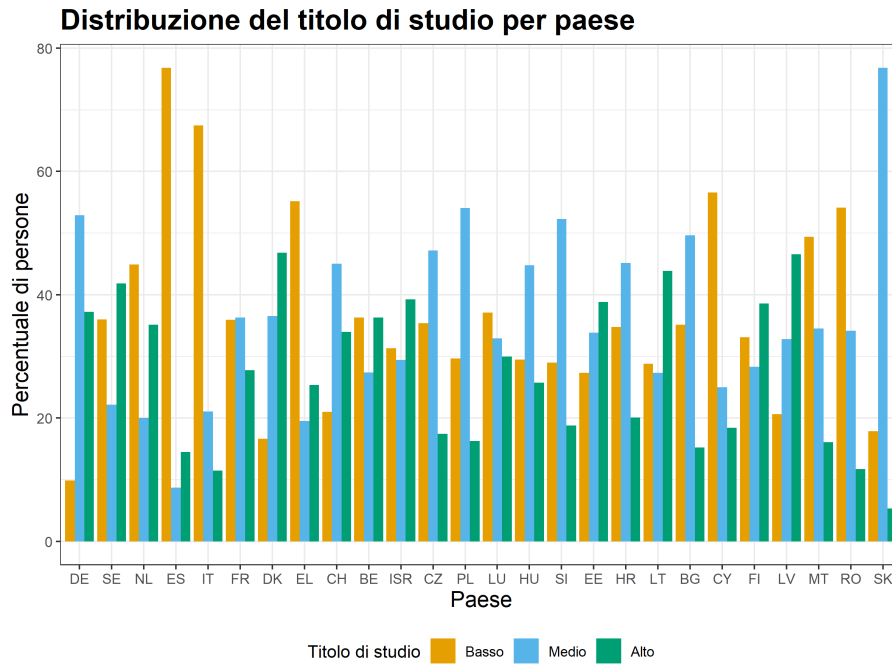
Figura 2.6 mostra il conteggio di persone con titoli di studio bassi, medi e alti in ciascuno dei 26 Paesi considerati ed intende evidenziare come la distribuzione del livello di istruzione possa variare tra i Paesi considerati, offrendo una panoramica diacronica e trasversale sulle differenze educative e socio-economiche nella popolazione anziana in Europa.



**Figura 2.6:** Istogramma sulla distribuzione del titolo di studio della popolazione, per Paese.

Questo istogramma restituisce una rappresentazione visiva della distribuzione dei titoli di studio per ogni Paese presente nei dati. Ogni barra rappresenta un Paese e le diverse sezioni colorate all'interno di ciascuna barra indicano, appunto, la proporzione di persone che hanno conseguito un determinato titolo di studio in quel Paese.

Invece, il grafico di Figura 2.7, (e la relativa Tabella 2.4 che riassume i dati riportati nell'istogramma, al fine di fornire una visione più specifica dell'analisi) presenta la distribuzione percentuale del titolo di studio di un campione di individui divisa per Paese di provenienza.



**Figura 2.7:** Istogramma sulla distribuzione percentuale del titolo di studio per Paese.

Paese	Basso	Medio	Alto	Paese	Basso	Medio	Alto
DE	9.86	52.90	37.26	LU	37.13	32.90	29.97
SE	36.01	22.17	41.82	HU	29.48	44.78	25.75
NL	44.89	20.00	35.11	SI	28.97	52.26	18.77
ES	76.81	8.70	14.49	EE	27.34	33.85	38.81
IT	67.48	21.04	11.48	HR	34.75	45.15	20.09
FR	35.96	36.29	27.74	LT	28.81	27.33	43.86
DK	16.64	36.55	46.81	BG	35.14	49.66	15.20
EL	55.14	19.50	25.36	CY	56.55	25.00	18.45
CH	20.98	45.03	33.99	FI	33.10	28.33	38.57
BE	36.31	27.39	36.31	LV	20.66	32.79	46.56
ISR	31.31	29.44	39.25	MT	49.40	34.52	16.07
CZ	35.36	47.18	17.46	RO	54.09	34.17	11.74
PL	29.66	54.04	16.30	SK	17.87	76.81	5.32

**Tabella 2.4:** Tabella sulle percentuali delle modalità di titolo di studio, per Paese.

I valori evidenziati in:

- Rosso: indicano i Paesi con le percentuali più *alte* di un *basso* di livello di istruzione.
- Verde: indicano i Paesi con le percentuali più *alte* di un *alto* livello di istruzione.

- Arancione: indicano i Paesi con le percentuali più *basse* di un *alto* livello di istruzione.

Si può notare che i pensionati anziani in Spagna e in Italia hanno una percentuale molto alta di *bassa istruzione* (rispettivamente il 76.87% e il 67.48%), mentre in Svezia, Danimarca, Lituania e in Lettonia la maggioranza dei pensionati anziani ha un *alto livello di istruzione* (tutti e quattro con percentuali maggiori del 40%). Inoltre, la percentuale di anziani pensionati con *istruzione media* è più alta in Repubblica Slovacca (76.81%).

In generale, viene evidenziato come il livello di istruzione dei pensionati anziani possa variare notevolmente tra i Paesi, con alcune nazioni in cui la maggioranza ha un alto livello di istruzione e altre in cui la maggioranza ha un basso livello di istruzione. Tuttavia, è importante sottolineare che questi risultati possono essere influenzati da vari fattori, come la qualità del sistema educativo di un Paese e le politiche sociali e previdenziali. Sicuramente le coorti più vecchie di 60 anni potrebbero aver avuto meno opportunità di istruzione rispetto a quelle più giovani, a causa di fattori economici e sociali. Inoltre, la durata dell'istruzione obbligatoria può variare tra i Paesi e tra le diverse epoche. Questo potrebbe avere un impatto sulla distribuzione dei titoli di studio nel campione, a seconda del Paese di appartenenza e della fascia di età. Alcuni Paesi potrebbero aver avuto programmi più avanzati o incentivi all'istruzione, portando a una maggiore percentuale di anziani con un titolo di studio superiore, mentre altri potrebbero aver avuto sistemi educativi meno sviluppati, con una maggiore percentuale di anziani con

un titolo di studio inferiore. Risulta quindi importante tener conto delle differenze storiche e culturali nell'istruzione tra i diversi Paesi.

**Breve approfondimento: l'Italia e l'istruzione del campione di riferimento**

Ci possono essere molteplici ragioni storiche per cui in Italia, rispetto ad altri Paesi, ci siano tassi più alti di anziani con un'istruzione bassa.

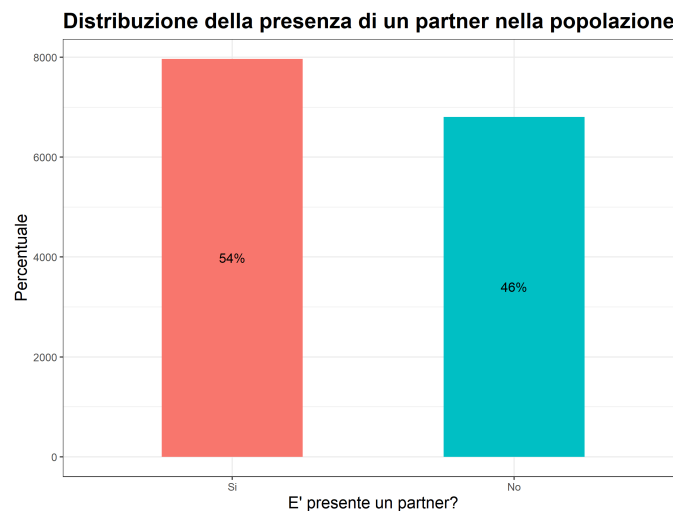
Alcuni fattori possono includere:

- Nel dopoguerra, l'Italia ha subito una forte emigrazione dalle aree rurali alle città, con molte persone che si sono spostate per trovare lavoro. Questo può aver reso difficile per molte famiglie garantire ai loro figli l'accesso all'istruzione, specialmente per quelli di origine povera.
- L'istruzione nel Paese è stata a lungo gestita dalla Chiesa cattolica, che ha spesso promosso un modello di istruzione basato sull'insegnamento religioso piuttosto che su quello secolare. Ciò può aver limitato l'accesso all'istruzione per molte persone, specialmente per coloro che non condividevano la fede cattolica.
- Negli anni '60 e '70, l'Italia ha subito un boom economico, che ha portato a una maggiore industrializzazione e urbanizzazione. Tuttavia, il sistema educativo non è stato in grado di tenere il passo con questi cambiamenti, il che può aver limitato l'accesso all'istruzione per molte persone.
- Infine, la disoccupazione giovanile in Italia è stata a lungo uno dei più alti in Europa. Ciò può aver reso difficile per molti giovani continuare i loro studi e migliorare il loro livello di istruzione, il che

potrebbe aver influenzato i livelli di istruzione dell'intera popolazione, compresi, quindi, gli anziani di oggi che entrano a far parte del nostro campione.

## 2.2.5 Analisi descrittiva per la presenza del partner

La Figura 2.8 illustra la distribuzione sul campione di riferimento. La variabile 'presenza di un partner' è stata codificata come una variabile binaria, con 'si' e 'no' come possibili categorie. Questo istogramma rappresenta la proporzione di soggetti per ogni categoria.



**Figura 2.8:** Istogramma sul confronto tra la presenza e l'assenza di partner nella popolazione.

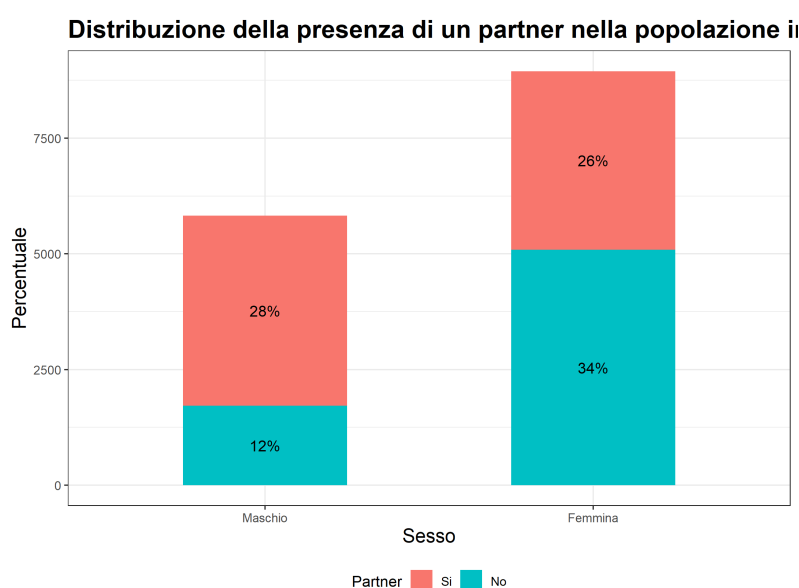
Dalla Figura emerge che la maggioranza del campione ha un partner (54%), anche se differenza tra i due gruppi non è molto marcata. Questi dati indicano comunque una buona percentuale di anziani che vivono soli, probabilmente a causa di vedovanza o di separazione/divorzio. È utile ricordare che la variabile "presenza di un partner" può influenzare diversi aspetti della vita individuale e sociale, tra cui la salute mentale, il benessere fisico e le relazioni interpersonali, e quindi potrebbe essere



un fattore importante nella valutazione della solitudine e del benessere individuale.

### Differenze di genere nella presenza di un partner

Dopo aver analizzato la presenza del partner nella popolazione di interesse, risulta interessante esaminare la distribuzione di questa variabile in base al genere dei soggetti. In Figura 2.9, si può osservare la distribuzione della variabile "Presenza del partner" in base al genere.



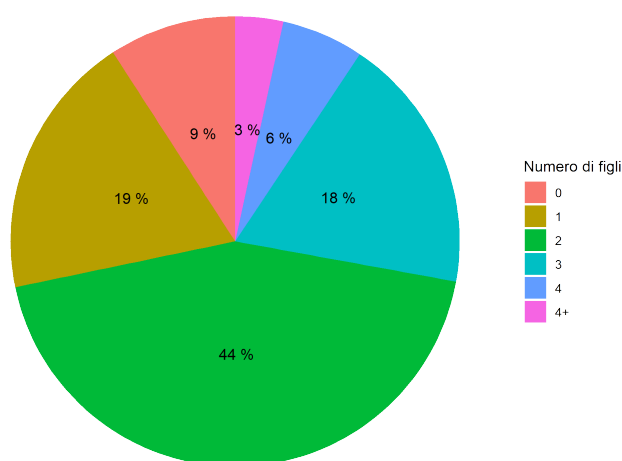
**Figura 2.9:** Istogramma per la distribuzione della presenza di un partner in base al sesso nel campione di studio.

Dai risultati emersi dal grafico, si può notare che la presenza di un partner varia tra i due generi, ma non in modo rilevante. Infatti, la percentuale di uomini con un partner (28%) è maggiore rispetto a quella delle donne (26%), mentre la percentuale di donne senza partner (34%) è più alta rispetto a quella degli uomini (12%).

## 2.2.6 Analisi descrittiva per numero di figli

Il numero di figli rappresenta un indicatore importante per valutare la demografia e la struttura della popolazione. Il grafico presentato in Figura 2.10 mostra la distribuzione percentuale del numero di figli nel campione, suddivisa in quattro categorie.

**Percentuale del numero di figli della popolazione**



**Figura 2.10:** Istogramma per la distribuzione percentuale del numero di figli nella popolazione, suddivisi in 4 categorie: 1 figlio, 2 figli, 3 figli, e 4 o più figli.

Osservando il grafico a torta della distribuzione dei figli nel campione, si nota che la maggior parte degli anziani ha avuto due figli (44%), seguiti da coloro che hanno avuto un solo figlio (19%), mentre una quota rilevante ha avuto tre figli (18%). Una quota minoritaria ha avuto quattro figli (6%) e solo il 3% ha avuto cinque o più figli.

### **Possibili cause alla base del numero di figli del campione**

Il fatto che la maggioranza degli anziani nel campione abbia avuto due figli può essere spiegato in diversi modi. Ad esempio, potrebbe riflettere la tendenza generale delle famiglie a ridurre le dimensioni, che ha avuto

inizio negli anni '60 e '70. Inoltre, l'aumento della qualità della vita e l'accesso a cure mediche migliori potrebbero aver contribuito a ridurre la mortalità infantile, portando quindi ad una maggiore sopravvivenza dei bambini e alla riduzione della necessità di avere un numero elevato di figli per garantire la continuità della famiglia. Inoltre, il fatto che una quota rilevante di anziani abbia avuto tre figli potrebbe essere spiegato dall'effetto di un trend demografico noto come *Baby Boom*, che ha avuto luogo negli anni '50 e '60.<sup>6</sup> Infine, il fatto che solo una quota molto ridotta di anziani abbia avuto quattro o più figli potrebbe essere indicativo di un cambiamento culturale che ha portato ad una riduzione della pressione sociale per avere figli numerosi.

---

<sup>6</sup>Questo trend ha portato ad un aumento rilevante del tasso di natalità in molti Paesi, compreso l'Italia, e potrebbe aver influenzato il numero di figli di molti anziani nel campione.

## Capitolo 3

# Analisi dei fattori associati alla solitudine degli anziani durante la pandemia di COVID-19

In questo capitolo, vengono presi in considerazione alcuni fattori associati alla solitudine degli anziani durante la pandemia di Covid-19.

### 3.1 Una misura di solitudine: introduzione alla variabile risposta e alle variabili esplicative

La variabile  $Y$  viene spesso chiamata "variabile risposta" o "variabile dipendente". Questo perché rappresenta la variabile di interesse che si vuole studiare, spiegare o predire.

Avendo a disposizione i dati di SHARE, è stato possibile costruire la variabile risposta e adattarla alle nostre esigenze. All'interno dei questionari relativi a entrambe le Wave (quella della pre-pandemia e quella del periodo relativo alla prima ondata), era presente una domanda che ha reso possibile il raggiungimento dell'obiettivo di questa relazione. Que-

sta è stata estrapolata dal modulo "Salute Mentale" (MH e CAMH), e si riferisce al sentimento di solitudine basato su una scala di misura che indica la frequenza con cui gli anziani si sentono soli. La domanda con le relative modalità (che sarà oggetto di tutte le analisi) è:

**Quanto spesso si sente solo/a?**

1. **Spesso**
2. **Qualche volta**
3. **Mai o quasi mai**

Come si può notare la scala è composta da tre livelli di risposta: "Spesso" indica che l'anziano prova frequentemente la sensazione di solitudine; "Qualche volta" indica che la solitudine è presente ma non è così frequente; "Mai o quasi mai" indica che l'anziano non si sente solitario in modo consistente. Questa scala di misura è utile per valutare l'esperienza soggettiva di solitudine degli anziani, che può influenzare la loro qualità di vita e il benessere psicologico. Inoltre, confrontando le risposte prima e durante la pandemia di Covid-19, si può valutare se ci siano state variazioni nella percezione di solitudine in relazione al cambiamento delle circostanze.

A tal fine, la  $Y$  è stata costruita in questo modo:

$$Y = SolitudineDurante^1 - SolitudinePrima^2 \quad (3.1)$$

---

<sup>1</sup>Solitudine Durante: si riferisce alla domanda "Quanto spesso si sente solo/a?" posta durante la prima ondata di Covid-19

<sup>2</sup>Solitudine Prima: si riferisce alla domanda "Quanto spesso si sente solo/a?" posta prima della pandemia

Le cui modalità sono rappresentate nella scala riportata in Tabella 3.1, con i relativi significati:

Modalità della Y	Significato
-2 e -1	La percezione di solitudine è peggiorata
0	La percezione di solitudine è rimasta invariata
1 e 2	La percezione di solitudine è migliorata

**Tabella 3.1:** Tabella delle modalità della variabile risposta Y.

In questo modo la variabile risposta rappresenterà la differenza nella percezione di solitudine degli anziani tra i due periodi distinti, permettendo di evidenziare il ruolo che ha avuto la pandemia. Si noti che, procedendo nel modo appena presentato, i valori della Y (ad esclusione del 2 e del -2) possono essere ottenuti in diversi modi:

il valore **0** può essere il risultato (della sottrazione) delle risposte:

- 1 (durante la pandemia) e 1 (pre-pandemia)
- 2 (durante la pandemia) e 2 (pre-pandemia)
- 0 (durante la pandemia) e 0 (pre-pandemia)

il valore **1** può essere il risultato (della sottrazione) delle risposte:

- 2 (durante la pandemia) e 1 (pre-pandemia)
- 1 (durante la pandemia) e 0 (pre-pandemia)

il valore **-1** può essere il risultato (della sottrazione) delle risposte:

- -2 (durante la pandemia) e 1 (pre-pandemia)
- -1 (durante la pandemia) e 0 (pre-pandemia)

Viene quindi ottenuta una misura di solitudine diversa rispetto a quella originariamente raccolta. Essa permette infatti di evidenziare differenze

nel periodo preso in esame indipendentemente dal livello di solitudine dichiarato in partenza. Infatti l'interesse è rivolto ai cambiamenti osservati dopo la prima ondata di Covid-19, per analizzare un possibile effetto dovuto alla pandemia stessa. Ad esempio, se un anziano si è sentito spesso solo sia prima che durante la pandemia, la sua percezione di solitudine, per quanto seria, non è peggiorata a causa del Covid-19. Sono state scelte minuziosamente anche diverse variabili esplicative che possono essere associate alla solitudine, alcune delle quali già discusse nel capitolo precedente. Quelle prese in considerazione al momento, sono: Paese, genere, anno di nascita, titolo di studio, presenza di partner, presenza e numero di figli, giudizio sulla percezione della situazione economica.

Tuttavia, all'interno del dataset sono presenti anche altre variabili che verranno invece introdotte nei modelli di regressione, stimati nel prossimo capitolo (come ad esempio la salute fisica, eventuali problemi di insonnia o alimentari e altre).

## **3.2 Analisi visiva della correlazione tra solitudine e fattori di contesto**

In questa sezione vengono presentati i risultati dell'analisi grafica della relazione tra la variazione di solitudine e le variabili esplicative rilevanti, nel campione di anziani pensionati.

Ogni analisi viene introdotta (come in quella precedente) da una breve descrizione che evidenzia le principali caratteristiche della relazione tra la variabile risposta e la variabile esplicativa in questione. Questo con-

sente di comprendere meglio la distribuzione dei valori della  $Y$  rispetto alle diverse categorie di ogni covariata e di individuare eventuali tendenze o relazioni rilevanti tra le variabili.

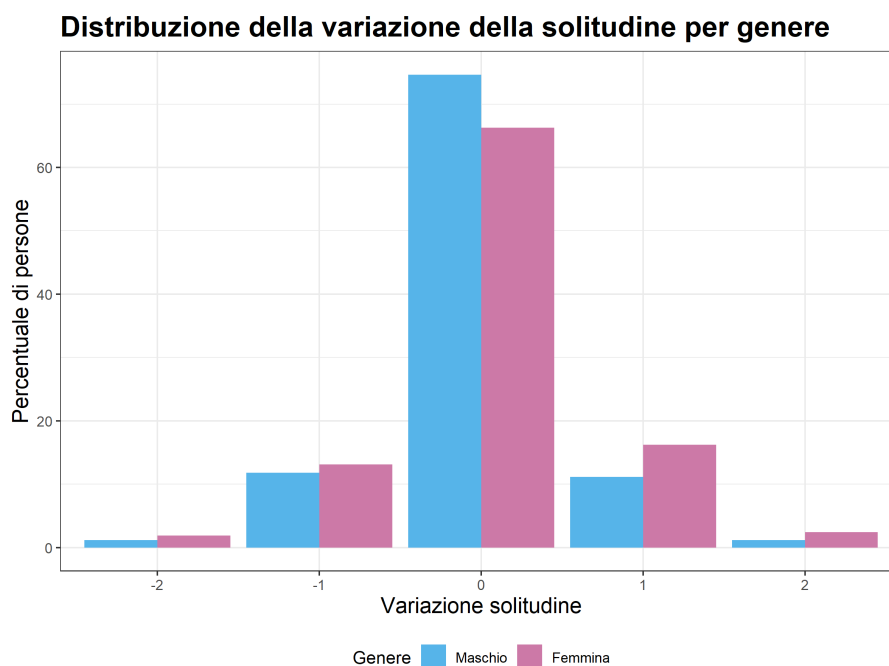
### 3.2.1 Solitudine e genere

Queste analisi descrittive possono rivelarsi utili per vedere se esistano differenze rilevanti nella solitudine tra donne e uomini anziani, e possono anche mostrare se siano determinati livelli di solitudine che siano più comuni in un genere rispetto all'altro.

La Tabella 3.2 e il grafico in Figura 3.1 mostrano la distribuzione della variazione della solitudine in base al genere.

	Y				
	-2	-1	0	1	2
Maschi	1.19	11.80	74.65	11.18	1.19
Femmine	1.91	13.14	66.29	16.22	2.44

**Tabella 3.2:** Tabella di frequenza relativa - Variazione % della solitudine per sesso.



**Figura 3.1:** Istogramma della variazione % di solitudine in base al sesso.



E' possibile notare come, per ogni valore della variabile  $Y$  (cioè -2, -1, 0, 1, 2), la tabella mostri la percentuale di maschi e femmine che hanno sperimentato quel livello di variazione nella solitudine durante lo studio. Ad esempio, il 13% dei maschi (contro il 15% delle donne) hanno sperimentato un peggioramento della loro solitudine durante il periodo preso in esame, mentre il 19% delle donne hanno riportato un miglioramento (contro il 12% maschile).

In generale, la tabella suggerisce che le donne sembrano avere maggiori variazioni (non particolarmente rilevanti) di solitudine rispetto agli uomini, poiché hanno percentuali un po' più elevate nelle modalità che riguardano il peggioramento ed il miglioramento (cioè i valori estremi). Dal grafico risulta molto evidente come la maggioranza dei partecipanti abbia riportato di non aver avvertito variazioni nella propria solitudine (per entrambi i sessi), ad indicare come il fenomeno sia rimasto invariato<sup>3</sup> nei due periodi presi in considerazione. Su un totale di 5822 maschi, il 74.65% non ha presentato alcun cambiamento, mentre per le femmine risulta una percentuale minore (ma comunque molto alta) di resilienza (61.17%).

### **3.2.2 Solitudine e Paese**

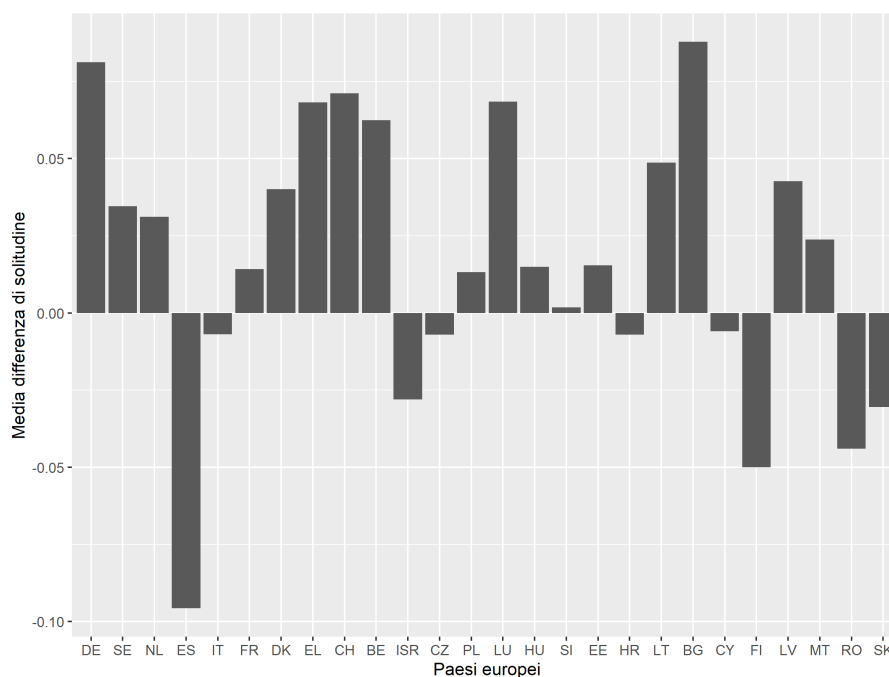
In questo paragrafo ci si concentra sull'analisi della variazione di solitudine esaminando anche le differenze tra i diversi Paesi europei.

Il grafico in Figura 3.2 e la Tabella 3.3 riassumono le medie sulle differenze di solitudine tra la condizione pre-Covid e la condizione durante

---

<sup>3</sup>Con il termine "invariato" si intende che se prima della pandemia il soggetto si sentiva solo, anche durante la pandemia si è sentito solo (in egual modo per la non percezione di solitudine).

il Covid per ciascun Paese europeo.



**Figura 3.2:** Istogramma sulla media della differenza di solitudine tra prima e durante il Covid-19 per ogni Paese europeo.

<b>PAESE</b>	<u>Media</u>	<b>PAESE</b>	<u>Media</u>
DE	0.08	LU	0.07
SE	0.03	HU	0.01
NL	0.03	SI	0.001
ES	-0.10	EE	0.02
IT	-0.01	HR	-0.01
FR	0.01	LT	0.05
DK	0.04	BG	0.09
EL	0.07	CY	-0.01
CH	0.07	FI	-0.05
BE	0.06	LV	0.04
ISR	-0.02	MT	0.02
CZ	-0.01	RO	-0.04
PL	0.01	SK	-0.03

**Tabella 3.3:** Tabella sulla media della Y per ogni Paese.

Questa analisi rappresenta la media della variazione di solitudine tra prima e durante il Covid-19 per ogni Paese europeo. I risultati suggeriscono che, in media, la differenza di solitudine non è stata esageratamente in-

fluenzata dal Covid-19 in molti Paesi. Tuttavia, ci sono alcune realtà in cui la differenza di solitudine è aumentata, ad esempio in Bulgaria, o in Germania.

Al contrario, ci sono anche alcuni Paesi in cui la media della differenza di solitudine è diminuita durante il Covid-19, ad esempio in Spagna, che nell'istogramma risulta abbastanza evidente, e a seguire la Romania e la Finlandia.

È importante notare che questi risultati si basano su una media delle differenze di solitudine, e quindi non tengono conto della distribuzione all'interno di ciascun Paese. Inoltre, la differenza di solitudine potrebbe essere influenzata da molti fattori oltre al Covid-19, come il livello di isolamento sociale prima della pandemia, il tipo di lavoro svolto, la situazione familiare e altri fattori socio-demografici.

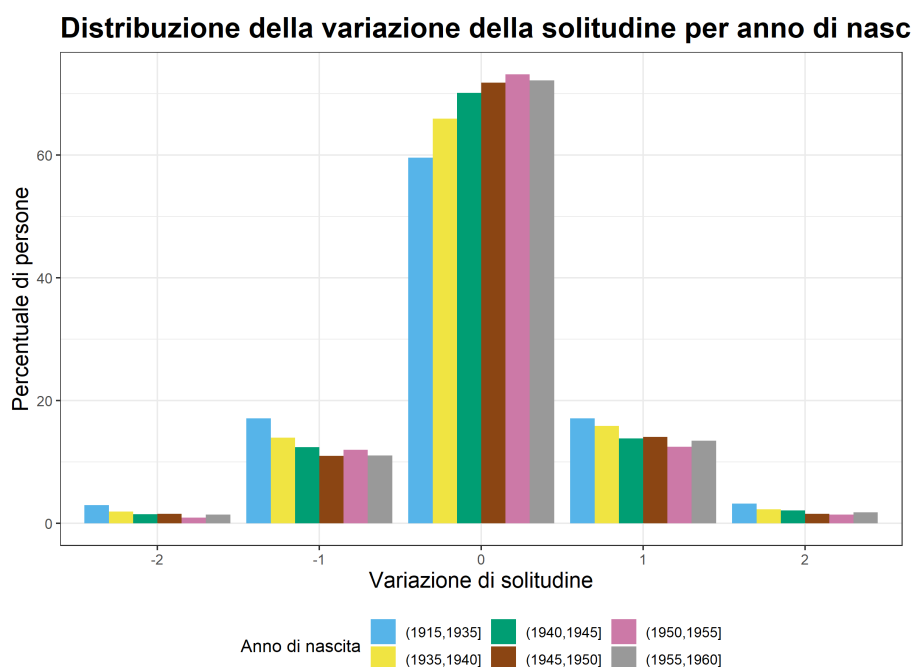
### **3.2.3 Solitudine ed età**

L'età avanzata è un fattore che può influenzare la percezione di solitudine. Gli anziani, come già descritto in precedenza, possono trovarsi a fronteggiare un cambiamento nella propria vita quotidiana, spesso caratterizzato da una minore interazione sociale e da una riduzione del supporto sociale. La solitudine può essere quindi una sfida per molte persone anziane: la pandemia di Covid-19 potrebbe aver ulteriormente aumentato il rischio di isolamento e, quindi, della solitudine.

In questa sezione, viene esplorato l'effetto dell'età sulla percezione di solitudine (Tabella 3.4 e Figura 3.3), ricordando che il campione comprende solo i soggetti che hanno dai 60 anni in su.

	Y				
	-2	-1	0	1	2
(1915,1935]	3.00	17.09	59.58	17.09	3.25
(1935,1940]	1.95	13.94	65.94	15.89	2.27
(1940,1945]	1.52	12.40	70.18	13.81	2.10
(1945,1950]	1.58	11.00	71.79	14.08	1.55
(1950,1955]	0.93	12.00	73.15	12.51	1.42
(1955,1960]	1.44	11.07	72.18	13.47	1.83

**Tabella 3.4:** Tabella di frequenza relativa - Variazione % della solitudine per anno di nascita.



**Figura 3.3:** Istogramma della variazione % di solitudine in base all'anno di nascita.

In particolare, l'analisi evidenzia che le fasce di età più giovani (nati negli anni '50 e '60) hanno subito variazioni minori rispetto alle fasce di età più anziane, che hanno registrato valori più alti di variazione. Infatti, la maggiore percentuale di individui che hanno segnalato un peggioramento (ma anche miglioramento) della solitudine, è quello compreso tra il 1915 e il 1935.

Il peggioramento potrebbe essere dovuto al fatto che le fasce di età più anziane hanno maggiori difficoltà ad adattarsi alle nuove tecnologie e ai

cambiamenti delle dinamiche sociali, mentre le fasce di età più "giovani" hanno maggiori opportunità di utilizzare strumenti tecnologici per mantenere le connessioni sociali. Il miglioramento invece può trovare spiegazione nel fatto che in molte società, i membri della famiglia o i caregiver possono aver dedicato più tempo e attenzione ai più anziani durante la pandemia per garantire che avessero supporto sociale e materiale adeguato.

In generale, i dati suggeriscono che la pandemia non ha avuto un impatto rilevante sulla solitudine degli individui, poichè le percentuali maggiori si concentrano nello 0 (come nelle analisi precedenti e come in quelle che vedremo), ad indicare che questi anziani europei hanno riportato di non aver avvertito alcuna variazione nella loro solitudine durante il periodo considerato.

Potrebbe anche essere che, nonostante ci sia stata una pandemia globale, ci siano state alcune persone che hanno continuato ad avere una vita sociale attiva, o che hanno trovato modi alternativi per mantenere i loro contatti sociali. Inoltre, è possibile che alcuni soggetti abbiano percepito il periodo del Covid come un'opportunità per rallentare e dedicarsi ad attività solitarie che non avevano tempo di fare in precedenza.

Risulta importante ricordare che gli anziani potrebbero essere meno colpiti dalla solitudine in quanto, a causa della loro età, normalmente potrebbero avere meno attività sociali o relazioni di amicizia rispetto ai giovani. Un'altra motivazione potrebbe essere la maggiore esperienza del campione considerato che ha come conseguenza l'essere in grado di affrontare la solitudine in modo più efficace rispetto, ad esempio, alla

popolazione giovane.

### 3.2.4 Solitudine e presenza del partner e/o figli

Questa sezione si propone di esplorare come la presenza del partner e/o dei figli abbia influenzato la variazione della solitudine in un periodo particolarmente difficile come quello della pandemia. L'obiettivo principale di questa analisi è quello di valutare se la presenza di una figura di supporto familiare abbia un effetto protettivo sulla solitudine delle persone, e se questo effetto vari in base alla presenza o all'età dei figli.

Saranno proposte 3 sotto-sezioni:

- a) Relazione tra  $Y$  e  $x_1 =$  "Presenza del partner"
- b) Relazione tra  $Y$  e  $x_2 =$  "Numero di figli"
- c) Relazione congiunta tra  $Y$ ,  $x_1$  e  $x_2$ .

#### a) $Y$ rapportata alla presenza o assenza del partner

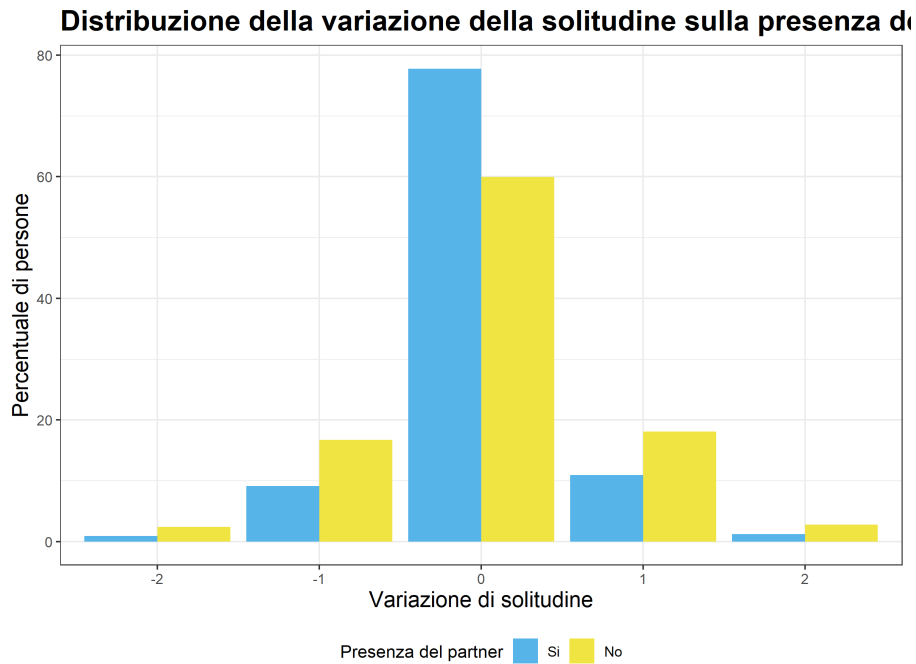
La pandemia ha portato a un cambiamento significativo nel modo in cui le persone si relazionano tra loro e si sentono socialmente. Molte persone hanno sperimentato un aumento della solitudine durante la pandemia a causa del distanziamento fisico, della quarantena, dell'isolamento e della chiusura di molte attività sociali. In questo contesto, il ruolo del partner può essere particolarmente importante per il benessere emotivo e la salute mentale delle persone.

Potrebbe, quindi, essere interessante esplorare se la presenza del partner ha un effetto protettivo contro la solitudine durante la pandemia, oppure se la pandemia ha influenzato in modo diverso la solitudine di persone

con partner rispetto a quelle senza. A tal fine, si è deciso di focalizzare l'attenzione sulla variazione di solitudine tra prima e durante la pandemia in relazione alla presenza o assenza di un partner, suddividendo i dati in base a queste due condizioni e costruendo la Tabella 3.5 con il relativo grafico di Figura 3.4.

		Y				
		-2	-1	0	1	2
Presenza del partner	Si	0.95	9.11	77.78	10.95	1.21
	No	2.41	16.71	59.99	18.08	2.81

**Tabella 3.5:** Tabella di frequenza relativa - Variazione % della solitudine in base alla presenza del partner.



**Figura 3.4:** Istogramma della variazione % di solitudine in base alla presenza del partner.

Risulta evidente come gli andamenti di questo grafico siano simili a quelli delle analisi precedenti, poichè, ancora una volta, le percentuali maggiori si concentrano nello zero, suggerendo che la solitudine percepita non è cambiata drasticamente durante la pandemia. Ma si può anche notare come le persone senza partner abbiano subito maggiori variazioni

nella solitudine rispetto agli individui con un partner. Infatti, nella riga "No" si osserva che i valori percentuali sono più elevati rispetto alla riga "Si" per le categorie di variazione della solitudine positive e negative, mostrando che le persone senza un partner presentano una maggiore percentuale di peggioramento della solitudine rispetto a chi ne ha uno. D'altra parte, i dati mostrano anche una maggiore percentuale di miglioramento della solitudine per le persone che vivono da sole. Questo potrebbe suggerire che la solitudine per alcune persone è stata mitigata dal fatto di doversi confrontare con essa più frequentemente, a causa della mancanza della presenza del partner. Un'altra causa possibile è che coloro che non hanno un partner potrebbero essere più motivate a cercare relazioni sociali al di fuori della coppia e a trovare modi creativi per mantenere il contatto con gli altri durante la pandemia.

#### **b) $Y$ rapportata al numero di figli**

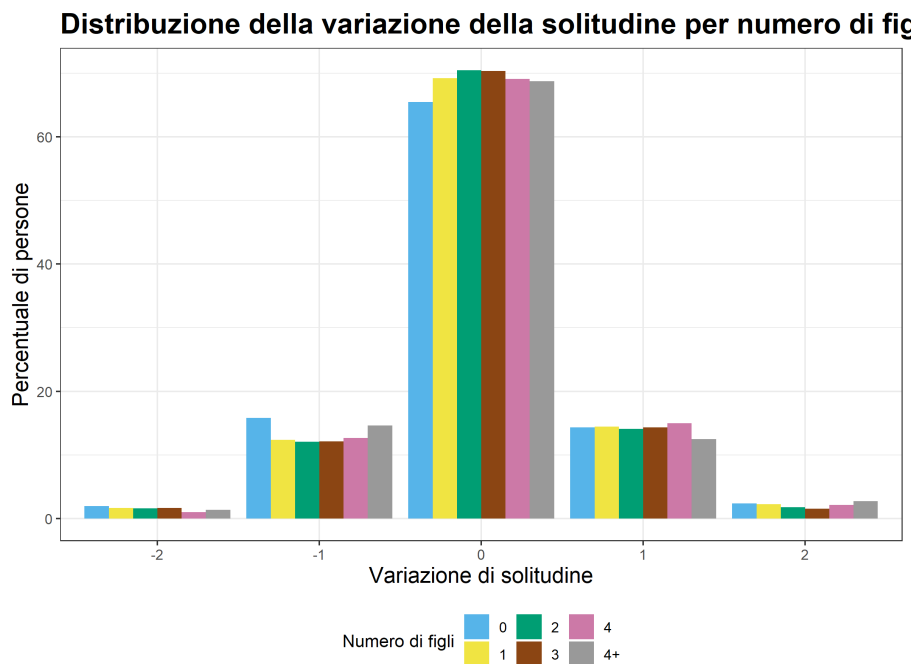
In questo caso è stata esplorata la possibile influenza del numero di figli sulla percezione di solitudine durante la pandemia. Questa analisi può essere importante per comprendere il ruolo dei figli all'interno dell'esperienza di solitudine degli anziani.

Il numero dei figli è spesso considerato un indicatore del supporto sociale e dell'aiuto pratico che una persona anziana può ricevere. In questo senso, un suo aumento o diminuzione potrebbe influire sulla percezione di solitudine.



		Y				
		-2	-1	0	1	2
Numero di figli	0	2.00	15.83	65.46	14.35	2.37
	1	1.70	12.40	69.18	14.45	2.27
	2	1.61	12.07	70.43	14.11	1.79
	3	1.65	12.13	70.36	14.33	1.54
	4	1.03	12.70	69.11	14.99	2.17
	4+	1.37	14.65	68.75	12.50	2.73

**Tabella 3.6:** Tabella di frequenza relativa - Variazione % della solitudine in base al numero di figli.



**Figura 3.5:** Istogramma della variazione % di solitudine in base al numero di figli.

In generale (Tabella 3.6 e Figura 3.5), i valori sono tutti relativamente vicini tra di loro e indicano una percentuale elevata di "nessuna variazione" nella solitudine durante il COVID-19. Tuttavia, ci sono alcune variazioni tra i diversi gruppi di figli. Nel grafico emerge solo un fattore che si può considerare importante: i partecipanti con zero figli hanno la percentuale più alta di valori negativi (soprattutto nella modalità -1), ma anche la percentuale più alta di miglioramento nella solitudine

(sebbene non relativamente più alta rispetto ad altri gruppi di figli), ad indicare, come nel caso precedente, variazioni maggiori.

I partecipanti con quattro o più figli sembrano aver avuto la percentuale più bassa di miglioramento nella solitudine <sup>4</sup>.

Questi risultati potrebbero essere spiegati in vari modi. Ad esempio, i partecipanti senza figli potrebbero avere meno supporto sociale rispetto a quelli con figli e potrebbero quindi essere stati più vulnerabili alla solitudine durante il Covid. D'altra parte, coloro che hanno quattro o più figli potrebbero aver avuto difficoltà ad avere abbastanza spazio e privacy, il che potrebbe aumentare la loro solitudine durante la pandemia.

Dalle analisi fin'ora condotte risulta quindi che la maggior parte degli individui si concentra nella modalità 0 <sup>5</sup>, non facendo emergere risultati esageratamente differenti (nonostante questa sia comunque un'informazione valida). Ma è importante ricordare che quelli che potrebbero essere classificati come gli anziani più "a rischio" cioè senza partner, senza figli, e le donne, hanno riportato maggiori variazioni (chiaramente non rilevanti), sia in termini positivi (1 e 2), ma soprattutto in termini negativi (-2,-1), abbassando il livello della percentuale della modalità 0, evidenziando in questo modo la maggiore variabilità della Y rispetto queste categorie di soggetti.

---

<sup>4</sup>Viene presa in considerazione la somma tra 1 e 2.

<sup>5</sup>Si ricorda che lo 0 non indica che gli individui non si sono sentiti soli, ma che la loro percezione di solitudine non è cambiata tra il pre pandemia e la prima ondata.

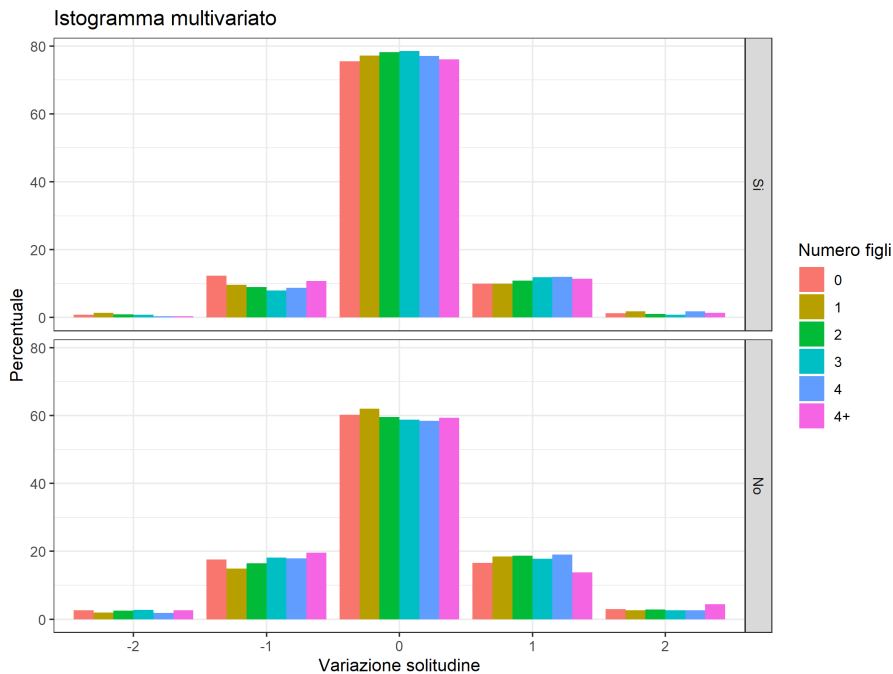
c)  $Y$  rapportata al numero di figli e alla presenza del partner congiuntamente

		<i>PARTNER NO</i>					
		Figli					
		0	1	2	3	4	4+
Y	-2	2.58	2.01	2.48	2.75	1.88	2.68
	-1	17.64	14.86	16.43	18.12	17.96	19.64
	0	60.22	62.05	59.55	58.79	58.45	59.38
	1	16.63	18.47	18.69	17.76	19.03	13.84
	2	2.92	2.61	2.86	2.58	2.68	4.46

**Figura 3.6:** Tabella trivariata - Variazione % della solitudine in base al numero dei figli, senza la presenza del partner.

		<i>PARTNER SI</i>					
		Figli					
		0	1	2	3	4	4+
Y	-2	0.87	1.35	0.98	0.87	0.40	0.35
	-1	12.34	9.63	8.96	7.92	8.78	10.76
	0	75.54	77.20	78.19	78.48	77.05	76.04
	1	9.96	9.93	10.84	11.92	11.98	11.46
	2	1.30	1.88	1.03	0.81	1.80	1.39

**Figura 3.7:** Tabella trivariata - Variazione % della solitudine in base al numero dei figli e con la presenza di un partner.



**Figura 3.8:** Istogramma della variazione % di solitudine in base al numero di figli e alla presenza del partner ("Si"=presenza del partner, "No"= assenza del partner).

Questa sezione mostra la percentuale di individui che hanno riportato una variazione nella percezione di solitudine, suddivisi per presenza/assenza di partner e per numero di figli.

In Tabella 3.6 (e il secondo pannello del grafico riportato nella Figura 3.8), si nota che la percentuale di individui che ha riportato una va-

riazione nella percezione di solitudine rispetto al periodo precedente è generalmente più alta rispetto alla situazione in cui c'è un partner. In particolare, è possibile notare che la percentuale di individui che hanno riportato un peggioramento nella percezione di solitudine sembra aumentare al crescere del numero di figli.

Possibili motivazioni di quest'ultima affermazione:

- Impegno maggiore per le persone con figli e senza un partner: le persone con figli potrebbero avere meno tempo libero a disposizione rispetto a chi non ha figli, a causa dell'impegno richiesto dalla cura dei figli stessi, il che potrebbe ridurre la loro capacità di partecipare ad attività sociali e interagire con gli altri. Inoltre, con la pandemia, molte persone con figli potrebbero aver dovuto affrontare il doppio carico di lavoro dovuto alla gestione della casa e dei figli e di conseguenza potrebbero sentirsi più soli.
- Minore supporto sociale: con le restrizioni alle visite e alle attività sociali, potrebbe essere stata limitata anche la possibilità di interagire con la famiglia, aumentando così la percezione di solitudine.
- Fattori socio-economici: questi soggetti potrebbero avere un reddito inferiore rispetto a chi non ha figli e potrebbero essere più colpite dalle conseguenze economiche della pandemia, come la perdita di lavoro o la riduzione delle entrate.

In Tabella 3.7 (e nel primo pannello del grafico riportato nella Figura 3.8), cioè quella relativa alla presenza del partner, si può notare che la maggior parte dei partecipanti ha riportato la stessa percezione di solitudine rispetto al periodo precedente, quindi non ci sono state variazioni

come prima. In particolare, per la maggioranza delle situazioni con un diverso livello di solitudine riportato, la percentuale è più alta per coloro che non hanno figli, ma in generale non sembrano esserci grandi differenze tra le diverse situazioni. Tuttavia, si può notare che, in presenza di partner, la percentuale di individui che hanno riportato un peggioramento nella percezione di solitudine sembra aumentare leggermente dai 4+ figli (le cui ragioni potrebbero essere le stesse riportate prima).

In generale, questi dati suggeriscono che la presenza di un partner sia associata a una minore probabilità di variazioni nella percezione di solitudine, mentre l'assenza sembra essere associata a una maggiore probabilità di variazioni, soprattutto in negativo, nella percezione di solitudine, in particolare per le famiglie con più figli.

### **3.2.5 Solitudine, titolo di studio e gestione economica**

La variabile "gestione economica" (nel testo verrà anche chiamata "valutazione sulla solvibilità"<sup>6</sup> per riassumere meglio il concetto) si riferisce alla percezione sulla capacità di far quadrare i conti, già presentata nel Capitolo 2 (2.2.3).

Come citato in precedenza, il titolo di studio può influire sulla capacità di una persona di relazionarsi con gli altri e di gestire la solitudine, mentre la valutazione della solvibilità può riflettere il grado di autonomia economica e di indipendenza dell'anziano, fattori che possono a loro volta influire sulla percezione di solitudine. Considerando che le persone con un livello di istruzione più elevato hanno maggiori opportunità la-

---

<sup>6</sup>Questo termine si riferisce alla capacità di un individuo, di una famiglia o di un'organizzazione di saldare i debiti e di far fronte alle spese correnti e future, senza dover affrontare situazioni di insolvenza o di fallimento.

vorative e tendono a guadagnare di più, potrebbero valutare la propria capacità di far fronte alla pandemia in modo diverso (più positivo) rispetto a coloro con un livello di istruzione inferiore. In questo contesto, potrebbe essere interessante esaminare se ci sia una relazione tra questa disparità socioeconomica e la variazione della solitudine degli anziani durante la pandemia.

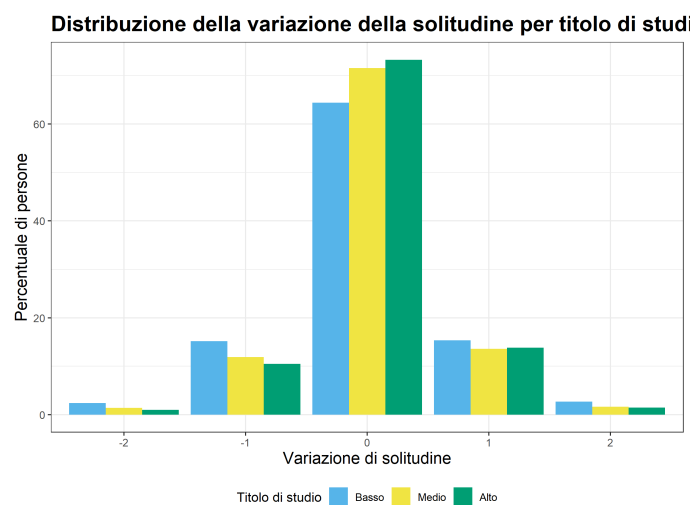
Anche qui vengono proposte 3 sotto-sezioni:

- a) Relazione tra  $Y$  e  $x_1 = \text{"Titolo di studio"}$
- b) Relazione tra  $Y$  e  $x_2 = \text{"Valutazione della solvibilità"}$
- c) Relazione congiunta tra  $Y$ ,  $x_1$  e  $x_2$ .

**a)  $Y$  rapportata al titolo di studio**

		Y				
		-2	-1	0	1	2
Titolo di studio	Basso	2.40	15.18	64.38	15.34	2.70
	Medio	1.39	11.90	71.53	13.56	1.62
	Alto	1.00	10.49	73.26	13.80	1.45

**Tabella 3.7:** Tabella di frequenza relativa - Variazione % della solitudine in base al titolo di studio.



**Figura 3.9:** Istogramma della variazione % di solitudine in base al titolo di studio.

Il grafico in Figura 3.9 mette in evidenza che l'andamento è molto simile a quello delle analisi precedenti: le maggiori percentuali si registrano nella modalità "nessun cambiamento".

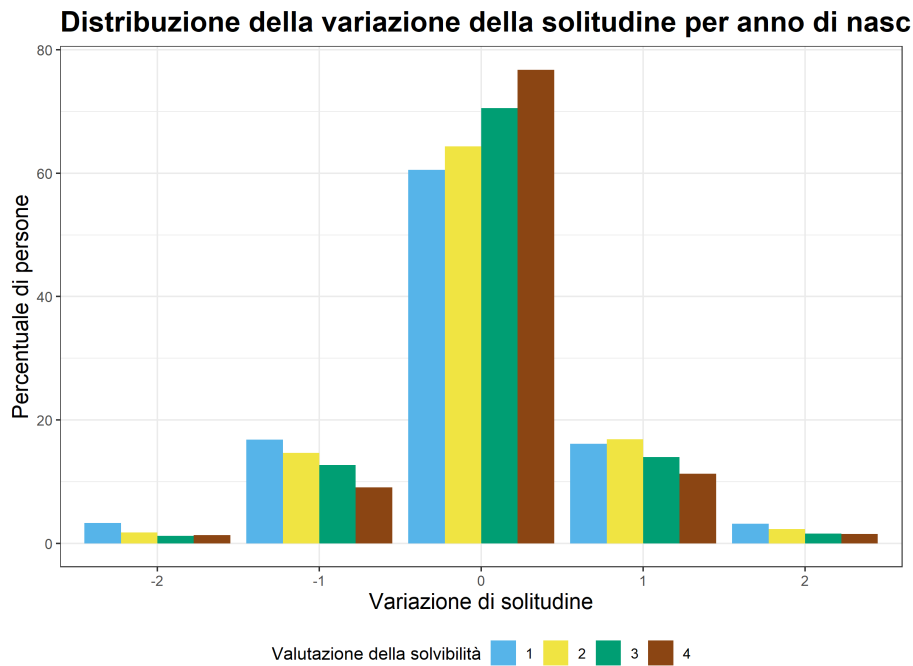
In particolare la Tabella 3.7 mostra la variazione maggiore negli individui con un titolo di studio più basso. Infatti, hanno una percentuale più alta di peggioramento della loro solitudine rispetto alle altre due categorie. Si nota anche che in minima parte sono i soggetti con le proporzioni un po' più alte di miglioramento, poichè chi ha un'istruzione bassa potrebbe anche avere meno aspettative iniziali rispetto alla propria situazione di vita e potrebbe essere più abituato a situazioni complesse da gestire.

Al contrario, le persone con un alto titolo di studio sembrano essere quelle con una minore percentuale di peggioramento, e una stabilità maggiore (% più alta nello zero). Una possibile motivazione va ricercata nelle maggiori risorse e strumenti a disposizione di chi ha questo tipo di istruzione, utilizzate per far fronte alle sfide della pandemia, come lavorare da remoto, utilizzare le tecnologie per rimanere in contatto con gli altri e accedere a fonti di informazione affidabili, per adattarsi meglio alla nuova situazione, riducendo così la variazione della solitudine.

**b) Y rapportata alla valutazione della solvibilità**

		Y				
		-2	-1	0	1	2
Valutazione della solvibilità	1=Con grande difficoltà	3.29	16.84	60.53	16.16	3.17
	2=Con qualche difficoltà	1.76	14.67	64.34	16.90	2.33
	3=Abbastanza facilmente	1.21	12.69	70.52	13.99	1.59
	4=Facilmente	1.34	9.06	76.77	11.31	1.52

**Tabella 3.8:** Tabella di frequenza relativa - Variazione % della solitudine in base alla valutazione della solvibilità



**Figura 3.10:** Istogramma della variazione % di solitudine in base al giudizio sulla solvibilità.

I dati riassunti nella Tabella 3.8 e nel grafico in Figura 3.10 mostrano che, all’aumentare del giudizio positivo sulla solvibilità, le percentuali di miglioramento e peggioramento diminuiscono, e quella di nessuna variazione (zero) aumenta, indicando quindi una minore variazione complessiva nella percezione di solitudine (è possibile notarlo molto bene dalle barre centrate sullo zero che salgono in maniera crescente). Come già incontrato nel resto delle analisi su altre variabili esplicative, potrebbe essere che le persone che percepiscono una maggiore solvibilità hanno una situazione finanziaria più stabile e sicura, quindi hanno meno probabilità di subire una variazione di solitudine durante la pandemia. Al contrario, all’aumentare del giudizio negativo sulla solvibilità, la percentuale di individui che riportano una variazione negativa aumenta.



c) *Y* rapportata al titolo di studio e alla valutazione della solvibilità

Il livello di istruzione e la percezione della solvibilità potrebbero essere considerati come indicatori di benessere economico e sociale, che possono avere un impatto sulla percezione di solitudine e sull'eventuale variazione rispetto alla situazione pre-covid.

Si propone quindi come ultimo studio di questa analisi descrittiva eseguita sulla *Y* in rapporto alle covariate ritenute più interessanti, grafici e tabelle trivariate, al fine di valutare dei possibili legami tra loro.

<i>Titolo di studio: BASSO</i>						
		Valutazione sulla solvibilità				
		0	1	2	3	4
Y	-2	3.68	2.27	1.83	2.42	1.88
	-1	17.67	16.06	15.07	12.21	17.96
	0	59.55	60.36	66.73	70.38	58.45
	1	15.78	18.33	14.35	12.21	19.03
	2	3.32	2.98	2.02	2.78	2.68

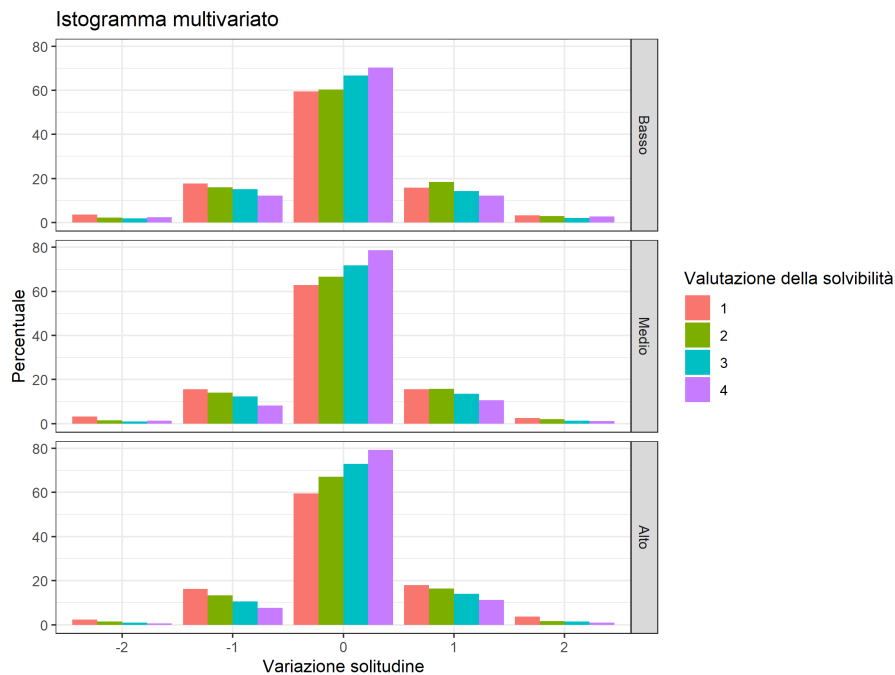
**Figura 3.11:** Tabella trivariata - Variazione % della solitudine in base ad un livello di istruzione basso e alla valutazione sulla solvibilità.

<i>Titolo di studio: MEDIO</i>						
		Valutazione sulla solvibilità				
		0	1	2	3	4
Y	-2	3.18	1.45	0.95	1.31	1.88
	-1	15.68	14.02	12.30	8.20	17.96
	0	62.92	66.65	71.79	78.60	58.45
	1	15.68	15.79	13.63	10.64	19.03
	2	2.54	2.08	1.33	1.25	2.68

**Figura 3.12:** Tabella trivariata - Variazione % della solitudine in base ad un livello di istruzione medio e alla valutazione sulla solvibilità.

<i>Titolo di studio: ALTO</i>						
		Valutazione sulla solvibilità				
		0	1	2	3	4
Y	-2	2.38	1.44	0.88	0.61	1.88
	-1	16.33	13.40	10.58	7.74	17.96
	0	59.52	67.11	73.01	79.36	58.45
	1	18.03	16.39	14.08	11.36	19.03
	2	3.74	1.66	1.46	0.92	2.68

**Figura 3.13:** Tabella trivariata - Variazione % della solitudine in base ad un livello di istruzione alto e alla valutazione sulla solvibilità.



**Figura 3.14:** Istogramma della variazione % di solitudine in base al titolo di studio e alla valutazione sulla solvibilità.

Le tre tabelle (in Figura 3.11, 3.12, e 3.13) con i relativi grafici (in Figura 3.14), riportano la distribuzione delle frequenze della variabile  $Y$ , divisa per i livelli di giudizio sulla solvibilità in quattro modalità, a seconda del titolo di studio dei pensionati. L'evidenza mostra che all'aumentare del titolo di studio aumentano leggermente le percentuali di  $Y$  pari a 0 (nessuna variazione della solitudine) e diminuiscono le percentuali di  $Y$  pari a -1 e -2, mentre le percentuali pari a 1 e 2 restano pressoché costanti.

Questo fenomeno è ormai diventato costante all'interno di questo studio, evidenziato anche dal grafico che mostra come le barre dello zero formino sempre più una scala crescente (da una valutazione negativa a una positiva) all'aumentare del titolo di studio, ad indicare una maggiore stabilità degli anziani che hanno un'istruzione più elevata e confermando tutte le analisi simili fatte in precedenza.



## Capitolo 4

# Analisi multivariata: Modelli di regressione lineare multipla

I modelli di regressione rappresentano un importante strumento di analisi, che consente di esplorare le relazioni tra le variabili coinvolte nello studio e di valutare l'effetto di ciascuna di esse sulla variabile di interesse, ovvero la variazione di solitudine. Essi sono dei modelli matematici che consentono di descrivere la relazione tra una  $Y$  e una o più variabili indipendenti (ad esempio, il titolo di studio, l'età, il sesso e altre che vedremo). L'obiettivo dei modelli di regressione è quello di valutare l'effetto delle variabili indipendenti sulla variabile dipendente, e di determinare il grado di associazione tra le variabili.

In questo studio, verrà utilizzato il modello di regressione lineare multipla, che consentirà di analizzare l'effetto congiunto delle diverse variabili indipendenti sulla  $Y$ <sup>1</sup>. In questo modo, saremo in grado di individuare i fattori più rilevanti che influenzano la variazione di solitudine nei pensionati over 60.

---

<sup>1</sup>La descrizione della variabile risposta è stata fornita nel capitolo 3.1.

Il modello di regressione lineare multipla <sup>2</sup> può essere così definito:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \epsilon_i \quad (4.1)$$

Dove:

- $Y_i$  è la variabile dipendente (o variabile risposta) per l' $i$ -esimo individuo;
- $X_{i1}, X_{i2}, \dots, X_{ip}$  sono le variabili indipendenti (o covariate) per l' $i$ -esimo individuo;
- $\beta_0, \beta_1, \beta_2, \dots, \beta_p$  sono i coefficienti di regressione rispettivamente dell'intercetta e delle variabili indipendenti;
- $\epsilon_i$  è un termine d'errore per l' $i$ -esimo individuo, che traduce l'incapacità del modello di riprodurre con esattezza la realtà osservata.

In questo capitolo, sarà analizzato l'output del modello di regressione multipla in R. L'output include diverse statistiche e misure che ci consentono di valutare la bontà di adattamento del modello e l'importanza di ciascuna variabile esplicativa.

Le stime dei coefficienti delle variabili indipendenti rappresentano gli effetti marginali di ciascuna variabile indipendente sul valore della variabile dipendente. In particolare, il coefficiente di una variabile esplicativa rappresenta l'incremento/decremento della variabile dipendente associato all'aumento di una unità della variabile esplicativa, mantenendo costanti tutte le altre variabili.

---

<sup>2</sup>Inteso come la funzione che lega la variabile dipendente a quelle indipendenti.

## 4.0.1 Le variabili esplicative presenti nel modello di regressione

Oltre alle covariate già discusse ampiamente nei capitoli precedenti (sesso, anno di nascita, titolo di studio, giudizio sulla solvibilità, presenza del partner e numero di figli), all'interno del modello di regressione lineare multipla saranno presenti anche altre variabili esplicative<sup>3</sup> (tutte qualitative), riportate nella Tabella 4.1 ai fini di una maggiore chiarezza:

Codifica della variabile	Domanda	Modalità di risposta
mh007_	Recentemente ha avuto problemi a dormire?	1= Sì 2= No
mh010_	Recentemente le capita di diventare nervoso/a o di arrabbiarsi?	1= Sì 5= No
mh011_	Com'è stato il suo appetito nell'ultimo mese?	1= Diminuzione del desiderio di cibo 2= Nessuna diminuzione 3= Risposta generica
ph	Direbbe che la sua salute è...	1= Ottima 2= Molto buona 3= Buona 4= Discreta 5= Cattiva
ph004_	Lei ha qualche problema di salute, malattia, invalidità o infermità a lungo termine?	1= Sì 5=No
cah002_	Se confronta la Sua salute con quella prima dell'inizio dell'epidemia, direbbe che la Sua salute è:	1= Migliorata 2= Peggiorata 3= Invariata
cah003_	Dalla nostra ultima intervista, le sono stati diagnosticati una grave malattia o un problema di salute?	1= Sì 5= No
cah020_	Nell'ultimo mese si è sentito/a nervoso/a, ansioso/a o al limite?	1= Sì 5= No
camh007_	Recentemente ha avuto problemi a dormire?	1= Sì 2= No
cac013_	Qualcuno vicino a Lei è morto a causa dell'infezione da coronavirus?	1= Sì 5= No
cas015_	Dall'inizio dell'epidemia, Lei ha svolto altre attività di volontariato?	1= Sì 5= No
cas020_	Dall'inizio dell'epidemia, ha ricevuto aiuto, da persone che non vivono con Lei, per procurarsi beni o servizi di prima necessità, ad es. cibo, farmaci o riparazioni domestiche di emergenza?	1= Sì 5= No

**Tabella 4.1:** Variabili esplicative e relative modalità presenti nel *Dataset*.

### Gestione delle variabili categoriali nei modelli di regressione: un esempio

Le variabili qualitative possono rappresentare un'importante sfida nella modellizzazione di un problema di regressione. Considerando che nel modello sono presenti soprattutto variabili di questo tipo, è utile esaminare brevemente come gestirle correttamente e come interpretare i risultati del modello.

<sup>3</sup>Risulta importante sottolineare che la variabile presente nella Tabella 4.1 con il termine "ph" è l'unica esplicativa che viene considerata numerica (le altre sono delle fattoriali, cioè variabili che assumono un insieme di valori discreti o categorici. Le variabili precedute dal prefisso "ca", si riferiscono alle domande del questionario poste durante la prima ondata di Covid.

Queste variabili qualitative sono generalmente convertite in variabili dummy per poter essere utilizzate in un modello di regressione lineare multipla.

In generale, un coefficiente positivo per un livello della variabile qualitativa indica che i soggetti che hanno quella caratteristica tendono ad avere un outcome maggiore rispetto ai soggetti senza la caratteristica (il livello di riferimento), e viceversa per un coefficiente negativo.

#### 4.0.2 Stima del modello di regressione lineare multipla completo

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.34196 -0.11449 -0.00771  0.07727  2.25341

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.050649   0.069002   0.734 0.462947
mh007_2      0.051783   0.012008   4.312 1.63e-05 ***
mh010_5      0.079945   0.012749   6.271 3.70e-10 ***
mh011_2      0.044729   0.018446   2.425 0.015328 *
mh011_3      0.055523   0.043473   1.277 0.201556
SessoFemmina 0.035881   0.011372   3.155 0.001607 **
ph004_5     -0.015895   0.011950  -1.330 0.183498
cah002_2     0.107659   0.035253   3.054 0.002263 **
cah002_3     0.027513   0.031085   0.885 0.376120
cah003_5    -0.012959   0.019853  -0.653 0.513948
cah020_5    -0.137450   0.012443 -11.046 < 2e-16 ***
camh007_2   -0.078840   0.012934  -6.096 1.12e-09 ***
cac013_5    -0.033743   0.033717  -1.001 0.316954
cas015_5     0.026247   0.023855   1.100 0.271235
cas020_5    -0.044736   0.011713  -3.819 0.000134 ***
partNo      -0.016918   0.011536  -1.466 0.142546
figli1       0.035346   0.021082   1.677 0.093644 .
figli2       0.024989   0.019202   1.301 0.193133
figli3       0.023937   0.021313   1.123 0.261415
figli4       0.050398   0.027706   1.819 0.068929 .
figli4+     0.014703   0.033094   0.444 0.656843
lavoro2      0.041306   0.018848   2.192 0.028426 *
lavoro3      0.029324   0.018644   1.573 0.115772
lavoro4      0.033381   0.019276   1.732 0.083335 .
anno(1935,1940] 0.015031   0.020867   0.720 0.471344
anno(1940,1945] 0.023145   0.020104   1.151 0.249645
anno(1945,1950] 0.022972   0.019847   1.157 0.247108
anno(1950,1955] 0.013956   0.020574   0.678 0.497592
anno(1955,1960] 0.034979   0.026573   1.316 0.188079
studioMedia  0.016276   0.012745   1.277 0.201611
studioAlta  0.033087   0.013802   2.397 0.016531 *
ph           -0.026945   0.006508  -4.141 3.48e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6339 on 14737 degrees of freedom
Multiple R-squared:  0.02373, Adjusted R-squared:  0.02167
F-statistic: 11.55 on 31 and 14737 DF, p-value: < 2.2e-16

```

Figura 4.1: Output del modello di regressione lineare multipla.

L'output di questo modello di regressione lineare multipla in Figura 4.1, indica che solo alcune delle variabili indipendenti considerate sono significativamente associate alla variazione di solitudine tra prima e durante la pandemia di Covid-19. In particolare, le variabili indipendenti significative al 5% di livello, includono *mh007\_2*, *mh010\_5*, *mh011\_2*, *SessoFemmina*, *cah002\_2*, *cah020\_5*, *camh007\_2*, *cas020\_5*, *lavoro2*, *studioAlta* e *ph*. Altri coefficienti non sono statisticamente significativi al 5% di livello, come tutte le variabili sull'anno di nascita, quella sull'assenza del partner e le altre visibili nel modello.

Questo modello non sembra avere un buon adattamento. Il valore del **residual standard error** (RSE) è di 0,6339, il che indica che i valori osservati si discostano in media di circa 0,63 unità dai valori predetti dal modello.

L' $R^2$  del modello è molto basso (0,0237), il che suggerisce che solo il 2,37% della variazione nella solitudine può essere spiegato dalle variabili indipendenti incluse nel modello. Da una parte il set di covariate incluse nel modello non è troppo ampio. Dall'altra non va dimenticato che la variabile dipendente ha un range di variazione limitato per costruzione (numeri interi da -2 a +2).

Anche il *p-value* del test F è molto basso e, quindi, significativo ( $< 0.05$ ), e ci indica che il modello preso in considerazione si adatta meglio ai dati rispetto ad un modello con sola intercetta.

In sintesi, i risultati suggeriscono che solo alcune delle variabili indipendenti considerate sono significativamente associate alla variazione di solitudine durante la pandemia di Covid-19, e queste variabili spiegano



solo una piccola percentuale della variazione nella solitudine.

## **ANALISI DEI COEFFICIENTI**

I coefficienti del modello indicano l'effetto di ciascuna variabile indipendente sulla variabile dipendente, tenendo conto degli effetti delle altre covariate. Come anticipato, di default R parametrizza le variabili categoriali rispetto al parametro d'angolo, quindi rispetto alla prima modalità assunta da ogni predittore categoriale. Risulta importante sottolineare che, nel caso preso in esame, un coefficiente negativo stimato per una variabile indipendente indica che un aumento unitario in quella variabile è associato ad un aumento del peggioramento della condizione di solitudine (cioè un valore di  $Y$  più vicino a -2 o -1). Al contrario, un coefficiente positivo stimato per una variabile indipendente indica che un aumento unitario in quella variabile è associato ad un aumento del miglioramento (cioè un valore di  $Y$  più vicino a 1 o 2).

Ad esempio, il coefficiente stimato per  $mh007\_2$  è 0.051783, il che significa che per un aumento di una unità in questa variabile, la  $Y$  aumenterà di 0.051783 unità, tenendo costanti tutte le altre variabili. Quindi, il modello suggerisce che coloro che hanno risposto "No" alla domanda su problemi di sonno hanno una media più alta di  $Y$  rispetto a coloro che hanno risposto "Sì", e che, quindi, questi soggetti (essendo il coefficiente stimato positivo) hanno subito un miglioramento della sensazione di solitudine.

Il risultato suggerisce anche che la variabile  $mh010\_5$  ha un effetto statisticamente significativo sulla variazione di solitudine, con un valore di

coefficiente stimato di 0.079945. Ciò indica che coloro che hanno risposto "No" alla domanda "recentemente le capita sentirsi nervoso o arrabbiato" hanno sperimentato un miglioramento della solitudine rispetto a quelli che hanno risposto "Sì". Sulla stessa linea, il risultato della variabile *mh011\_2* suggerisce che coloro che hanno risposto "nessuna diminuzione del desiderio di cibo" alla domanda su com'è stato l'appetito nell'ultimo mese, hanno sperimentato un miglioramento della solitudine rispetto a quelli che hanno risposto "diminuzione".

Alcune variabili significative di questo modello (come anche di tutti i prossimi che vedremo), come *cah002\_2*, *cah020\_5* e *camh007\_2*, che riguardano rispettivamente aspetti sulla salute percepita rispetto a prima del Covid-19, sensazioni di rabbia o nervosismo e eventuali problemi di insonnia, hanno un'interpretazione un po' complessa. I segni dei coefficienti danno dei risultati a prima vista controintuitivi (per esempio, per chi nell'ultimo mese non si è sentito nervoso e non ha avuto problemi a dormire, ha una probabilità maggiore di peggioramento della condizione di solitudine). Tuttavia potrebbe essere che la mancanza di stress o problemi di sonno potrebbe aver impedito alla persona di sviluppare strategie di coping<sup>4</sup> efficaci per affrontare il senso di solitudine e isolamento. Quando la solitudine diventa più intensa e prolungata, potrebbe essere difficile per la persona adattarsi e trovare nuove modalità di connessione sociale.

Una possibile motivazione per cui chi ha avuto un peggioramento della salute ha avuto un miglioramento nella condizione di solitudine tra

---

<sup>4</sup>Sono le strategie che le persone utilizzano per gestire le emozioni e i pensieri negativi associati a situazioni avverse o a eventi traumatici.

prima e durante il Covid (vedi positività del coefficiente di *cah002\_2*), potrebbe essere legata al fatto che l'esperienza di malattia può portare ad un maggior senso di connessione con gli altri. Infatti, durante la malattia, molte persone cercano e ricevono supporto emotivo da amici, familiari e professionisti sanitari. Questo supporto può aumentare il senso di appartenenza e di connessione sociale, diminuendo il senso di solitudine. Inoltre, la malattia può portare ad un aumento della consapevolezza della propria vulnerabilità e della dipendenza dagli altri, favorendo un maggior senso di empatia e di comprensione per gli altri. Quindi, in questi casi, è importante non soffermarsi sui risultati superficialmente e avere ben chiaro che tipo di analisi si sta conducendo, per capire se in realtà alcune variabili colgano molto di più. E nel caso non si riesca a fornire una spiegazione logica è opportuno indagare altri aspetti (per esempio, errori nella raccolta dati, di misura o nel questionario).

La stima del coefficiente associato alla variabile *SessoFemmina* è di 0.035881, il che significa che, tenendo conto di tutte le altre variabili nel modello, le donne hanno un aumento medio di 0.036 unità nella variazione della solitudine tra prima e durante la pandemia rispetto agli uomini. Perciò questo risultato positivo implica un miglioramento, per quanto marginale, che fa capire che il livello di solitudine delle donne anziane è migliorato rispetto a quello maschile.

Allo stesso modo, il risultato della variabile *mh011\_2* indica che coloro che hanno risposto "No" alla domanda "Dall'inizio dell'epidemia, ha ricevuto aiuto, da persone che non vivono con Lei, per procurarsi beni o servizi di prima necessità, ad es. cibo, farmaci o riparazioni domestiche

di emergenza?", hanno avuto in media un peggioramento della solitudine rispetto a coloro che hanno risposto "Sì". Questo suggerisce che ricevere aiuto da persone esterne durante l'epidemia può avere un effetto positivo sulla solitudine delle persone.

Risulta interessante anche la significatività della variabile *studioAlta*, e la relativa stima del coefficiente positivo, ad indicare che chi ha un titolo di studio alto ha avuto un miglioramento rispetto le altre due categorie (basso e medio).

La variabile *ph* ha un coefficiente stimato di -0.026945 con un errore standard di 0.006508 e un valore p molto basso, segnalando che la variabile è altamente significativa. Inoltre, la stima del coefficiente negativo indica che all'aumentare di questa variabile <sup>5</sup> sulla percezione della propria salute fisica, la variazione evidenzia un peggioramento della condizione di solitudine, cioè c'è una relazione inversa tra queste due variabili. Concretamente, più il soggetto giudica la salute fisica peggiore, più è stato rilevato un peggioramento del livello di solitudine nei due periodi presi in considerazione.

Questi risultati confermano alcune delle analisi fatte precedentemente.

### 4.0.3 Selezione del modello tramite criterio dello step AIC

Il criterio step AIC (Akaike Information Criterion) è un metodo utilizzato per selezionare le variabili più importanti in un modello di regressione lineare.

---

<sup>5</sup>Essendo *ph* una numerica, quando si dice "all'aumentare della variabile", si intende passare dalla modalità 1 (Ottima) alla modalità 2 (Molto buona) e così via fino alla 5 (Cattiva). Quindi, se si aumenta la variabile, peggiora la percezione sulla salute fisica dell'individuo.

Il criterio step AIC utilizza una procedura a passi per selezionare queste variabili del modello. Inizialmente, il modello utilizza tutte le variabili disponibili. Successivamente, seleziona una variabile da rimuovere dal modello, esegue nuovamente l'addestramento del modello senza la variabile selezionata, e confronta il criterio AIC del nuovo modello con quello del modello precedente. Se il criterio AIC del nuovo modello è inferiore a quello del modello precedente, la variabile selezionata viene rimossa dal modello e il processo viene ripetuto. Se, invece, quello del nuovo modello è maggiore o uguale a quello del modello precedente, la procedura di selezione delle variabili si ferma e il modello selezionato viene considerato completo.

Questo criterio è utile per selezionare il modello più semplice e parsimonioso possibile che spieghi, però, al meglio i dati. Il suo utilizzo per selezionare le variabili del modello aiuta ad evitare il problema dell'overfitting, ovvero quando il modello è troppo complesso e si adatta troppo bene ai dati di allenamento, ma non è in grado di generalizzare bene su nuovi dati.

La formula del criterio step AIC per selezionare il modello migliore è la seguente:

$$AIC = -2\log(L) + 2k \quad (4.2)$$

dove  $L$  è la funzione di verosimiglianza del modello e  $k$  è il numero di parametri del modello. La funzione di verosimiglianza  $L$  descrive quanto bene il modello rappresenta i dati osservati, e il numero di parametri  $k$  rappresenta la complessità del modello.

In sintesi, la formula del criterio step AIC utilizza il valore AIC per

selezionare il modello migliore, che tiene conto della bontà del modello (tramite la funzione di verosimiglianza) e della complessità del modello (tramite il numero di parametri  $k$ ).

Viene riportato in Figura 4.2 l'output del modello con il valore di AIC più basso (-13452), ovvero quello che meglio bilancia la bontà di adattamento e la complessità del modello.

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.34167 -0.11227 -0.00800  0.07681  2.24549

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.109131   0.047498   2.298 0.021600 *
mh007_2      0.051970   0.011996   4.332 1.49e-05 ***
mh010_5      0.079374   0.012718   6.241 4.47e-10 ***
mh011_2      0.047001   0.018379   2.557 0.010556 *
mh011_3      0.057081   0.043422   1.315 0.188680
SessoFemmina 0.037646   0.011297   3.332 0.000863 ***
ph004_5      -0.016887   0.011927  -1.416 0.156846
cah002_2      0.106804   0.035108   3.042 0.002353 **
cah002_3      0.024438   0.030913   0.791 0.429222
cah020_5     -0.137201   0.012408 -11.057 < 2e-16 ***
camh007_2    -0.079093   0.012908  -6.127 9.17e-10 ***
cas020_5     -0.043495   0.011318  -3.843 0.000122 ***
partNo       -0.021689   0.011131  -1.948 0.051377 .
studioMedia  0.019668   0.012476   1.576 0.114948
studioAlta   0.034959   0.013522   2.585 0.009738 **
ph           -0.028181   0.006323  -4.457 8.39e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6339 on 14753 degrees of freedom
Multiple R-squared:  0.02278,    Adjusted R-squared:  0.02179
F-statistic: 22.93 on 15 and 14753 DF,  p-value: < 2.2e-16

```

**Figura 4.2:** Output del modello di regressione lineare multipla con selezione delle variabili utilizzando il criterio dello step AIC.

Il software restituisce un modello di regressione lineare multipla con selezione delle variabili attraverso l'algoritmo di selezione stepwise backward, che parte con tutte le variabili del modello e gradualmente elimina le variabili meno significative fino a raggiungere il miglior modello possibile. Il modello contiene 15 predittori e una variabile dipendente, che è la variazione di solitudine divisa in 5 modalità.

Il valore finale di  $R^2$  (**Multiple R-squared**) è 0.02278, il che indica che il modello spiega solo una piccola parte della variazione nella variabile dipendente. Inoltre, alcuni coefficienti hanno un valore p significativamente inferiore a 0.05, come *mh007\_2*, *mh010\_5*, *SessoFemmina*, *cah002\_2*, *cah020\_5*, *camh007\_2*, *cas020\_5*, *studioAlta* e *ph*, il che significa che queste variabili indipendenti hanno una relazione statisticamente significativa con la variazione di solitudine.

Tuttavia, la varianza residua (**Residual standard error**) è 0.6339, il che suggerisce che il modello non si adatta bene ai dati.

#### **4.0.4 La creazione di altre due variabili dipendenti per un'analisi più approfondita**

Per come è stata costruita la variabile dipendente non risulta essere molto adatta ai modelli (anche dalle analisi grafiche fatte precedentemente era evidente la gran numerosità sullo 0). E, nonostante ci siano alcune variabili statisticamente significative, alcune non sono di facile interpretazione.

Si è quindi deciso di prendere in considerazione la creazione di altre due variabili dipendenti dummy per vedere se emergessero relazioni più chiare. Queste due nuove variabili dipendenti sono così costruite (equazioni 4.3 e 4.4):

$$Y_1 = \begin{cases} 1 & \text{se } Y \in \{1, 2\} \\ 0 & \text{altrimenti} \end{cases} \quad (4.3)$$

$$Y_2 = \begin{cases} 1 & \text{se } Y \in \{-1, -2\} \\ 0 & \text{altrimenti} \end{cases} \quad (4.4)$$

Per come sono definite queste due variabili, la prima indica il miglioramento e la seconda il peggioramento<sup>6</sup>. In particolare,  $Y_1$  (4.3) può essere interpretata come una variabile che indica se la solitudine è migliorata (assumendo, come nel resto delle analisi, che i valori 1 e 2 corrispondano a un miglioramento), mentre  $Y_2$  (4.4) può essere interpretata come una variabile che indica se la solitudine è peggiorata (assumendo che i valori -1 e -2 corrispondano a un peggioramento).

In altre parole,  $Y_1$  e  $Y_2$  forniscono una misura binaria del cambiamento nella solitudine, indicando se il cambiamento è stato positivo o negativo rispetto alla condizione iniziale.

Vengono ora ristimati i due modelli di regressione lineare multipla e riportate prima alcune statistiche (Tabella 4.2), e successivamente la Tabella 4.3 con i coefficienti stimati e i loro  $p$  - *value*.

$Y_1$	Residual standard error: 0.36 on 14737 degrees of freedom Multiple R-squared: 0.04647, Adjusted R-squared: 0.04447 F-statistic: 23.17 on 31 and 14737 DF, p-value: <2.2e-16
$Y_2$	Residual standard error: 0.343 on 14737 degrees of freedom Multiple R-squared: 0.03984, Adjusted R-squared: 0.03782 F-statistic: 19.73 on 31 and 14737 DF, p-value: <2.2e-16

**Tabella 4.2:** Alcune statistiche del modello di regressione lineare multipla per la  $Y_1$  e  $Y_2$ .

<sup>6</sup>Cioè 1 = tutti quelli che hanno avuto un peggioramento, 0= chi non peggiora cioè tutti quelli che hanno avuto un miglioramento, e chi non ha avuto alcuna variazione



Il primo modello ( $Y_1$ ) ha un errore standard residuo leggermente più alto rispetto al secondo ( $Y_2$ ) (0.36 contro 0.3428), ma ha anche un  $R^2$  multiplo e  $R^2$  aggiustato più elevati (0.04647 e 0.04447 rispettivamente contro 0.03984 e 0.03782). Questo suggerisce che il primo modello è leggermente migliore nel descrivere la variazione dei dati rispetto al secondo. Entrambi i modelli hanno un  $p$  – *value* del test F significativo ( $< 0.05$ ), quindi almeno una delle variabili indipendenti è utile per spiegare le variazioni della variabile dipendente.

In generale, rispetto ai modelli visti precedentemente, sicuramente questo presenta dei valori statisticamente migliori sotto alcuni punti di vista (errori standard residui minori,  $R^2$  e  $F$  più elevati), senza comunque raggiungere un adattamento soddisfacente del modello ai dati.

#### **Possibili cause di un basso valore del p-value in un modello statistico**

Il  $p$  – *value* generale del modello indica la significatività globale del modello, ovvero la probabilità che almeno una delle variabili esplicative abbia un effetto significativo sulla variabile dipendente. Ci sono diverse ragioni per cui potrebbe essere molto basso:

1. Forte relazione tra la variabile dipendente e le variabili esplicative: se questo succede, il modello sarà in grado di spiegare gran parte della varianza della variabile dipendente, portando ad un p-value molto basso.
2. Campione ampio: con campioni molto ampi, anche effetti di piccola entità possono raggiungere significatività statistica, portando ad un p-value basso.

3. Variabili esplicative significative: se ci sono molte variabili esplicative nel modello e almeno una di esse ha un effetto significativo sulla variabile dipendente, il modello sarà in grado di spiegare parte della varianza della variabile dipendente, portando ad un p-value basso.
4. Modello sovradimensionato: se il numero di variabili esplicative è maggiore del numero di osservazioni il modello può soffrire di overfitting e perdere la capacità di generalizzare su nuovi dati. In questo caso, il modello può avere un p-value molto basso, ma le sue previsioni potrebbero essere poco affidabili.
5. Violazione delle assunzioni del modello: se il modello non soddisfa le assunzioni per la sua costruzione, come ad esempio l'omoschedasticità degli errori o la normalità delle distribuzioni, il p-value può essere distorto e portare a conclusioni errate sulla significatività del modello.

Nel caso preso in esame va sicuramente escluso il p.to 4, in quanto questo non accade, e presi in considerazione tutti gli altri, poichè sembrano rispecchiare le analisi qui condotte.

Vengono riportati in Tabella 4.3 i modelli delle due  $Y$  costruite, dove sono presenti le stime dei coefficienti e i relativi  $p - value$  per valutare la significatività e poter fare un confronto.

Per la variabile dicotomica  $Y_1$ , il coefficiente di regressione positivo (negativo) indica un aumento (una diminuzione) della probabilità di avere un miglioramento (peggioramento) della condizione di solitudine, men-

tre per la  $Y_2$ , il coefficiente di regressione positivo (negativo) indica un aumento (diminuzione) della probabilità di avere un peggioramento (miglioramento).

	$Y_1$		$Y_2$	
	Estimate	Pr(> t )	Estimate	Pr(> t )
<i>(Intercept)</i>	0.1493188	0.000139 ***	0.142742	0.000131 ***
<i>mh007_2</i>	0.0060511	0.374870	-0.031767	1.01e-06 ***
<i>mh010_5</i>	0.0149611	0.038797 *	-0.047543	5.57e-12 ***
<i>mh011_2</i>	0.0039644	0.705089	-0.035551	0.000367 ***
<i>mh011_3</i>	0.0222865	0.366657	-0.023174	0.324273
<i>SessoFemmina</i>	0.0245929	0.000141 ***	-0.008690	0.157647
<i>ph004_5</i>	-0.0031837	0.638955	0.015101	0.019463 *
<i>cah002_2</i>	0.0620207	0.001951 **	-0.040405	0.034067 *
<i>cah002_3</i>	0.0023176	0.895546	-0.028225	0.093156 .
<i>cah003_5</i>	0.0044167	0.695240	0.012902	0.229474
<i>cah020_5</i>	-0.0916517	<2e-16 ***	0.017715	0.008482 **
<i>camh007_2</i>	-0.0425189	7.22e-09 ***	0.019197	0.006065 **
<i>cac013_5</i>	0.0166866	0.383489	0.040073	0.027982 *
<i>cas015_5</i>	0.0194852	0.150339	-0.004855	0.706665
<i>cas020_5</i>	-0.0227837	0.000616 ***	0.011032	0.081595 .
<i>partNo</i>	0.0638918	<2e-16 ***	0.079374	<2e-16 ***
<i>figli1</i>	0.0059287	0.620451	-0.027767	0.014882 *
<i>figli2</i>	0.0092986	0.393789	-0.018524	0.074457 .
<i>figli3</i>	0.0113184	0.349708	-0.018148	0.115374
<i>figli4</i>	0.0213951	0.173894	-0.021978	0.142436
<i>figli4+</i>	-0.0062775	0.738359	-0.010833	0.544989
<i>lavoro2</i>	0.0145867	0.172939	-0.019355	0.057589 .
<i>lavoro3</i>	-0.0083401	0.430849	-0.032146	0.001434 **
<i>lavoro4</i>	-0.0234394	0.032259 *	-0.053930	2.32e-07 ***
<i>anno(1935,1940]</i>	-0.0059701	0.614394	-0.021040	0.062268 .
<i>anno(1940,1945]</i>	-0.0125105	0.273166	-0.031094	0.004242 **
<i>anno(1945,1950]</i>	-0.0046689	414 0.678684	-0.030210	0.004889 **
<i>anno(1950,1955]</i>	-0.0129594	0.267356	-0.023138	0.037582 *
<i>anno(1955,1960]</i>	0.0031347	0.835437	-0.028430	0.047898 *
<i>studioMedia</i>	-0.0070943	0.327002	-0.022979	0.000858 ***
<i>studioAlta</i>	0.0007336	0.925433	-0.030750	3.81e-05 ***
<i>ph</i>	0.0064629	0.080332 .	0.027103	1.43e-14 ***

**Tabella 4.3:** Tabella delle stime dei coefficienti e p-value della regressione lineare multipla della  $Y_1$  e  $Y_2$ .

Appare sicuramente evidente la forte significatività delle variabili di  $Y_2$  rispetto quella rilevata nelle variabili della prima. Un'altra differenza che risulta importante sottolineare è l'alta significatività, presente solo in  $Y_1$ , della variabile *SessoFemmina*, che, avendo il coefficiente positivo, indica che le donne hanno una probabilità maggiore di miglioramento rispetto gli uomini.

Notando che la variabile risposta sul miglioramento non apporta particolari modifiche a livello di significatività rispetto a quella originaria, si concentrerà l'attenzione su quella che riguarda il peggioramento, cioè  $Y_2$ , al fine di enfatizzare le differenze importanti riscontrate tra quest'ultima e la  $Y$ . Ciò che è emerso di interessante ai fini di questa analisi, è<sup>7</sup>:

- Risultano altamente significative le variabili *mh007\_2*, *mh010\_5*, *mh011\_2*, (prima \* ora \*\*\*)<sup>8</sup>.
- Genere: *SessoFemmina* non è più statisticamente significativa.
- Presenza del partner: interessante la variabile *partNo*, che ora è altamente statisticamente significativa, con la stima del coefficiente che ha segno positivo, ad indicare l'alta probabilità di peggioramento della condizione di solitudine durante il Covid degli anziani senza partner. Un risultato che potrebbe sembrare ovvio, ma che solo attraverso l'analisi di questa variabile è stato possibile riscontrare.

---

<sup>7</sup>Si ricorda il significato degli asterischi :

1. '\*\*\*' indica una significatività statistica molto alta (p-valore inferiore a 0.001)
2. '\*\*' indica una significatività statistica alta (p-valore compreso tra 0.001 e 0.01)
3. '\*' indica una significatività statistica moderata (p-valore compreso tra 0.01 e 0.05)
4. '.' indica una significatività statistica bassa (p-valore compreso tra 0.05 e 0.1)
5. '' indica una non-significatività statistica (p-valore superiore a 0.1)

<sup>8</sup>Prima= nel modello di regressione lineare multipla fatto sulla  $Y$ . Ora= nel modello di regressione lineare multipla fatto sulla  $Y_2$

- Numero di figli: *figli1*, ora, ha una significatività statistica moderata con il segno negativo, ad indicare che la presenza di un figlio rispetto a non averne influiva sul peggioramento.
- Giudizio sulla solvibilità: Prima *lavoro2* era moderatamente significativo (\*) ora non lo è, mentre *lavoro3* e *lavoro4* sono altamente significative con segno negativo (3= Abbastanza facilmente, 4= Facilmente), a segnalare che chi giudicava di arrivare a fine mese senza problemi, riscontrava un miglioramento della sensazione di solitudine.
- Anno di nascita: anche qui un risultato che potrebbe risultare interessante, gli intervalli d'età dal 1940 (escluso) al 1960 diventano statisticamente significativi (nel modello con la *Y* nessuno lo era), con il segno delle stime dei coefficienti che indicano un miglioramento di questa fascia d'età rispetto a quella del 1915-1935.
- Titolo di studio: *studioMedia* non era significativa e *studioAlta* lo era poco; ora risultano entrambe altamente significative con segni negativi per mettere in luce come un'alta istruzione possa modificare positivamente la condizione di solitudine percepita, rispetto agli anziani che hanno un titolo di studio basso.
- Salute fisica percepita: anche in questo caso la variabile *ph*, come nel modello per la *Y* è altamente significativa. Un'analogia che prosegue anche nel segno del coefficiente (positivo) che indica che più si percepisce una salute fisica precaria, più la solitudine risulta peggiorata durante la prima pandemia di Covid-19.

# Capitolo 5

## Conclusioni

Le analisi descrittive svolte in questa trattazione, attraverso i dati di SHARE, dimostrano, innanzitutto, che la maggior parte del campione considerato non ha subito variazioni importanti, concentrando quasi tutta la distribuzione attorno lo zero. Anche questo è sicuramente un risultato importante, poichè evidenzia che i soggetti non hanno subito forti cambiamenti a causa del Covid-19. Possiamo attribuire questo esito a diversi possibili fattori:

- **Mantenimento delle abitudini sociali:** molte persone potrebbero aver mantenuto le loro abitudini sociali durante la pandemia, anche se sono state costrette a fare questo in modo virtuale o a distanza. Ciò potrebbe significare che le persone che erano già solitarie prima della pandemia potrebbero non aver avuto molte occasioni per costruire nuove relazioni o mantenere quelle esistenti.
- **Riduzione dell'opportunità sociale:** il Covid-19 ha portato a restrizioni sociali e di viaggio che hanno limitato l'opportunità di incontrare nuove persone o di partecipare ad attività sociali. Ciò potrebbe aver limitato l'opportunità per le persone solitarie di incontrare

nuove persone o di partecipare ad attività sociali che avrebbero potuto aiutarli a superare la solitudine.

- Incremento dei livelli di stress: la pandemia ha portato a un aumento dei livelli di stress e di ansia in molte persone. Ciò potrebbe aver reso più difficile per le persone solitarie affrontare la loro condizione e potrebbero aver ridotto la motivazione a cercare di cambiare la situazione.
- Incremento dei problemi di salute mentale: la pandemia ha portato a un aumento dei problemi di salute mentale, tra cui ansia, depressione e stress post-traumatico. Ciò potrebbe aver reso più difficile per le persone solitarie cercare aiuto per affrontare la loro condizione.
- Mancanza di supporto: molte persone potrebbero non aver ricevuto il supporto adeguato dalle loro reti di supporto esistenti durante la pandemia, a causa delle restrizioni sociali o della situazione economica difficile. Ciò potrebbe aver fatto sentire alle persone solitarie ancora più isolate e potrebbe aver limitato la loro capacità di uscire dalla solitudine.

È importante ricordare che gli individui che potrebbero essere classificati come gli anziani più “a rischio” cioè senza partner, senza figli, le donne, chi possedeva un titolo di studio più basso e chi riteneva di non riuscire ad arrivare a fine mese facilmente, hanno riportato maggiori variazioni positive e negative. Avere maggiori variazioni significa che le persone in questione hanno sperimentato fluttuazioni più ampie nel loro benessere mentale e fisico rispetto ad altre persone durante la pandemia. Ad

esempio, se consideriamo l'esperienza di solitudine, una persona che ha riportato maggiori variazioni potrebbe aver sperimentato periodi in cui si sentiva molto sola e periodi in cui si sentiva meno sola, mentre una persona che ha riportato variazioni minori potrebbe aver sperimentato una solitudine costante o quasi costante durante la pandemia.

A riguardo dei modelli di regressione lineare analizzati, un primo risultato che emerge è un adattamento tendenzialmente povero, per cui questi modelli non sono in grado di spiegare o prevedere correttamente la variabilità nei dati osservati. Non essendo verificate le assunzioni del modello di regressione lineare multipla, questo non risulta particolarmente utile. Potrebbe essere necessario apportare modifiche al modello, oppure utilizzare un modello completamente diverso.

In ogni caso, i risultati emersi e riportati confermano alcune conclusioni delle analisi descrittive sui soggetti più a rischio del campione. Per esempio, il miglioramento della condizione avuta da chi possedeva un titolo di studio più alto. Anche sulla base di alcune domande simili poste prima e durante la prima ondata di Covid-19, si possono notare alcune differenze, che evidenziano che alcuni individui hanno rilevato dei miglioramenti durante la pandemia nonostante presentassero maggiori problemi (insonnia, salute fisica e altro), facendo emergere interpretazioni a favore del possibile aumento degli aiuti ricevuti durante quel periodo complesso.

Riassunto, l'importanza degli interventi di supporto per il benessere mentale degli anziani è indiscutibile ai fini di un invecchiamento sano. Al



fine di analizzare l'incidenza della pandemia e delle misure di controllo dell'epidemia sulla salute mentale degli anziani, sono necessarie ulteriori ricerche per ottenere risultati più attendibili. Si auspica, inoltre, lo sviluppo di interventi innovativi basati sulla tecnologia per migliorare la connessione sociale di questa popolazione. Le tecnologie mobili possono rappresentare un'importante risorsa in quanto hanno rivoluzionato i nostri modi di interagire, accedere alle informazioni, alle risorse e ai servizi. Infine, il sostegno tra pari offerto dai social media può migliorare gli effetti del supporto professionale basato sull'evidenza, quali le risorse informative, la promozione della salute, la consulenza e la risoluzione dei problemi.

# Bibliografia

- Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**(6), 716–723.
- Atzendorf, Josefine, & Grube, Stefan. 2021. *Depression and Loneliness during the Pandemic — share-eric.eu*. <http://www.share-project.org/share-publications/share-working-paper-series.html>.
- Burnham, Kenneth P, & Anderson, David R. 2004. *Multimodel inference: Understanding AIC and BIC in model selection*. Springer Science & Business Media.
- Gombas, Lauren. 2022. *The Difference Between Solitude and Loneliness Post*. <https://www.coachtrainingedu.com/blog/solitude-vs-loneliness>.
- Grigoletto, Matteo, Ventura, Laura, & Pauli, Francesco. 2017. *Modello lineare: teoria e applicazioni con R*. G Giappichelli Editore.
- Heinrich, Liane M, & Gullone, Eleonora. 2006. Social isolation, social support, and health. *In Handbook of adolescent health psychology*, 31–44.

- Perlman, Daniel, Peplau, Letitia Anne, & Goldston, SE. 1984. Loneliness research: A survey of empirical findings. *Preventing the harmful consequences of severe and persistent loneliness*.
- Share. 2004a. *Infrastructure* — *share-eric.eu*. <https://share-eric.eu/infrastructure>.
- Share. 2004b. *Questionnaires* — *share-eric.eu*. <https://share-eric.eu/data/data-documentation/questionnaires>.
- Weiss, Robert S. 1973. Loneliness: The experience of emotional and social isolation. *MIT press*.
- World Health Organization. 2020. Mental health and psychosocial considerations during the COVID-19 outbreak. *World Health Organization*.