

UNIVERSITÀ DEGLI STUDI DI GENOVA

**SCUOLA DI SCIENZE SOCIALI
DIPARTIMENTO DI ECONOMIA**

*Corso di laurea in Economia e Management Marittimo e
Portuale*



Elaborato scritto per la prova finale in
Analisi di mercato per il trasporto marittimo di persone

Strategie di controllo dell'errore nelle indagini di mercato e nella
redazione di un questionario

Docente di riferimento:
Luca Persico

Candidato:
Lorenzo Frondana

Anno accademico 2022-2023

Sommario

Ringraziamenti	4
Abstract.....	5
Abstract.....	6
CAPITOLO I. Il questionario	7
1.1. Storia della ricerca attraverso i questionari.	7
1.2. Aspetti definatori e peculiarità.	9
1.3 Domande e aspetti cognitivi delle risposte.	12
1.4 Metodi di somministrazione: criticità e tecniche di intervista.....	24
1.4.1 Le interviste postali	24
1.4.2 Interviste online	28
1.4.3 Interviste telefoniche	32
1.4.4 Interviste personali	35
CAPITOLO II. Le Strategie di campionamento.....	38
2.1 L'inferenza statistica e il campionamento	38
2.2 Il campionamento probabilistico	43
2.3 Il campionamento non probabilistico	48
2.3.1 Tecniche di campionamento non probabilistico: vantaggi e svantaggi	51
CAPITOLO III. Gli errori	57
3.1 L'errore complessivo del questionario.....	57
3.2 Errori Campionari.....	62
3.3 Errori non campionari.....	67
CAPITOLO IV. Analisi dei dati.....	77
4.1 Metodi di analisi dati	77
4.2 Affidabilità e validità dei dati raccolti.	81
4.3 Le scale di misurazione e la <i>cluster analysis</i>	86

4.3.1 La scala Likert.	91
4.3.2 Data <i>cleaning</i>	93
Conclusioni.	98
Bibliografia.	100

Ringraziamenti

Un ringraziamento sentito va al relatore, il professor Luca Persico, per i suggerimenti e le indicazioni che mi ha dato e per la disponibilità dimostratami durante la stesura dell'elaborato e durante il corso delle lezioni.

Ringrazio i miei genitori, mia sorella e tutta la mia famiglia per avermi concesso l'opportunità di studiare in tutti questi anni e per avermi spronato a fare sempre meglio.

Ringrazio Chiara, compagna inseparabile di vita e di studio. Grazie per esserci sempre stata.

Ringrazio gli amici di sempre, per non aver mai fatto mancare la loro presenza in tutti questi anni.

Abstract

Attualmente le indagini di mercato rivestono un ruolo di estrema rilevanza per quanto riguarda i processi decisionali delle aziende, dal momento che possono fornire numerose ed importanti informazioni circa le preferenze dei consumatori e le tendenze del mercato, e uno degli strumenti più comunemente impiegati per raccogliere questa tipologia di informazioni è il questionario. Tuttavia, per garantire l'affidabilità e la validità dei dati raccolti, diventa fondamentale per chi conduce l'indagine adottare alcune accortezze nelle fasi di redazione e di conduzione di un sondaggio e durante l'analisi dei dati, al fine di non assumere decisioni sulla base di dati viziati da errori. Pertanto, il presente elaborato di tesi ha l'obiettivo di individuare alcune delle principali strategie di controllo dell'errore nell'ambito della redazione di un questionario e nelle indagini di mercato. Questa disamina, che è stata condotta revisionando in maniera attenta la letteratura scientifica relativa alle tematiche trattate e facendo ricorso ad alcuni esempi pratici, si è concentrata dapprima sulle tecniche di costruzione di un questionario e sulle strategie di campionamento, per poi spostarsi sulla trattazione degli errori e, infine, sull'analisi dei dati, dove è stata proposta una tecnica per eliminare i dati errati raccolti da un questionario.

Abstract

Actually, market surveys play an extremely important role in the decision-making processes of companies, as they can provide a lot of important information about consumer preferences and market trends, and one of the most commonly used tools to collect this information is the questionnaire. However, to ensure the reliability and validity of the collected data, it becomes crucial for the researcher to take certain precautions when drafting and conducting a survey and when analysing the data, in order not to make decisions based on flawed data. Therefore, the aim of this dissertation is to identify some of the main error control strategies in questionnaire writing and market surveys. This examination, which was carried out by carefully reviewing the scientific literature on the topics discussed and by resorting to some practical examples, first focused on the techniques of constructing a questionnaire and sampling strategies, and then moved on to the treatment of errors and, finally, to data analysis, where a technique was proposed for eliminating erroneous data collected from a questionnaire.

CAPITOLO I. Il questionario

1.1. Storia della ricerca attraverso i questionari.

La raccolta dati sugli individui, seppure con forme differenti da quelle attuali, ha una lunga storia. Sebbene ad oggi siano rimasti pochi frammenti, si può affermare che il primo esempio di *data gathering* sia rappresentato da un censimento condotto in Cina più di 4000 anni fa, mentre sono pervenute ai giorni nostri informazioni più dettagliate circa i censimenti condotti nell'Antico Egitto, durante il cosiddetto Nuovo Regno (1550-1070 a.c), e in Cina, durante la dinastia (206 a.c- 220 d.c.), realizzati in entrambi i casi per raccogliere informazioni sulla popolazione e per scopi fiscali e amministrativi. Ben più noti sono gli esempi riportati nella Bibbia, in cui dai censimenti condotti dai romani emerge anche una finalità militare. In Europa, una delle fonti più autorevoli è il "Domesday Book" (1086), commissionato da Guglielmo il Conquistatore, che aveva il compito di documentare le proprietà terriere e il bestiame della maggior parte della popolazione inglese nel medioevo. Il più antico censimento nominale completo documentato – cioè contenente i nomi dei membri della popolazione – è stato realizzato in Islanda nel 1703, dove sono stati raccolti i seguenti dati come nome, età e posizione nel nucleo familiare, oltre a informazioni sulla salute. A partire dal XIX secolo d.c. si iniziano ad avere evidenze documentali che certificano il lavoro di raccolta dati da parte di paesi come Francia, Stati Uniti, Svezia, Norvegia, Germania e Italia. Il primo censimento ufficiale in Italia ebbe luogo alla conclusione del processo di unificazione nazionale nel 1861. Questo aveva lo scopo di raccogliere dati demografici, sociali ed economici per delineare una panoramica precisa dello stato appena costituito.

È proprio a partire dall'Ottocento che cominciano ad essere documentate le prime forme di questionario, che venivano condotte da privati cittadini e filantropi. Tuttavia, l'indagine veniva interpretata con un concetto di fondo differente rispetto a come avviene attualmente, come dimostra il caso di Charles Booth tra il XIX e il XX secolo, il quale si interessò alle tematiche legate alla povertà e alla disoccupazione che in quel momento storico affliggevano la città di Londra e non trovando risposte adeguate dagli interlocutori politici, decise di mettere insieme le informazioni a sua disposizione. Si impegnò in dibattiti e discussioni con politici, socialisti e assistenti sociali che non erano in grado di rispondere adeguatamente alla sua preoccupazione principale: "come

vivevano esattamente i poveri, quanto erano scontenti, come potevano essere aiutati concretamente” (Fried e Elman, 1968: xvi). Mettendo insieme in modo creativo le informazioni disponibili, come i dati dei censimenti, e attingendo a una pluralità di approcci per la raccolta dei dati, Booth si propose di rispondere a questa domanda e concentrò la sua attenzione inizialmente sull’East End di Londra.

1.2. Aspetti definatori e peculiarità.

In primo luogo, occorre fornire una definizione di questionario. Un questionario è uno strumento di raccolta delle informazioni composto da un insieme strutturato di domande e relative categorie di risposta definite a priori da chi lo costruisce, ovvero di domande “chiuse”, dove all’intervistato viene richiesto di individuare tra le risposte presentate quella che più si avvicina alla propria posizione, e/o domande “aperte” che non prevedono delle risposte predeterminate (Zammuner, 1996). Tuttavia, in letteratura esistono anche altre definizioni, che pongono l’attenzione su elementi fondamentali concernenti il questionario. Pertanto, si può definire il questionario anche come un metodo sistematico per raccogliere informazioni da un campione di entità allo scopo di costruire descrittori quantitativi degli attributi della popolazione più ampia di cui le entità fanno parte (Groves et al. 2009). Quest’ultimo approccio, a differenza del primo, si concentra su due aspetti, ovvero sui descrittori quantitativi e sul campione. I sopracitati descrittori quantitativi sono rappresentati dalle statistiche, ovvero sintesi quantitative di osservazioni raccolte su un campione. Le statistiche possono essere definite descrittive, quando rappresentano le dimensioni e le distribuzioni di differenti attributi in una popolazione, oppure analitiche, quando vanno ad osservare come due o più variabili sono legate tra loro. Mentre per quanto riguarda il campione, si tratta di un sottoinsieme rappresentativo di una popolazione che è stato selezionato per partecipare ad uno studio statistico. La raccolta e l’analisi dei dati da un campione sono utilizzate per fare inferenza sulla popolazione di origine (Snedecor e Cochran, 1989).

Quindi, da queste prime definizioni si può capire come il questionario sia uno strumento di raccolta dei dati basato su una serie di domande che vengono poste ad alcune persone, che formano un campione, con lo scopo di ottenere informazioni su un certo argomento. Occorre fornire una breve descrizione delle peculiarità di un questionario.

In primis, esistono due tipologie di questionario: strutturato e non strutturato. Il questionario strutturato consiste in una serie di domande poste al medesimo modo a tutti gli intervistati, mentre quello non strutturato si concretizza in una conversazione a domande aperte. Si osserva, dunque, la struttura delle domande. Le domande devono essere concise e senza ambiguità e possono essere aperte, concedendo all’intervistato la possibilità di rispondere liberamente, oppure possono essere chiuse, prevedendo risposte

predefinite. Quindi, le domande del questionario devono essere espresse in un linguaggio semplice, grammaticalmente corrette, specifiche e volte ad ottenere informazioni ben definite. Inoltre, devono poter essere interpretate senza difficoltà da tutti e non devono fare affidamento sulla memoria dell'intervistato. Infine, devono garantire l'anonimato e la riservatezza del dato. Diventa fondamentale, quindi, ottenere risposte corrette alle domande e per realizzare tale obiettivo è necessario che gli intervistati debbano avere la capacità e la volontà di fornire l'informazione ed essere in grado di comprendere la domanda. Oltre a questo aspetto, è fondamentale che gli autori del questionario abbiano chiaro l'obiettivo per cui si realizza il questionario. Continuando a fare riferimento alla fase di formulazione delle domande, è necessario evitare di inserire un numero troppo elevato di alternative tra le risposte possibili, al fine di non creare confusione e di non rendere difficile la valutazione delle opzioni da parte del rispondente. Un altro aspetto da evitare è quello legato alla possibilità che ha l'intervistatore di influenzare la risposta, che si può concretizzare con l'inserimento di frasi all'interno dei quesiti che facciano sentire la pressione sociale sull'intervistato. Inoltre, una formulazione non corretta delle domande può condurre i rispondenti ad abbandonare il questionario prima di completarlo e a rispondere in maniera casuale alle domande. Oltre alla struttura delle domande, occorre soffermarsi anche sul loro ordine. L'ordine secondo il quale i quesiti vengono posti all'interno del questionario è un aspetto da tenere in considerazione in quanto può influenzare le risposte degli intervistati, per cui si rende necessario adattare l'ordine delle domande agli obiettivi della ricerca per ottenere risposte attendibili.

Sebbene sul campionamento e sulle relative modalità di realizzazione l'elaborato si concentrerà in seguito, è necessario introdurre brevemente un'ulteriore peculiarità del questionario, ovvero quella legata al campione. Il campione deve rappresentare la popolazione di interesse per la ricerca che si intende effettuare ricorrendo allo strumento del questionario. Infatti, un questionario viene redatto per essere somministrato ad uno specifico gruppo di persone, per cui la sua validità e la sua utilità dipende dalla sua capacità di raggiungere il target a cui si rivolge. Utilizzare un campione che differisca dalla popolazione di riferimento può condurre alla raccolta di risultati distorti dalla scarsa rappresentatività.

Assumono una particolare rilevanza la scala di misurazione e l'ordine delle domande. La scala di misurazione rappresenta le modalità con la quale vengono raccolte e quantificate le risposte degli intervistati ed influisce in maniera diretta sulla qualità e sulla precisione dei dati raccolti. Dal momento che esistono differenti tipologie di scala di misurazione, tra le quali si possono citare la scala nominale, la scala ordinale e la scala Likert, assume un ruolo fondamentale la scelta di una scala di misurazione opportuna a seconda della tipologia di ricerca che si intende condurre. Se la scala di misurazione è ben strutturata restituisce una misura precisa e puntuale delle opinioni e dei comportamenti dei rispondenti, perché semplifica l'interpretazione delle risposte per gli intervistati e, per coloro che effettuano la ricerca, riduce possibili fraintendimenti e ambiguità nelle risposte. La scala di valutazione, assegnando valori numerici ai dati raccolti, riduce la variabilità delle risposte, restringendo il *range* di risposte possibili, e permette di aggregare con maggior facilità le risposte fornite dai rispondenti.

In genere, i questionari costituiscono una modalità efficiente ed economica per raccogliere dati personalizzati su larga scala. Pertanto, bisogna considerare gli ambiti di applicazione del questionario e comprendere le motivazioni che giustificano il ricorso a tale strumento. Tra gli ambiti di applicazione più comuni dei questionari si includono la ricerca scientifica, allo scopo di raccogliere dati qualitativi e quantitativi, le indagini di mercato, per conoscere le preferenze dei consumatori, le loro abitudini di acquisto e altre informazioni di mercato, le risorse umane, al fine di selezionare il personale o di valutarne le prestazioni, l'istruzione, per valutare programmi accademici o per raccogliere i feedback degli studenti e, infine, le ricerche di opinione, allo scopo di conoscere l'opinione pubblica su questioni politiche, sociali ed economiche.

1.3 Domande e aspetti cognitivi delle risposte.

In letteratura esistono differenti approcci per quanto riguarda la stesura delle domande, in quanto la costruzione di domande precise e di facile comprensione, che consentano all'intervistato di rispondere in modo corretto, è uno dei punti fondamentali nella redazione di un questionario. La costruzione e la selezione dei quesiti dipendono dal contesto in cui il questionario è inserito, ovvero gli scopi dell'indagine, le caratteristiche dell'intervistatore e del rispondente, come le domande vengono poste e le peculiarità delle risposte. Per queste ragioni, si può affermare come prima di cominciare la stesura delle domande del questionario, è necessario seguire una serie di passaggi.

Il primo di questi *steps* concerne la comprensione approfondita dello scopo del sondaggio, tenendo in considerazione anche le esigenze del *team* che conduce la ricerca e la natura dei soggetti che devono rispondere ai quesiti. La finalità del questionario è rappresentata dal risultato che quest'ultimo spera di ottenere, normalmente gli intervistatori hanno in mente scopi generali, quali la misurazione della soddisfazione dei clienti che usufruiscono di un determinato prodotto o servizio oppure la classificazione dei clienti in base ai loro comportamenti d'acquisto.

Dopodiché occorre chiarire i termini utilizzati per indicare il fine ultimo dell'indagine, in quanto, talvolta, i termini impiegati possono essere generici e comprendere significati non inerenti all'indagine che si sta svolgendo. Ciò può comportare ambiguità di contenuto che possono condurre l'intervistato a fornire una risposta differente rispetto a quella che avrebbe potuto dare se la domanda non avesse incluso termini equivocabili. Può essere utile inserire alcune definizioni per facilitare la comprensione dell'accezione che si intende attribuire ad uno specifico termine. A tal proposito, la scelta della definizione può ricadere su argomenti già noti e pubblicati oppure possono essere frutto di un'elaborazione propria del team, o del singolo soggetto, che porta avanti l'indagine. Tuttavia, è opportuno considerare che qualora la scelta dovesse ricadere su quest'ultima fattispecie si potrebbe incorrere nel rischio di non riuscire a convincere i lettori della validità delle definizioni, ragion per cui, quando è possibile, si fa ricorso a definizioni già rispettate e testate. In genere, i ricercatori tendono a risolvere il problema dando agli intervistati una definizione da utilizzare e chiedendo loro di fare il lavoro di classificazione.

Il passaggio successivo consiste nella definizione degli obiettivi specifici dell'indagine, ovvero delle finalità precise che si intendono raggiungere con le informazioni raccolte. Gli obiettivi possono essere espressi sotto forma di ipotesi di ricerca, di domande oppure di affermazioni e possono riguardare differenti aspetti, come l'identificazione delle esigenze dei clienti, la determinazione delle differenze che sussistono nelle esigenze tra le differenti tipologie di clienti e la determinazione dei servizi più importanti. Le domande, pertanto, dovranno essere specifiche e coerenti con gli obiettivi che l'intervistatore intende perseguire. La differenza tra l'obiettivo di una domanda e la domanda stessa è una distinzione fondamentale, in quanto l'obiettivo definisce il tipo di informazioni necessarie, mentre la progettazione della domanda o delle domande specifiche per raggiungere l'obiettivo è un passo completamente diverso. In sede di redazione del questionario, sarà necessario porre particolare attenzione al processo che porta dall'obiettivo a comporre una domanda, le cui risposte consentiranno di raggiungere lo scopo prefissato. È buona norma produrre un elenco dettagliato degli obiettivi delle domande e un piano di analisi che delinei le modalità di utilizzo dei dati prima di progettare lo strumento di indagine.

Conoscere il rispondente è un'altra fase del processo corretto di costruzione dei quesiti. Le domande del sondaggio devono essere scritte in modo da soddisfare le esigenze di dati del gruppo di ricerca, ma devono anche essere formulate in modo che gli intervistati possano rispondere facilmente e con precisione. È necessario verificare l'adeguatezza del livello di lettura delle domande proposte, poiché le capacità di lettura degli intervistati possono variare notevolmente, anche per lo stesso sondaggio. Inoltre, è opportuno ricordare che il questionario è uno strumento che è in grado di raggiungere una moltitudine di persone, ognuna delle quali in possesso di differenti capacità interpretative e comprensive, per cui si possono trovare soggetti che, per esempio, sanno leggere molto bene e altri che, invece, leggono male oppure si può incorrere in persone che hanno gradi di istruzione differenti. Sebbene nella pratica sia molto difficile riuscire a scrivere una domanda che sia comprensibile da tutti allo stesso modo, esistono alcuni aspetti che aumentano la probabilità di incomprensione dei quesiti da parte del lettore. Infatti, le domande che includono gergo, termini non familiari o tecnici o frasi astratte hanno maggiori probabilità di essere comprese in modo incoerente, per cui la scelta dei vocaboli nella stesura di una domanda deve tenere conto delle caratteristiche della

popolazione campionata, la quale, ad esempio, potrebbe non avere le competenze tecniche per capire pienamente le domande che gli vengono poste. Quando una domanda include parole o frasi che un rispondente non è in grado di definire, gli intervistati possono cercare di indovinare cosa chiede la domanda e rispondere comunque, possono saltare la domanda e non rispondere affatto, oppure possono scegliere una risposta a caso. Tutte queste opzioni sono dannose per l'affidabilità dei dati. Al fine di evitare tali situazioni, è necessario che i ricercatori si assicurino che il campione della popolazione tenuto in considerazione abbia conoscenze sufficienti per rispondere alle domande poste. Inoltre, sarà responsabilità dell'intervistatore fornire al rispondente tutti gli elementi necessari, come esempi o definizioni, perché possa comprendere adeguatamente. Oltre a comprendere il vocabolario utilizzato in una domanda, è anche importante che i rispondenti capiscano per quale periodo di tempo devono rispondere. Un altro approccio che è possibile adottare quando è difficile osservare un fenomeno attraverso una singola domanda, è quello che prevede l'introduzione di domande supplementari per coprire tutti quegli eventi che vengono omessi. L'uso di domande multiple per coprire tutti gli aspetti di ciò che deve essere segnalato, piuttosto che cercare di racchiudere tutto in un'unica definizione, è uno dei modi più semplici per assicurarsi di non perdere nessuna informazione legata agli aspetti di interesse ai fini dell'indagine. Per queste ragioni, le domande dovranno essere formulate in modo da massimizzare il numero di intervistati capaci di leggere e comprendere correttamente le domande.

Il ricercatore deve effettuare un attento confronto tra ciò che serve sapere e il tempo che hanno a disposizione per ottenere l'informazione. Il numero di domande che si possono includere in un sondaggio dipende in larga misura dal tempo a disposizione degli intervistati per rispondere alle domande, da ciò che si vuole sapere e dal numero di domande necessarie per ottenere una misurazione adeguata.

L'ultimo *step* consiste nella standardizzazione dell'intervistatore. Con tale termine si intende che l'intervistatore pone tutte le domande allo stesso modo in ogni circostanza. Quando due intervistatori che conducono la stessa indagine sono indistinguibili l'uno dall'altro per quanto riguarda la somministrazione e i risultati, quindi sono standardizzati. Gli intervistatori standardizzati possono essere umani, come nelle interviste faccia a faccia o telefoniche, oppure possono assumere la forma di questionari

autosomministrati. Un questionario autosomministrato può essere spedito ai rispondenti o compilato in un'area appositamente designata, come la sala d'attesa di una clinica, un'aula o un ufficio del personale.

I procedimenti sopra analizzati, se svolti tutti correttamente, permettono ai ricercatori di cominciare a stilare le domande in modo opportuno. È necessario ricordare che le domande, e le conseguenti risposte, sono lo strumento fondamentale per misurare adeguatamente il fenomeno oggetto di studio. Per questo, in primo luogo, occorre che il *team* che sta svolgendo l'indagine si interroghi su quali siano le domande corrette da inserire all'interno del sondaggio. Si può definire una domanda corretta quando produce risposte che sono misure affidabili e valide di qualcosa che si intende descrivere. L'affidabilità è usata qui nel senso classico della psicomètria, ovvero della misura in cui le risposte sono coerenti: Quando lo stato di ciò che viene descritto è coerente, anche le risposte lo sono (Nunnally, 1978). La validità, a sua volta, è la misura in cui le risposte corrispondono a un ipotetico "valore vero" di ciò che si sta cercando di descrivere o misurare (Cronbach & Meehl, 1955)¹. Inoltre, una corretta progettazione delle domande significa assicurarsi che il ricercatore e tutti gli intervistati utilizzino le stesse definizioni e gli stessi termini quando classificano le persone o conteggiano gli eventi.

Tuttavia, si rende necessario approfondire con un maggior grado di dettaglio le caratteristiche vere e proprie che devono possedere i quesiti. In prima istanza, bisogna porre domande mirate e concrete. Le domande mirate sono quelle logicamente collegate agli obiettivi del sondaggio, per cui si fa riferimento ai quesiti riguardanti, ad esempio, la qualità del servizio, la puntualità del servizio, la qualità del cibo e così via. Mentre una domanda concreta è precisa e non ambigua. Nello specifico, le domande possono essere definite precise e non ambigue quando, senza essere sollecitati, due o più potenziali rispondenti concordano sulle parole usate nella domanda. Solitamente, per rendere ancor più concreta una domanda si aggiunge un periodo di tempo che possa indirizzare gli intervistati a fornire una risposta maggiormente precisa. Pertanto, più aumentano il livello di dettaglio e la concretezza delle domande più sarà probabile ottenere risposte più affidabili ed attendibili. Un altro aspetto da considerare quando si parla di domande formulate in misura corretta è legato all'impiego di periodi di tempo

¹ La valutazione della validità e gli aspetti legati alla valutazione dei dati ottenuti dal questionario verranno approfonditi nel corso della trattazione.

collegati all'importanza del quesito. Ragion per cui, dovranno essere utilizzati intervalli temporali più lunghi, come uno o più anni, per gli interrogativi attinenti ad eventi particolarmente importanti nella vita di un intervistato, come possono essere l'acquisto di una casa o la nascita di un figlio. Mentre, per eventi che hanno relativamente minore importanza si ritiene una soluzione migliore inserire archi temporali più brevi, in quanto periodi più lunghi potrebbero indurre la persona a sforzarsi di ricordare e ad effettuare supposizioni in merito quel determinato episodio che non risultino aderenti alla realtà della propria esperienza. Bisogna comunque prestare attenzione alla ristrettezza dell'intervallo temporale considerato dal momento che un periodo troppo breve rischia di riferirsi ad eventi che non si sono ancora verificati. Come accennato in precedenza, è fondamentale adottare un linguaggio convenzionale. In particolare, le domande di un questionario si fondano sulla grammatica, sulla punteggiatura e sull'ortografia basilari, ciò avviene con la finalità di massimizzare la comprensione di tutti coloro che partecipano al sondaggio. Un primo passo che va nella direzione di un utilizzo convenzionale del linguaggio è rappresentato dall'impiego di frasi complete, sottoforma sia di affermazioni sia di domande, poiché esprimono pensieri chiari e completi. Inoltre, si devono evitare abbreviazioni e acronimi nei quesiti, ad eccezione per gli acronimi più noti di cui si è certi che vengano comunemente compresi, come le sigle dei paesi o delle città. Le domande, inoltre, non devono contenere slang, espressioni gergali, colloquiali e tecniche, in quanto possono non essere conosciute da tutti i componenti del campione². In aggiunta a quanto appena detto, si sottolineano potenziali problematiche legate all'uso delle suddette espressioni qualora si prevedesse di riferire i risultati del sondaggio a un pubblico generale, il che renderebbe necessario tradurre il gergo, con il rischio che una traduzione meno esperta potrebbe comportare una perdita di significato. Bisogna evitare parole e frasi distorte. Le parole e le frasi distorte suscitano reazioni emotive che possono avere poco a che fare con le questioni affrontate dall'indagine. Vengono considerate tendenziose perché scatenano reazioni emotive o pregiudizi. Il sondaggio può essere influenzato da pregiudizi se non si comprendono appieno la cultura e i valori degli intervistati e si pongono domande inavvertitamente offensive. Per evitare questa eventualità, è necessario che tutte le domande vengano riviste e testate,

² Esistono tuttavia alcune eccezioni, come un'indagine su un gruppo omogeneo che condivide un linguaggio particolare, come ad esempio i lavoratori che svolgono lo stesso lavoro o la stessa professione, per cui è possibile utilizzare uno slang familiare agli intervistati.

prima di utilizzarle, da esperti e da potenziali rispondenti. Gli esperti sono persone esperte nella stesura dei quesiti del sondaggio o nell'argomento trattato dal sondaggio. Gli esperti possono indicare quali domande del questionario appaiono troppo complesse per essere somministrate facilmente e quali sono troppo lunghe o difficili per poter rispondere con precisione. Per quanto concerne, i potenziali rispondenti, invece, questi sono i soggetti che hanno i requisiti per far parte del campione del sondaggio. La revisione da parte dei potenziali intervistati aiuta a garantire che le domande del sondaggio siano significative e includano tutti i fattori ritenuti fondamentali per ottenere l'informazione desiderata. Un'altra accortezza che dev'essere osservata dal team che conduce l'indagine è quella di evitare domande a doppio taglio, ovvero un quesito che contiene due idee distinte. Per capire meglio la fattispecie esaminata, si fa riferimento ad un esempio contenuto nel manuale "How to ask survey questions" di A. Fink, che dice: "Pensa che dovremmo continuare a usare il denaro delle tasse per sostenere i programmi artistici e sportivi nelle scuole pubbliche?". Questa domanda è in realtà duplice: "Pensa che dovremmo continuare a usare il denaro delle tasse per sostenere i programmi artistici?" e "Pensa che dovremmo continuare a usare il denaro delle tasse per sostenere i programmi sportivi?". Certamente, alcune persone sarebbero favorevoli a spendere i soldi delle tasse per sostenere i programmi artistici, altre per sostenere gli sport, altre ancora per entrambi e altre ancora per nessuno dei due. In ultima analisi, è opportuno evitare domande contenenti negazioni, in quanto sono difficili da rispondere per molti intervistati perché richiedono un esercizio di pensiero logico. In tale fattispecie, il rischio consiste nel fatto che alcuni rispondenti potranno attribuire alla domanda un valore positivo, mentre altri un significato negativo. Per tali ragioni è necessario sottolineare la parola negativa qualora si dovesse decidere di inserirla ugualmente all'interno dell'indagine che si sta portando avanti. È possibile, inoltre, adattare le domande utilizzate con successo in altri sondaggi. Questa tecnica, laddove possibile, è largamente impiegata dal momento che si tratta di domande che sono già state esaminate, utilizzate e che hanno dimostrato di raccogliere informazioni accurate. Solitamente, all'interno dei questionari si effettua un ampio ricorso a domande brevi per ridurre al minimo il tempo di lettura da parte dell'intervistato e quindi lo sforzo che quest'ultimo deve sostenere³. Le risposte a queste domande tendono anche a fornire

³ È evidente come sia possibile ricorrere all'utilizzo di questa tipologia di domande anche

informazioni meno dettagliate. Le domande più lunghe spesso forniscono informazioni di base agli intervistati e li aiutano a ricordare o a riflettere sul motivo per cui hanno fatto qualcosa o hanno una particolare opinione.

In ogni caso, la valutazione delle scelte progettuali dei ricercatori avviene tramite il processo di risposta. La reazione dell'intervistato alla formulazione, all'ordine delle domande, alla modalità di raccolta dei dati e così via, riveleranno la qualità delle decisioni assunte dagli intervistatori. Inoltre, l'analisi delle strategie di risposta, ovvero i processi cognitivi e comunicativi che guidano gli intervistati, rivelerà i tipi di bias che possono compromettere la qualità dei dati e indicherà i rimedi e le soluzioni appropriate. Tuttavia, la risposta fornita dall'intervistato può dipendere da una serie di fattori di natura psicologica, che caratterizzano un determinato contesto. Per questa ragione, si deve prestare particolare attenzione ai riscontri che emergono dopo la somministrazione del questionario, in quanto questi potrebbero essere influenzati dalle suddette distorsioni. In primo luogo, è necessario verificare se la risposta sia sincera o meno. Infatti, l'individuo potrebbe rispondere in modo non veritiero in base a ragioni legate alla mancanza di voglia o di tempo per leggere e rispondere correttamente ai quesiti contenuti nel sondaggio, ma anche per aspetti legati alla desiderabilità sociale. Con quest'ultimo termine si intende definire quell'effetto di disturbo che subentra quando, nell'ambito di una ricerca o di uno studio, il soggetto rispondente ad un'intervista o ad un questionario ha la possibilità di dare risposte che possono essere considerate socialmente più accettabili rispetto ad altre: questo fa sì che le persone cerchino di comportarsi in modo idealistico, ovvero tendano a sembrare più normali possibili, nel senso di più adeguate alla norma (Roccatò, 2003). Tanto premesso, risulta evidente come tale meccanismo possa influenzare la validità delle risposte, portando gli intervistati a mostrarsi sotto una luce socialmente desiderabile che non corrisponde all'espressione sincera delle proprie convinzioni o comportamenti. Pertanto, gli intervistatori devono cercare di ridurre l'effetto di questa distorsione attraverso la formulazione delle domande, ad esempio, cercando di sdoganare alcuni comportamenti o di inserire dettagli nella domanda per incoraggiare l'intervistato a concentrarsi sui fatti. Un'altra metodologia che può essere applicata da coloro i quali conducono l'indagine è la cosiddetta risposta randomizzata, ovvero si accoppia una domanda

quando si vuole ottenere una risposta breve.

soggetta a bias di desiderabilità sociale con una domanda innocua, la cui distribuzione delle risposte è nota da un'altra indagine. Così facendo, l'intervistato risponde casualmente ad una sola delle due domande, senza dire all'intervistatore a quale dei quesiti ha risposto. Un'idea simile è la tecnica dell'*item count* o *unmatched count*, in cui l'item soggetto a desiderabilità sociale è nascosto in un elenco di atteggiamenti o comportamenti innocui, per cui ad un intervistato viene data una lista con o senza l'elemento stigmatizzato e gli viene chiesto di indicare il numero di atteggiamenti che approva o di comportamenti che ha messo in atto. La stima desiderata si ottiene sottraendo i conteggi ottenuti quando l'item soggetto a bias di desiderabilità sociale è e non è incluso nella domanda. Oltre alla desiderabilità sociale, un altro aspetto che può avere implicazioni sulle risposte è rappresentato dall'acquiescenza. Con questo termine si sottendono quelle forme di accondiscendenza e remissività, a volte anche incondizionate, che si concretizzano nel rispondere positivamente a tutte le domande, a prescindere dal contenuto di quest'ultime. In altri termini, l'acquiescenza, detta anche bias di risposta concorde, è la tendenza di chi risponde a un sondaggio a rispondere positivamente, indipendentemente dal contenuto della domanda e tale tendenza può essere osservata in termini di soddisfazione o di risposta con il minimo sforzo, specialmente se la domanda è formulata in modo complesso o se l'intervistato non ha un'opinione su un certo argomento. Sebbene possa verificarsi anche per le domande sui fatti, l'acquiescenza è tipicamente un problema nella misurazione degli atteggiamenti e la sua entità risulta imprevedibile e difficile da misurare. La strategia tradizionale consiste nell'invertire la direzione di alcuni item, in modo che gli intervistati che rispondono in modo coerente siano d'accordo con alcune affermazioni e in disaccordo con altre. Questa soluzione, tuttavia, presenta notevoli svantaggi. Infatti, un *item* invertito può suonare innaturale e quindi le risposte sono soggette a errori di risposta. Inoltre, la necessità per l'intervistato di individuare brevi parole che cambiano la direzione della domanda può aumentare l'errore. Sebbene la modifica della direzione di circa la metà degli item elimini la distorsione in una scala additiva⁴, la distorsione nei singoli item non è influenzata. Un modo complicato per risolvere questo problema è quello di randomizzare la direzione di ogni singolo item. Se l'acquiescenza è una

⁴ Con "scala additiva" si fa riferimento ad un tipo di misurazione che consente l'aggiunta o la somma di quantità. Tuttavia, le scale di misura saranno approfondite nel proseguo della trattazione.

preoccupazione seria, una soluzione migliore è cambiare le domande in un formato in cui la risposta non implica l'accordo.

Non è solo importante valutare la sincerità del rispondente, ma esistono altri fattori capaci di influenzare la validità delle risposte. Tra questi, si possono citare i bias di auto-selezione, ossia quando le risposte che vengono fornite ad un questionario sono rappresentative solamente di una parte della popolazione, in quanto sussiste il rischio che le persone che decidono di partecipare ad un sondaggio potrebbero avere opinioni o esperienze differenti rispetto a coloro che scelgono di non parteciparvi. Oltre alla formulazione delle domande, anche l'ordine di quest'ultime può condizionare le risposte e quando ciò si viene a verificare, si parla di effetto ordine. Esistono due aspetti legati all'effetto ordine, ossia, l'effetto di *Primacy* e l'effetto di *Recency*. L'effetto di *Primacy* si realizza quando sono le prime domande del sondaggio ad influenzare le risposte successive⁵, mentre quello di *Recency* ha luogo quando sono le ultime domande ad influire su quelle precedenti e questo può accadere, solitamente, quando le ultime domande esercitano un'influenza sulla percezione del questionario che i rispondenti possiedono oppure quando suscitano emozioni forti. Per ridurre l'effetto ordine, i ricercatori possono attuare strategie, quali la casualizzazione dei quesiti e la presentazione di versioni differenti del sondaggio con domande poste secondo ordini diversi⁶. Tornando all'influenza che alcuni elementi possono esercitare sulle risposte ad un questionario, occorre tenere in considerazione anche l'effetto del contesto, il livello di conoscenza e i cambiamenti nel corso del tempo. L'effetto del contesto fa riferimento al fatto che il significato di una domanda potrebbe essere condizionata da domande precedenti o dal contesto in cui viene posta, da qui si intuisce come la comprensione di una domanda possa essere modificata dalle informazioni fornite precedenza e quindi sussiste il rischio che il rispondente possa rispondere in maniera distorta perché influenzato dai quesiti precedenti. Per quanto riguarda, invece, il livello di conoscenza, si può affermare come tale aspetto giochi un ruolo fondamentale nella qualità e nell'accuratezza delle risposte, dal momento che i soggetti che possiedono una

⁵ Ad esempio, se le prime domande vengono formulate negativamente, vi è il rischio che gli intervistati possano rispondere in maniera più critica o pessimista alle domande che verranno successivamente.

⁶ Somministrare differenti versioni dello stesso questionario, ognuna delle quali caratterizzata da un ordine delle domande differente, è fondamentale affinché l'ordine non influenzi sistematicamente le risposte.

conoscenza maggiormente approfondita su un determinato argomento potrebbero rispondere in maniera più completa rispetto a coloro che non sono correttamente informati o che comunque hanno una conoscenza della materia oggetto di indagine piuttosto limitata. Infine, si consideri che le risposte ad un questionario possono variare nel corso del tempo per differenti ragioni come cambiamenti nelle circostanze personali, nell'esperienza di vita o nelle percezioni dell'individuo.

Soffermandosi con maggior attenzione sulle implicazioni dei processi mentali messi in moto dalle domande del questionario, diversi ricercatori hanno tentato nel corso degli anni di delinearne i tratti distintivi e il risultato di questo tentativo si è concretizzato nell'individuazione di quattro gruppi di processi, individuabili in: comprensione, recupero, giudizio e reporting. In alcuni casi, è importante prendere in considerazione anche i processi cognitivi che hanno luogo prima dell'intervista, quando l'intervistato sperimenta per la prima volta gli eventi in questione. La codifica è il processo di formazione dei ricordi a partire dalle esperienze e, in generale, i progettisti dei sondaggi hanno un impatto limitato su questi processi, ma è dimostrato che le domande dei sondaggi possono essere migliorate se prendono in considerazione come gli intervistati hanno codificato le informazioni che le domande del questionario cercano di raccogliere. Tuttavia, come dimostrato dagli studi condotti in letteratura, non si deve fare l'errore di pensare che i rispondenti seguiranno una sequenza fissa di processi cognitivi che inizia con la comprensione delle domande e termina con la comunicazione della risposta, in quanto spesso gli intervistati saltano alcuni dei processi mentali sopra elencati e ne mettono in pratica altri in modo approssimativo. Nonostante ciò, la struttura composta da comprensione delle domande, recupero dell'informazione, giudizio e reporting della risposta risulta essere, seppur con i limiti sopracitati, la più corretta e coerente per spiegare le implicazioni psicologiche sottostanti alla risposta. Per questo motivo, occorre approfondire con un maggior grado di dettaglio le fasi che compongono la sequenza.

Il primo passaggio di tale sequenza è costituito dalla comprensione, la quale include processi come l'attenzione alla domanda e alle istruzioni che la accompagnano, l'attribuzione di un significato alla forma superficiale della domanda e la deduzione del senso della domanda, ovvero l'identificazione delle informazioni richieste. Una componente chiave dell'interpretazione di una domanda è la determinazione

dell'insieme delle risposte ammissibili e spesso i questionari offrono agli intervistati un aiuto in questo compito fornendogli un insieme di categorie di risposte o di linee guida chiare riguardanti la forma che le risposte devono possedere. È probabile che l'interpretazione comporti processi quali il *parsing*⁷ della domanda, l'assegnazione di significati agli elementi sostanziali chiave, l'inferenza dello scopo dietro la domanda, la determinazione dei confini e della potenziale sovrapposizione tra le risposte ammissibili. Nello svolgere questi compiti interpretativi, gli intervistati possono andare oltre le informazioni fornite dalla domanda stessa, guardando alle voci precedenti del questionario o agli ulteriori spunti disponibili dall'intervistatore e dal contesto.

Il secondo passaggio della sequenza è dato dal recupero, ovvero il processo di richiamo dalla memoria a lungo termine⁸ delle informazioni rilevanti per rispondere alla domanda. Ciononostante, esiste un certo consenso sul modo in cui le persone recuperano le informazioni sugli eventi che hanno vissuto. I rispondenti a un sondaggio potrebbero iniziare con una serie di indicazioni che includono un estratto della domanda stessa. Un altro spunto di recupero può essere rappresentato da un richiamo alla memoria, dal momento che fornisce un indizio che aiuta a richiamare le informazioni dalla memoria a lungo termine. In molte situazioni, però, la memoria non fornirà la risposta esatta, ma fornirà informazioni rilevanti, sottoforma di ulteriori spunti che portano alla risposta. Le inferenze e gli altri ricordi basati su questi spunti autogenerati contribuiscono a limitare la risposta e offrono ulteriori spunti per sondare la memoria. Questo ciclo di generazione di spunti e di recupero delle informazioni continua finché l'intervistato non trova le informazioni necessarie o rinuncia. Diversi fattori influenzano il successo del recupero. In parte, il successo dipende dalla natura degli eventi in questione. Alcuni aspetti sono generalmente più difficili da ricordare, poiché sono legati ad eventi che non sono molto caratteristici e che quindi non hanno lasciato il segno nella mente dell'intervistato. Naturalmente, un altro fattore importante che determina la difficoltà o la facilità di ricordare qualcosa è il tempo trascorso, dal momento che con il passare del tempo diventa più difficile ricordare attitudini comuni, avvenimenti poco rilevanti e così via. Un altro fattore che può influire sull'esito del recupero è il numero e

⁷ Si tratta di un processo che analizza un flusso continuo di dati in entrata, in maniera tale da determinare la correttezza della struttura grazie ad una certa grammatica formale.

⁸ Con memoria a lungo termine si intende il sistema di memoria che immagazzina i ricordi autobiografici e le conoscenze generali.

la ricchezza degli spunti che avviano il processo e, in generale, si può affermare che gli spunti che più aiutano il rispondente nel recupero dell'informazione sono quelli maggiormente ricchi di dettagli, mentre quando gli spunti forniti dalla domanda non corrispondono alle informazioni effettivamente immagazzinate nella memoria, il recupero può fallire.

Per quanto concerne il giudizio, con questo termine si intendono i processi di combinazione o integrazione di ciò che l'intervistato ha recuperato. I giudizi possono basarsi sul processo di recupero del giudizio, se è stato difficile o facile recuperare un determinato giudizio, e possono colmare le lacune di ciò che viene ricordato, combinare i prodotti del recupero o correggere le omissioni nel recupero.

Infine, l'ultima fase è quella del reporting. Il reporting è il processo di selezione e comunicazione di una risposta, comprende la mappatura della risposta sulle opzioni di risposta della domanda e la modifica della risposta in base alla coerenza con le risposte precedenti, all'accettabilità percepita o ad altri criteri. Esistono due tipi principali di domande che si basano sui formati di risposta. Una domanda "chiusa" presenta agli intervistati un elenco di risposte accettabili. Le domande "aperte" consentono agli intervistati di fornire le risposte nei loro termini, anche se, in genere, la risposta è comunque abbastanza circoscritta⁹. Il modo in cui gli intervistati scelgono di riportare le loro risposte dipenderà in parte dalla corrispondenza tra le informazioni che recuperano (o stimano) e i vincoli imposti dalla domanda. Oppure, quando non vengono fornite opzioni di risposta, gli intervistati possono dover decidere quanto esatta debba essere la loro risposta e arrotondare le risposte di conseguenza. Gli intervistati possono anche dare maggior peso a certe opzioni di risposta, a seconda dell'ordine in cui sono presentate.

⁹ Approssimativamente le domande aperte sono come le domande a riempimento di spazi in un test, mentre le domande chiuse sono come le domande a scelta multipla.

1.4 Metodi di somministrazione: criticità e tecniche di intervista.

Tra i differenti metodi di somministrazione del questionario si può affermare che non esiste un procedimento che sia migliore di un altro, dal momento che ognuno di questi presenta i propri punti di forza e di debolezza. Pertanto, le decisioni relative alla scelta del metodo di raccolta delle informazioni devono essere prese caso per caso, per cui occorre conoscere quali sono le metodologie e quali sono i loro vantaggi e svantaggi.

Nella scelta del metodo di raccolta dei dati, si devono considerare tre tipologie di fattori: fattori amministrativi o legati alle risorse, problemi legati al questionario e problemi di qualità dei dati. Nel considerare gli aspetti amministrativi e le risorse disponibili, è essenziale valutare sia il tempo a disposizione per condurre la ricerca sia il denaro dedicato al reclutamento di intervistatori e/o codificatori, all'acquisizione di software, hardware e altri equipaggiamenti e alla costruzione di un campione da sottoporre all'intervista. Gli aspetti relativi al questionario, invece, includono il numero e la tipologia di domande da porre per misurare i concetti e raggiungere gli obiettivi della ricerca, per cui si dovrà andare a decidere quale metodo di raccolta sia il più efficace dal punto di vista dei costi o dal punto di vista della comunicazione, ovvero se quella determinata tecnica produce meno errori di comunicazione rispetto alle altre. Infine, per quanto concerne la qualità dei dati, ci si deve domandare se è probabile ottenere la collaborazione di un maggior numero di intervistati con un metodo di raccolta piuttosto che con un altro, se si ottengono risposte più accurate e complete impiegando un intervistatore e se un metodo include la popolazione che si intende studiare.

Premesso ciò, le indagini campionarie effettuate attraverso la somministrazione di un sondaggio possono essere eseguite secondo differenti modalità: interviste postali, interviste online, interviste telefoniche e interviste personali.

1.4.1 Le interviste postali

Le interviste postali prevedono l'invio di una breve lettera di preavviso, di una lettera dettagliata di presentazione e di un questionario a una persona o a un indirizzo

specifico. La lettera di presentazione deve indicare lo scopo dell'indagine, il soggetto¹⁰ che raccoglie i dati, chi è preposto alla compilazione del questionario, perché è importante rispondere, una garanzia di riservatezza e i termini entro i quali si deve restituire il questionario compilato. I questionari per posta richiedono un questionario totalmente autoesplicativo, per cui è necessario prestare dichiarazioni chiare e semplici che possano essere comprese dalla parte più ampia della popolazione. Si tratta di un momento molto importante, siccome i rispondenti, se trovano troppo difficili i contenuti espressi all'interno del questionario oppure se considerano il sondaggio troppo lungo, potrebbero commettere errori nella compilazione e potrebbero essere indotti a non completarlo. Quindi, le lettere di accompagnamento devono fornire un numero di telefono che gli intervistati possono contattare in caso di dubbi sulla legittimità del sondaggio o di difficoltà nell'interpretazione delle domande¹¹. In sintesi, le interviste condotte per posta cercano di ottenere alti tassi di risposta inviando una cartolina di ringraziamento, una lettera di presentazione, due copie del questionario e un contatto speciale solamente per coloro che non rispondono per telefono e posta prioritaria, ma anche impiegando incentivi monetari e non monetari¹².

Uno dei vantaggi dei questionari inviati per posta è quello di possedere un costo molto più contenuto rispetto a quello che viene sostenuto per un'intervista telefonica o faccia a faccia. Si tratta di costi irrisori legati ad aspetti pratici come l'affrancatura, l'acquisto delle buste e dei toner per le stampanti che dovranno stampare il questionario, oltre a quelli sostenuti per assumere un personale capace di assemblare correttamente il materiale da inviare, di tener traccia di coloro che rispondono e di coloro che non rispondono e di effettuare un'analisi dei dati raccolti dal campione. In un'indagine per posta i dati non vengono raccolti da intervistatori, con un chiaro vantaggio relativo al costo. Infatti, il costo dell'affrancatura per spedire e restituire un questionario¹³ è di gran

¹⁰ Con soggetto è da intendersi una persona fisica o un gruppo di ricercatori oppure anche una persona giuridica, come un'associazione o un ente statale.

¹¹ Va detto che solamente una piccola percentuale del campione chiama il numero indicato nella lettera, in particolare, dalla letteratura è emerso che circa l'1% dei rispondenti utilizza il numero verde.

¹² Come incentivi non monetari si possono individuare buoni, sconti, coupon, ecc. Solitamente, si tratta di una soluzione adottata da associazioni, aziende e altre persone giuridiche.

¹³ Nelle interviste postali, in genere, si sceglie di inviare questionari non troppo lunghi, in modo da non scoraggiare il rispondente a portare a compimento la compilazione. In genere, si prediligono questionari da otto pagine, ovvero quattro fogli stampati fronte-retro.

lunga inferiore a quello che si dovrebbe sostenere per retribuire un intervistatore. Siccome il costo di spedizione di un questionario è grossomodo identico sia che questo venga spedito in un altro continente sia che questo venga inoltrato nella medesima città, è possibile condurre un'indagine su un campione molto ampio in misura efficace dal punto di vista dei costi, senza dover sostenere spese telefoniche o senza dover affrontare viaggi. Un ulteriore vantaggio dei questionari impiegati nelle indagini per corrispondenza è rappresentato dal fatto che i rispondenti possono consultare fonti online e cartacee e i propri conoscenti per compilare il sondaggio. Questo costituisce un punto importante in quanto permette agli intervistati di non dover ricorrere alla memoria per rispondere alle domande. Così facendo, i ricercatori avranno a loro disposizione risultati maggiormente accurati e privi di distorsioni dovuti all'effetto memoria.

Nel corso degli anni, i questionari postali hanno avuto successo nella raccolta di dati su argomenti sensibili, poiché i rispondenti si sentono più a loro agio di fronte a tali domande, che alle volte possono risultare quasi imbarazzanti per quest'ultimi, se si trovano in una zona di comfort e se il sondaggio è autosomministrato. Il tempo necessario per condurre tale metodologia di raccolta dati va da un minimo di quattro settimane a un massimo di dieci settimane, indipendentemente dalle dimensioni del campione e dalla sua distribuzione geografica. Nel considerare il periodo di tempo per la somministrazione, si deve prevedere un intervallo sufficiente per far arrivare i questionari agli intervistati, per farli completare e per la restituzione dei sondaggi compilati. Tuttavia, è opportuno specificare che il tempo può essere un vantaggio o uno svantaggio in base alle specificità del questionario e delle risorse, in termini economici e di personale, disponibili.

Nella valutazione del metodo di raccolta dei dati si deve considerare il tasso di risposta. Innanzitutto, il tasso di risposta può essere definito come il numero di membri del campione idonei che completano un questionario diviso per il numero totale di membri del campione idonei (American Association for Public Opinion Research). L'importanza ricoperta da tale tasso è evidente, in quanto rappresenta un indicatore di qualità del sondaggio, motivo per il quale, più è alto il tasso e migliore è la qualità del questionario somministrato. Esistono alcuni mezzi che i ricercatori possono impiegare per aumentare

i tassi di risposta per posta, ovvero gli incentivi e i ripetuti tentativi di *follow-up*¹⁴. Nella prassi, gli incentivi monetari e quelli non monetari pregiati risultano essere gli strumenti che più aumentano i tassi di risposta, mentre gli incentivi condizionati alla restituzione del questionario¹⁵ compilato non hanno generalmente alcun impatto. Nello specifico, si possono ottenere tassi di risposta piuttosto elevati da un questionario per posta quando l'argomento è molto importante per il rispondente.

Dopo aver osservati i potenziali vantaggi, occorre concentrare l'attenzione sugli svantaggi che tale metodo possiede. In primo luogo, i sondaggi per posta sono soggetti a bias di non risposta perché non raggiungono buoni tassi di risposta da parte di persone con un basso livello di istruzione, persone che non amano scrivere, persone che hanno difficoltà a leggere e persone che non hanno interesse per l'argomento. Più in generale, il bias di non risposta è potenzialmente maggiore nelle indagini per posta rispetto ad altri metodi, perché gli intervistati possono ignorare più facilmente questi sondaggi. Per questo motivo, la lettera di presentazione e l'aspetto del questionario sono di fondamentale importanza per incoraggiare la collaborazione e minimizzare i bias di risposta. Di solito l'unico contatto con gli intervistati avviene attraverso questi materiali, che quindi devono essere convincenti. Per valutare la potenziale distorsione delle mancate risposte, è necessario conoscere il più possibile i non rispondenti, ossia è importante determinare se coloro che rifiutano o che non possono essere raggiunti per un'intervista sono diversi, in termini di variabili dipendenti, da coloro che collaborano. Oltre ai problemi di mancata risposta, alcune caratteristiche delle indagini per corrispondenza possono renderle meno efficaci di quelle somministrate dagli intervistatori. Siccome gli intervistati possono esaminare il sondaggio prima di decidere se compilarlo o meno, il questionario non può essere molto lungo, né può sembrare complesso o difficile da compilare. Inoltre, un questionario inviato per posta deve essere completamente autoesplicativo, perché non c'è nessuno che possa assistere l'intervistato qualora non dovesse capire qualche punto. Infine, il ricercatore ha uno scarso controllo sull'ordine in cui gli intervistati rispondono alle domande o su chi effettivamente

¹⁴ Con *follow-up* si intende la procedura con la quale vengono controllati nel tempo i soggetti facenti parte di un campione, al fine di capire se sia o meno sviluppato il risultato al quale si è interessati.

¹⁵ Ovvero una serie di incentivi che vengono erogati dopo la consegna del sondaggio, come la partecipazione a estrazioni o concorsi, la possibilità di accedere a contenuti esclusivi

compila il questionario. Infine, le risposte alle domande aperte sono in genere meno approfondite e dettagliate nei sondaggi per corrispondenza rispetto a quelli somministrati dagli intervistatori, questo perché, a meno che non siano sollecitati dagli intervistatori, molti rispondenti forniscono solo il minimo delle informazioni necessarie per rispondere alla domanda aperta. Questa tendenza è particolarmente vera per gli intervistati con un basso livello di istruzione e per quelli che non amano scrivere, mentre per quelli più istruiti o più articolati, questo potrebbe non essere un problema.

1.4.2 Interviste online

Negli ultimi anni, grazie alla rapida e capillare espansione della rete internet, i sondaggi online hanno progressivamente soppiantato le interviste per corrispondenza. Infatti, se nei primi anni successivi all'introduzione dei sondaggi online sussisteva un problema di copertura, poiché non tutti gli individui avevano la possibilità di utilizzare internet con frequenza, la situazione odierna è ben diversa con la maggior parte degli intervistati che può accedere ad una connessione internet quotidianamente. Rispetto ai sondaggi per corrispondenza, l'unico svantaggio, che può comunque essere superato, è rappresentato dalla difficoltà di ottenere un elenco di indirizzi e-mail. Anche in questo caso riveste una grande importanza il contatto iniziale, che ha lo scopo di spiegare le motivazioni e l'importanza del questionario, di identificare il ricercatore, di garantire e di fornire le istruzioni per accedere all'indagine. In alternativa, può essere utilizzato un messaggio introduttivo breve, che dichiara lo scopo dell'indagine e incoraggi la partecipazione. Questo messaggio dovrebbe essere fornito nella prima schermata o pagina di benvenuto del questionario web, insieme a istruzioni semplici per inserire il codice identificativo unico¹⁶ e per accedere alla prima pagina del questionario. Anche in questo caso, dovrebbe essere fornito un numero di telefono, oppure un indirizzo mail, per i rispondenti che possono avere difficoltà nel rispondere alle domande. Il questionario online deve essere progettato con cura per essere accessibile e comprensibile ai rispondenti con diversi livelli di istruzione, alfabetizzazione informatica, hardware e software del computer e accesso a Internet. Inoltre, per aumentare i tassi di risposta è essenziale contattare più volte le persone incluse nel campione che non si collegano al

¹⁶ È possibile, per chi somministra il sondaggio, assegnare un codice identificativo unico che consente al rispondente di compilare il questionario un'unica volta.

sito web dell'indagine o non inviano i questionari compilati entro determinati periodi di tempo. Se l'obiettivo della ricerca è stimare i parametri della popolazione o supportare un altro tipo di inferenza o analisi, nella scelta della modalità di indagine il ricercatore deve considerare i fattori, come il tasso di copertura, i tassi di risposta o le distorsioni nelle risposte fornite dal campione, che più probabilmente influiscono sulla validità dello studio. Se si volesse soffermarsi sulle peculiarità di tale metodologia, si potrebbe affermare che i questionari web possono includere schemi di salto¹⁷, invisibili ai rispondenti, complessi, una serie di aiuti visivi come istruzioni a comparsa, elenchi a discesa, immagini, video clip, animazioni e persino l'audio. In un questionario online è anche possibile incorporare la risposta a una domanda precedente in una domanda successiva. I questionari online devono essere relativamente brevi. Per evitare alti tassi di non risposta, di non risposta all'item e di interruzioni, i ricercatori che hanno sperimentato diverse lunghezze di questionari hanno scoperto che i questionari online non dovrebbero richiedere più di 15 minuti per essere completati e che è preferibile una durata massima di 10 minuti

I vantaggi principali dei questionari online sono principalmente due: la velocità nella raccolta dei dati e il costo molto contenuto. I questionari online eliminano i costi sostenuti per la retribuzione dell'intervistatore in un'intervista personale, i costi legati alla carta, alla riproduzione dei questionari, i costi per l'affrancatura e i costi connessi all'inserimento dei dati ottenuti dai sondaggi per posta. In aggiunta a ciò, la dimensione del campione e la sua distribuzione geografica non incidono sul costo di un sondaggio via internet, siccome raccogliere dati da un campione nazionale o addirittura internazionale via Internet non è più costoso che sondare un campione geograficamente concentrato. Poiché i costi connessi alla raccolta dei dati costituiscono una delle voci più significative del costo totale di un sondaggio, le interviste online consentono ai ricercatori di aumentare le dimensioni del campione e di effettuare un *follow-up* successivo ad un costo aggiuntivo minimo. L'altro grande vantaggio delle interviste online rispetto ad altri metodi è la velocità di raccolta dei dati. Infatti, il periodo di raccolta dei dati per i sondaggi via internet è compreso tra i dieci ed i venti giorni e, in

¹⁷ Sono tecniche che consentono di indirizzare gli intervistati verso domande specifiche sulla base delle loro risposte.

taluni casi, può essere significativamente più breve¹⁸. La velocità nella raccolta dei dati dipende da differenti fattori, tra i quali la disponibilità degli utenti a compilare il questionario, il metodo di distribuzione e il design del sondaggio. Per questo motivo, è opportuno considerare in maniera attenta questi aspetti durante la progettazione e l'esecuzione del questionario al fine di massimizzare la partecipazione e ottenere dati affidabili e significativi.

Uno dei principali svantaggi delle interviste online è rappresentata dal fatto che sussistono differenze demografiche piuttosto consistenti tra coloro che hanno e coloro che non hanno accesso ad Internet, anche se nel corso del tempo è cresciuta l'alfabetizzazione digitale degli utenti e, di conseguenza, è diminuita la percentuale di individui che non sono capaci a navigare in rete. Nonostante ciò, persiste una quota parte della popolazione che non possiede l'accesso ad Internet, principalmente si tratta di persone anziane o persone sotto la soglia di povertà, e questo rende impossibile per i ricercatori l'implementazione di sondaggi online su determinate problematiche che possono interessare in modo particolare gli individui che rientrano in una determinata fascia e questo potrebbe rappresentare una difficoltà per i ricercatori che vorrebbero svolgere inferenza su un'intera popolazione. Per queste ragioni, uno svantaggio dei questionari online è dato dalla difficoltà di sviluppare una buona struttura di un campione basato sulla probabilità¹⁹. Per risolvere questo problema per la popolazione generale è stata talvolta utilizzata una strategia in due fasi, che vede nella prima un campione probabilistico della popolazione generale reclutato con una copertura accettabile e, nella seconda fase, l'intervista viene somministrata come sondaggio sul web. È stata notata in letteratura la tendenza che vede un aumento delle risposte corrispondenti a frasi come "non so" oppure risposte più estreme negli item con scala accordo/disaccordo²⁰. I bassi tassi di risposta e il conseguente potenziale bias di non risposta sono altri svantaggi significativi dei sondaggi via Internet. Infatti, sebbene le

¹⁸ Basti pensare che sono stati riscontrati in letteratura casi di studio che hanno evidenziato come alcuni questionari siano stati compilati e consegnati entro i primi cinque giorni dal suo invio.

¹⁹ Con campione basato sulla probabilità si intende un campione selezionato da una popolazione in modo che ciascun membro della popolazione conosciuta e non nulla di essere incluso nel campione.

²⁰ Si tratta di domande o affermazioni che prevedono una scala che consenta agli intervistati di rispondere esprimendo il proprio livello di accordo o di disaccordo con la domanda o l'affermazione che è stata presentata nel questionario.

informazioni sulle indagini probabilistiche condotte via Internet siano scarse, sembra che i tassi di risposta siano in genere inferiori a quelli dei sondaggi somministrati in altre modalità. Inoltre, per quanto riguarda la possibilità di distorsione delle risposte nei sondaggi su Internet, sembra che le persone con un livello di istruzione più basso, quelle che hanno un'esperienza minima nell'uso del computer e quelle che dispongono di apparecchiature informatiche più vecchie e di browser di fascia bassa abbiano meno probabilità di completare un questionario online rispetto alle persone con un livello di istruzione elevato, con un'alfabetizzazione informatica e con apparecchiature più aggiornate. Si possono, inoltre, annoverare alcune analogie tra i questionari online e quelli per corrispondenza e ciò li rende meno efficaci rispetto a quelli somministrati da un intervistatore. Per capire meglio quanto affermato, come un questionario per posta, un questionario online deve essere completamente autoesplicativo, perché non c'è un intervistatore a spiegare istruzioni o domande confuse o complesse e deve apparire uguale a tutti gli intervistati. Quest'ultima caratteristica risulta molto difficile da mettere in pratica, in quanto coloro che progettano il questionario devono comprendere i fattori che influenzano l'aspetto di un questionario online, come le dimensioni e la risoluzione dei monitor dei computer, i sistemi operativi, browser e così via. Una volta individuati tali elementi, l'obiettivo sarà quello di minimizzare il più possibile le differenze nell'aspetto di un questionario online. Un ulteriore inconveniente di tale metodologia è rappresentato dal fatto che i sondaggi possono essere progettati in modo tale che i partecipanti debbano rispondere ad ogni domanda nell'ordine presentato senza avere la possibilità di passare alla domanda successiva e, in taluni casi, questa peculiarità può provocare tassi di abbandono piuttosto elevati, poiché gli intervistati si possono trovare dinnanzi a domande alle quali non vogliono o non possono rispondere. Inoltre, tale metodologia di somministrazione può non essere molto efficace nella raccolta di dati riguardanti argomenti sensibili, vista la preoccupazione che alcuni partecipanti all'indagine possiedono riguardo la possibilità che le informazioni fornite possano superare l'anonimato del sondaggio e finire in possesso di terzi. Questo timore può tradursi in alti tassi di non risposta alle domande riguardanti le domande su temi sensibili, in un elevato tasso di abbandono del questionario e in una minore onestà nelle dichiarazioni.

1.4.3 Interviste telefoniche

I numeri impiegati nell'ambito di un'indagine telefonica possono essere selezionati in modi differenti, come l'estrazione casuale da un elenco oppure attraverso una selezione casuale delle cifre. I sondaggi telefonici hanno un costo che si attesta tra i questionari per corrispondenza e quelli personali. A differenza delle due precedenti metodologie, per realizzare un'intervista telefonica è necessario che sia presente un intervistatore, il quale avrà bisogno di uno spazio dove saranno presenti telefoni, computer e tutti gli altri strumenti necessari perché possa svolgere adeguatamente il proprio lavoro. Risulta evidente come il campione della popolazione debba essere sviluppato o acquistato. Infatti, come accennato in precedenza, il campione può provenire da elenchi, come quelli telefonici o elenchi soci, oppure può essere costruito dal *team* di ricerca, prendendo come riferimento, ad esempio, il numero di telefono degli abitanti di una determinata città. Oltre a quanto descritto, è necessaria la presenza di personale deputato a tener traccia dei risultati delle interviste. Nella prassi, si assiste ad un fenomeno del quale occorre tenere accortezza, ovvero capita molto spesso che gli intervistatori non siano i ricercatori, ma dei loro sostituti. Questo aspetto può avere dei risvolti positivi, come la possibilità di completare più rapidamente l'indagine rispetto a quanto potrebbe accadere se fossero i ricercatori a lavorare in autonomia, ma anche dei riscontri negativi, soprattutto in termini di costi²¹, dal momento che gli intervistatori dovranno essere formati al fine di porre le domande nel modo corretto e di non condizionare le risposte. All'intervistatore viene consigliato di effettuare non più di quattro richiami in giorni diversi della settimana e in diverse ore del giorno.

Uno dei riscontri è costituito dall'ampia copertura della popolazione che tale metodologia di indagine è in grado di raggiungere visto che la grande maggioranza della popolazione possiede un numero di telefono. Oltre alla vasta copertura, molte persone sono disposte a farsi intervistare per telefono e questo riduce il rischio di non risposta²², anche se alcuni gruppi di lingua straniera non si sentono a proprio agio ad essere intervistati in una lingua straniera, per cui il successo di un questionario può dipendere

²¹ Questo metodo di raccolta dati, insieme all'intervista personale, sono i metodi più costosi dal momento che richiedono un maggior impiego di manodopera.

²² Tuttavia, il tasso di risposta non è facilmente quantificabile, dal momento che dipende da diversi fattori, come la dimensione del settore, la lunghezza del questionario, la tipologia di domande e così via.

dal fatto che gli intervistatori conducano l'indagine nella lingua dei partecipanti. I tassi di risposta sono di solito migliori per i sondaggi telefonici che per quelli via posta e Internet, principalmente perché l'intervistatore cerca di convincere personalmente l'intervistato dell'importanza della ricerca, fissa l'intervista in un momento conveniente per l'intervistato, legge e registra le risposte dell'intervistato. Un altro vantaggio è dato dalla durata del periodo di raccolta dei dati è di solito breve o più breve per le indagini telefoniche rispetto alla maggior parte degli altri metodi. La telefonata è un modo rapido per avviare un contatto con una persona o una famiglia e per effettuare richiami, un fattore importante per contattare persone difficili da raggiungere. La distribuzione geografica del campione può essere ampia perché è facile ed economico acquistare un quadro di campionamento dei prefissi telefonici di un determinato paese e il costo ad esso connesso è abbastanza contenuto, per queste ragioni i costi del tempo di intervista sono grossomodo uguali per le indagini svolte su scale locali, statale e nazionale. L'impiego di intervistatori offre una serie di vantaggi per l'efficacia del questionario e la qualità dei dati, perché con intervistatori sufficientemente formati la qualità delle risposte registrate dovrebbe essere molto elevata. Siccome gli intervistatori sono addestrati su come porre ogni domanda e in quale ordine, la struttura o il layout possono essere complessi, per cui possono essere impiegati schemi di salto multipli, in cui le risposte ad un certo tipo di domande determinano quali sono le domande che vengono poste o saltate dagli intervistati. Un altro vantaggio derivante dall'impiego degli intervistatori è la possibilità di controllare l'ordine delle domande. I rispondenti non hanno idea di quale domanda verrà posta successivamente o del numero di quesiti del questionario. Se si accorgono del numero di domande del sondaggio o del fatto che la risposta "sì" comporterà una serie di domande successive, potrebbero essere meno propensi a farsi intervistare o potrebbero rispondere "no" alla prima domanda della serie. Infine, è possibile per gli intervistatori stabilire un rapporto con gli intervistati al telefono e quindi convincerli a completare l'intervista, a credere nell'autenticità e nella rilevanza della ricerca e a fornire risposte complete e accurate anche a domande delicate. Rispetto alle interviste per corrispondenza e online, il questionario telefonico riesce ad essere più incisivo nella raccolta di informazioni su argomenti sensibili, proprio perché la presenza dell'intervistatore stabilisce un legame più forte con il rispondente.

A fronte di quanto descritto, è necessario dire che uno dei più grandi svantaggi delle interviste telefoniche è rappresentato dal fatto che non tutti gli intervistati sono propensi e disponibili a rispondere all'intervista, perché molte volte i partecipanti all'indagine possono percepire queste chiamate come indesiderate, possono non avere tempo a disposizione per rispondere alle domande oppure possono scambiare questi numeri per *spam* e quindi potrebbero non rispondere neanche al telefono. Qualora si dovesse verificare tutto ciò, evidentemente si otterrebbe una diminuzione del tasso di risposta. In caso contrario, gli intervistati potrebbero non tenere a mente tutte le informazioni. Per quanto riguarda il questionario, le domande del sondaggio telefonico devono essere brevi e abbastanza semplici, e anche le opzioni di risposta che vengono lette agli intervistati devono essere poche, brevi e semplici. Le frasi devono essere limitate a venti parole o meno, il linguaggio deve essere semplice e non devono essere lette più di quattro o cinque categorie di risposte brevi per ogni domanda. Quando le frasi sono lunghe o le categorie di risposta sono numerose, gli intervistati possono ricordare solo le prime parti o le ultime²³. Anche l'impossibilità di utilizzare ausili visivi come immagini, campioni di prodotti o elenchi di alternative di risposta per domande con molte categorie di risposte può essere uno svantaggio. Sono stati tentati vari metodi per aggirare questa limitazione, come l'invio presso le abitazioni degli intervistati del materiale necessario per la piena comprensione dei quesiti, ma con risultati generalmente scarsi. Anche l'impossibilità per l'intervistatore di controllare l'ambiente di risposta e la difficoltà degli intervistati di consultare fonti durante un'intervista telefonica possono rappresentare degli svantaggi. Quando si intervista un intervistato, non si sa mai dove si trovi, se si trovi da solo oppure se c'è qualcun altro nella stanza e quanto l'intervistato sia a suo agio nel rispondere alle domande. Un altro svantaggio, che potrebbe non essere evidente, riguarda le risposte limitate alle domande aperte. Potrebbe sembrare che un intervistatore telefonico sia in grado di sondare le risposte degli intervistati e di registrare le risposte testuali così come un intervistatore faccia a faccia, e in effetti gli intervistati telefonici possono rispondere alle domande aperte. Solitamente gli intervistati rispondono con una frase o con poche frasi brevi. Se la richiesta di "altri motivi", spiegazioni o chiarimenti da parte dell'intervistatore crea silenzio o "tempi morti" e l'intervistato non riesce a pensare ad altro da dire, può

²³ Si tratta degli effetti di *primacy* e di *recency* precedentemente analizzati.

iniziare a sentirsi ansioso o nervoso. Infine, le informazioni sui rifiuti e sui mancati contatti sono piuttosto limitate con le indagini telefoniche, a meno che la struttura di campionamento non includa nomi e indirizzi o altre informazioni identificative.

1.4.4 Interviste personali

Nelle interviste personali le informazioni vengono solitamente raccolte dagli intervistatori presso le case dei partecipanti all'indagine, presso uffici, nei pressi di luoghi affollati come aeroporti o stazioni ferroviarie e in altri luoghi comodi per l'intervistato. L'elemento imprescindibile è che intervistato e intervistatore si trovino insieme nello stesso luogo. Dei metodi analizzati questo risulta essere il più costoso, a causa dei costi di viaggio, che dovranno essere sostenuti per fare in modo che l'intervistatore sia presente nel luogo in cui deve intercettare i rispondenti, e dei costi legati al tempo per raccogliere i dati, siccome l'intervistatore dovrà segnare le risposte e trasmetterle al *team* di ricercatori. In letteratura, i costi legati agli spostamenti e alle attività di redazione delle risposte corrispondono ad un intervallo compreso tra il 60% e il 75% del costo totale del sondaggio. In particolari condizioni, come nel caso di un'indagine svolta presso punti nevralgici di una metropoli in cui si concentrano grandi quantitativi di persone, i tempi di viaggio costituiscono una parte piuttosto consistente del tempo totale degli intervistatori. Ration per cui occorre ottimizzare la durata degli spostamenti che quest'ultimo dovrà effettuare per muoversi da un punto verso un altro, in quanto eventuali ritardi o tempi troppo lunghi potrebbero impedirgli di intervistare soggetti di particolare interesse ai fini dell'indagine.

Nonostante sia la metodologia di raccolta dati più costosa, l'intervista personale può essere considerata migliore rispetto ad altre tecniche di indagine poiché presenta una serie di vantaggi in termini di campionamento e di aumento dei tassi di risposta. Specialmente per quanto riguarda i tassi di risposta, il vantaggio è quello che risulta molto più difficile per un rispondente rifiutarsi di interloquire di persona con un intervistatore rispetto a quanto non lo sarebbe telefonicamente. In termini di qualità dei dati, invece, un vantaggio è costituito sicuramente dal fatto che diminuirà consistentemente la distorsione delle risposte e, contestualmente, aumenterà il tasso di collaborazione degli intervistati. La qualità delle risposte registrate nelle indagini faccia a faccia è giudicata molto buona anche perché gli intervistatori, i quali devono rivolgersi

in modo cortese e attento al fine di creare un rapporto con gli intervistati, ricevono una formazione approfondita per porre le domande e registrare le risposte. Anche il rapporto tra intervistatore e rispondente è migliore rispetto alle altre modalità di somministrazione anche perché il partecipante all'indagine ha la possibilità di vedere davanti a sé la persona con cui sta parlando e che raccoglie le informazioni. Inoltre, tale tipologia di intervista consente un maggior controllo sulle risposte, dal momento che l'intervistatore potrebbe intervenire qualora si accorgesse che l'intervistato stia rispondendo in modo non corrispondente al vero. Trattandosi di un incontro faccia a faccia, l'intervistatore dovrà prestare anche la massima attenzione ai modi e alle espressioni facciali poiché possono influenzare potenzialmente i rispondenti. Molti dei vantaggi del metodo faccia a faccia riguardano il questionario stesso. Infatti, può essere più complesso perché viene somministrato da un intervistatore esperto e, poiché intervistatore e intervistato si trovano nello stesso luogo, è possibile porre compiti o domande complessi. Le interviste faccia a faccia oltre all'uso di ausili visivi, consentono all'intervistatore, per le domande lunghe o le categorie di risposte lunghe, di consegnare all'intervistato una copia della domanda o delle categorie di risposte, rendendo possibile al rispondente seguire la lettura delle voci, al fine di facilitarne la comprensione di cosa gli viene chiesto. Nelle interviste faccia a faccia gli intervistatori hanno anche il controllo dell'ordine delle domande. Come accennato pocanzi, il metodo di indagine faccia a faccia è anche il migliore per le domande aperte, perché queste indagini consentono un'atmosfera e un ritmo più rilassati rispetto alle interviste telefoniche. Di conseguenza, è più facile per l'intervistatore sondare per ottenere ulteriori informazioni e gli intervistati non si sentono a disagio con le lunghe pause tra le risposte, perché possono vedere cosa sta facendo l'intervistatore. Un altro aspetto positivo del metodo faccia a faccia è che molto spesso è possibile per l'intervistato consultare fonti²⁴ per cercare di rispondere in maniera corretta alle domande. Tuttavia, il rispondente potrebbe non essere in grado di reperire le informazioni appropriate in presenza dell'intervistatore, a meno che non sia stato avvisato in anticipo della necessità di consultare fonti e documenti. A questo proposito, le indagini per posta e via Internet

²⁴ Non sempre è possibile per il partecipante consultare fonti e documenti, ma solo in particolari contesti, come i sondaggi condotti presso le abitazioni o gli uffici.

possono essere più vantaggiose perché gli intervistati possono trovare e consultare i dati con maggiore calma e tempo.

Si potrebbero riassumere gli svantaggi del metodo faccia a faccia in quattro categorie principali. Il primo, come spiegato in precedenza, è il costo. Infatti, l'intervista personale è costosa per differenti motivi, come completare l'elenco delle unità abitative e compensare gli intervistatori per il tempo di viaggio e le altre spese, per cui un'indagine nazionale faccia a faccia costerebbe in genere più del doppio di un'indagine telefonica, online o per corrispondenza. Un altro svantaggio evidente è il tempo necessario per completare un'indagine faccia a faccia. In generale, l'intero processo di intervista di un sondaggio personale è due volte più lungo di quello di un sondaggio telefonico simile. Considerando la distribuzione capillare degli intervistatori nell'area di indagine, i tempi di formazione, il trasferimento dei questionari compilati e dei materiali da e verso gli intervistatori e altre considerazioni logistiche, la fase di raccolta dei dati di un'indagine faccia a faccia potrebbe facilmente richiedere molto più tempo rispetto a un'analoga indagine condotta con altre modalità. Un altro possibile svantaggio delle interviste faccia a faccia è l'esitazione degli intervistati a riferire comportamenti molto personali. L'aspettativa è che quanto più personale è il metodo di raccolta dei dati, tanto meno è probabile che gli intervistati riferiscano comportamenti sensibili. Tuttavia, i risultati sono contrastanti, poiché hanno riscontrato tassi di segnalazione più bassi nelle interviste faccia a faccia per gli argomenti sensibili, mentre altri non hanno riscontrato differenze significative in base al metodo di segnalazione. Le ricerche disponibili coprono una vasta gamma di argomenti, un fattore i cui effetti sono difficili da separare dalla modalità. Un altro inconveniente correlato è che gli intervistati hanno maggiori probabilità di fornire risposte socialmente desiderabili in un'intervista faccia a faccia rispetto a quanto non avviene nei sondaggi per corrispondenza o telefonici.

CAPITOLO II. Le Strategie di campionamento

2.1 L'inferenza statistica e il campionamento

Nella ricerca empirica, l'inferenza statistica può essere definita come quel processo che si basa sull'impiego della statistica per trarre conclusioni su un aspetto non noto della popolazione, riferendosi ad un campione casuale estratto dalla popolazione di riferimento. Volendo scendere maggiormente nel dettaglio, l'inferenza statistica si basa su un modello formale di probabilità che collega un set di dati generato casualmente da una popolazione più ampia di riferimento e aiuta a comprendere e a valutare il contributo del singolo studio, tenendo in considerazione l'incertezza della stima derivante dall'errore casuale. L'inferenza statistica si divide in due aree distinte, ovvero le stime e il test di ipotesi. L'obiettivo delle stime è descrivere un aspetto non noto della popolazione, mentre quello del test di ipotesi è decidere quale tra due ipotesi²⁵ formulate sulla suddetta popolazione sia vera. A tal proposito, si deve introdurre il concetto di stimatore ed illustrarne alcune delle sue peculiarità più significative. Uno stimatore, noto anche come statistica campionaria, è una regola, un metodo, una formula o una procedura che indica come stimare una grandezza della popolazione, come la media, la varianza o il coefficiente di correlazione della popolazione, ed in genere viene espressa come funzione dei valori campionari. Il valore numerico assunto da uno stimatore in un determinato campione è noto come stima. Poiché il suo valore varia da campione a campione, lo stimatore è una variabile casuale e avrà una distribuzione di probabilità o di campionamento. La distribuzione di campionamento è una distribuzione di probabilità che spiega l'andamento di uno stimatore quando sono campionati dati provenienti dalla medesima popolazione. Inoltre, è una misura di grande importanza per valutare l'affidabilità della stima rispetto al valore della popolazione. Alcune delle proprietà statistiche più importanti dello stimatore sono sicuramente rappresentate dalla non distorsione, dalla consistenza, dall'efficienza e dall'invarianza. Per quanto riguarda la prima caratteristica, uno stimatore non è distorto quando, in media, la propria stima possiede lo stesso valore del parametro che si intende stimare. Mentre uno stimatore è

²⁵ Nell'ambito di un test di ipotesi si deve scegliere quale sia l'ipotesi vera tra l'ipotesi nulla e l'ipotesi alternativa. In genere, l'ipotesi nulla è quella che rappresenta la presenza di una certa caratteristica all'interno della popolazione, mentre quella alternativa indica l'assenza della suddetta caratteristica.

consistente se al crescere della numerosità campionaria la sua stima converge verso il valore corretto del parametro. La condizione di efficienza si verifica quando lo stimatore risulta essere, tra tutti gli stimatori non distorti, quello con la varianza minima per il parametro di interesse. Infine, uno stimatore è invariato se la propria stima non varia quando i dati sono assoggettati ad una trasformazione appropriata. Tanto premesso, si può affermare come l'inferenza statistica, attraverso corrette ed appropriate metodologie statistiche²⁶, permetta di generalizzare i risultati osservati sul campione all'intera popolazione di riferimento. Da quanto si può intuire, il collegamento tra inferenza statistica e campionamento è molto significativo, dal momento che il campionamento, come si avrà modo di osservare a breve, fornisce i dati sui quali le analisi statistiche devono essere basate. Quindi, la relazione che sussiste tra inferenza e campionamento consente di ottenere informazioni significative che si possono estendere sulla popolazione oggetto di studio, permettendo la comprensione e la spiegazione dei fenomeni studiati nell'ambito dell'indagine che si sta portando avanti. In estrema sintesi, il processo di inferenza statistica è quel procedimento attraverso il quale si utilizzano i dati raccolti da un campione, estratto da una popolazione *target* più ampia, per trarre conclusioni valide sulla popolazione che si intende osservare al fine della ricerca.

Passando, invece, all'analisi del campionamento, si può definire come una delle tecniche più importanti impiegate nella ricerca di mercato allo scopo di raccogliere i dati. Il ricorso a tale pratica è fondamentale perché consente un ingente risparmio di costi e di tempo per i ricercatori. Infatti, analizzando un numero ristretto di unità statistiche rispetto alla totalità della popolazione oggetto di indagine, i ricercatori possono ottenere informazioni più dettagliate in termini qualitativi e quantitativi. Prima di entrare maggiormente nel dettaglio delle strategie di campionamento, bisogna soffermarsi sui passaggi operativi presenti all'interno di un'indagine campionaria.

In primo luogo, occorre definire gli obiettivi della ricerca e della popolazione oggetto di indagine, per cui è necessario che chi conduce l'indagine abbia chiari gli obiettivi che la ricerca intende perseguire e l'insieme degli individui che possono soddisfare i criteri stabiliti per la ricerca. In particolare, per arrivare ad una definizione puntuale e precisa

²⁶ Si fa riferimento ai test di ipotesi, alle stime dei parametri e allo studio delle relazioni tra variabili.

di quest'ultimo aspetto, il *team* che conduce l'indagine deve considerare fattori quali le caratteristiche demografiche, come età, genere e livello di istruzione, le caratteristiche geografiche, tra cui si possono annoverare luogo di residenza e regione o area geografica, le caratteristiche comportamentali, quindi abitudini, attitudini e stili di vita, e le caratteristiche specifiche del contesto, per cui occorre osservare se si tratta di un'azienda, di un'associazione o ente pubblico oppure di una comunità.

Una volta appurati questi aspetti, si procede alla valutazione della lista di campionamento. Si tratta di una fase molto importante nell'economia di un'indagine campionaria, in quanto determina l'efficacia e la rappresentatività del campione che è stato estratto dal totale di una popolazione. Questo momento dell'indagine è fondamentale nella determinazione della rappresentatività del campione selezionato e nella riduzione del rischio di distorsioni o errori. Nell'ambito della valutazione della lista sono coinvolti alcuni passaggi come:

L'esame della completezza della lista. In questa sottofase, i ricercatori dovranno assicurarsi che nella lista di campionamento sia presente ognuna delle unità statistiche della popolazione che si vuole prendere come riferimento, eventuali lacune o omissioni potrebbero influenzare negativamente sulla rappresentatività del campione stesso.

La valutazione della precisione e dell'aggiornamento dei dati a disposizione nella lista di campionamento.

Assicurarsi che la lista di campionamento rifletta la stratificazione della popolazione. Quindi, se la popolazione oggetto di studio si suddivide in differenti strati, la lista dovrà rifletterli proporzionalmente. In questo modo si garantisce che il campione sia rappresentativo delle stratificazioni presenti all'interno della popolazione.

Valutare se la lista di campionamento include sufficienti unità statistiche per soddisfare i requisiti legati alla dimensione del campione stabiliti sulla base degli obiettivi dell'indagine da perseguire.

Esaminare la rappresentatività della popolazione campionata.

A questo punto, si passa alla scelta del piano di campionamento, che può essere probabilistico o non probabilistico. La scelta dipenderà dalla natura dell'indagine, dai suoi obiettivi, dalle risorse disponibili e dal livello di accuratezza ricercato per i risultati. Con campionamento probabilistico si fa riferimento a quella fattispecie in cui ogni unità statistica compresa nella popolazione ha una probabilità non nulla di essere inclusa nel campione, mentre il campionamento non probabilistico viene realizzato su base soggettiva e, di conseguenza, non è nota la probabilità di ogni unità di essere estratta. In questo frangente si è preferito accennare alle due modalità di campionamento che verranno approfondite nel corso della trattazione.

Scelto il piano di campionamento, si procede con la decisione del disegno campionario, ovvero delle diverse tipologie di campionamento. Quindi, qualora i ricercatori optassero per eseguire un campionamento di tipo probabilistico, questi dovrebbero scegliere quale metodo adottare nello specifico, per cui potranno scegliere tra campionamento casuale semplice, campionamento stratificato, campionamento a due stadi e così via.

Dopodiché, si può definire la numerosità campionaria, ossia il numero delle unità statistiche che faranno parte del campione di riferimento. Si tratta di uno dei momenti più importanti nel corso dell'indagine campionaria dal momento che le scelte adottate in sede di definizione della numerosità possono impattare in modo significativo l'accuratezza e l'affidabilità dei risultati. Anche in questo caso la numerosità campionaria può dipendere da una serie di fattori, tra i quali il livello di precisione, maggiore è la numerosità e più le stime saranno precise, il livello di confidenza desiderato, ossia l'intervallo di valori in cui è probabile che sia contenuto il parametro stimato, e la variabilità della popolazione, perché se la popolazione è molto omogenea sarà sufficiente estrarre un campione più ristretto per ottenere stime più precise e accurate.

Infine, gli ultimi due passaggi operativi sono rappresentati dalla valutazione dei costi ed estrazione del campione. La valutazione dei costi è una parte fondamentale nell'ambito di un'indagine perché garantisce un utilizzo efficiente delle risorse che si hanno a disposizione. I costi rappresentano un indicatore fondamentale che può condizionare la pianificazione e l'esecuzione di una ricerca, in quanto è necessario per i ricercatori considerare il *pay-off* tra risultati più accurati e risorse che si hanno a disposizione.

Mentre l'estrazione del campione afferisce al processo attraverso il quale verranno selezionati gli individui che entreranno a far parte del campione, che dovrà essere rappresentativo della popolazione che si intende prendere ad esame.

2.2 Il campionamento probabilistico

Esposto alla comunità scientifica per la prima volta negli anni Trenta del secolo scorso²⁷, il campionamento probabilistico rappresenta quella tipologia di campionamento in cui ciascuna delle unità statistiche appartenenti alla popolazione ha una probabilità non nulla e calcolabile di essere inclusa all'interno del campione. I partecipanti sono selezionati in modo casuale, in modo che tutti i membri della popolazione abbiano le stesse possibilità di essere selezionati, eliminando la possibilità di un bias di selezione del campione. Il campionamento probabilistico comprende le tecniche utilizzate per selezionare un campione che rappresenti chiaramente una popolazione specifica. Tale metodologia rende possibile, attraverso la teoria del calcolo delle probabilità, controllare l'errore campionario e impiegare l'inferenza statistica per generalizzare i risultati emersi dal campione alla popolazione di riferimento nella sua interezza. Dunque, nel campionamento probabilistico, la selezione delle unità per il campione dell'indagine è guidata da una misura di probabilità su tutti i campioni possibili e tale misura di probabilità²⁸ elimina distorsioni associate ai metodi di campionamento soggettivi e ad altri metodi non probabilistici. Ancora più importante è il fatto che il campionamento probabilistico consente di ottenere affermazioni rigorose sulla popolazione sconosciuta con margini di errore controllati attraverso campioni di dimensioni adeguate, in particolare adottando strumenti statistici che sono in grado di fornire stime puntuali e intervalli di confidenza per i parametri della popolazione, tra i quali si possono trovare la media, la moda, la mediana, la varianza, la deviazione standard e così via. Per garantire l'affidabilità e l'accuratezza delle stime è fondamentale controllare le dimensioni del campione, siccome al crescere della numerosità del campione si riduce il margine di errore delle stime e, di conseguenza, aumenterà il livello di precisione dei risultati ottenuti dallo studio. È opportuno sottolineare, però, che la casualità del campione è strettamente connessa al metodo di campionamento, in quanto ciascuna di queste tecniche si fonda sull'impiego di probabilità note e diverse da zero. Si pensa, in genere, che i metodi di campionamento probabilistico siano sempre più accurati nel catturare le caratteristiche essenziali di una

²⁷ Jerzy Neyman, Varsavia, 1934, "On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection.

²⁸ Con questo termine si intende far riferimento a quella funzione che collega all'esito di un determinato evento la probabilità che il suddetto esito si realizzi.

popolazione, ma non è sempre così. Infatti, per i metodi di campionamento probabilistico, le dimensioni della popolazione devono essere note, oppure stimate con precisione, e le persone che ne fanno parte devono essere inserite in un quadro di campionamento. I metodi di campionamento probabilistico non possono essere utilizzati quando i membri della popolazione sono sconosciuti oppure quando non esiste un quadro di campionamento. Un'altra asperità è rappresentata dal fatto che le popolazioni difficili da raggiungere o nascoste rappresentano una percentuale relativamente piccola della popolazione generale, quindi, le interpretazioni statistiche sono spesso inaffidabili a causa dei piccoli campioni e dei frequenti errori di campionamento. Un'altra difficoltà è che le persone che fanno parte di popolazioni difficili da raggiungere o nascoste sono spesso riluttanti a partecipare alla ricerca per timore di essere stigmatizzate o per paura di conseguenze negative. I metodi di campionamento probabilistico funzionano meglio quando la popolazione di interesse è ampia e quando è possibile identificare le variazioni rispetto ai parametri reali della popolazione. Le variazioni rispetto ai parametri reali della popolazione sono chiamate errori di campionamento e possono essere identificate al meglio quando si scelgono campioni sufficientemente grandi. Il campionamento probabilistico è una procedura che si divide in tre fasi. Il primo passo consiste nell'identificare la popolazione target. In questa fase viene definita ed identificata la popolazione di interesse e si procede alla creazione del cosiddetto *frame* di campionamento, ossia un elenco di tutte le unità statistiche che compongono una popolazione, da cui è possibile ricavare un campione. Il secondo passo ha come scopo quello di delineare la struttura di campionamento, ovvero il metodo che sarà impiegato per estrarre il campione dalla popolazione target. Infine, il terzo passaggio si sostanzia nella selezione casuale del campione richiesto dalla struttura di campionamento individuata nella fase precedente. Questo procedimento si svolge utilizzando tecniche di selezione casuale, come il campionamento casuale semplice, affinché ciascuna unità statistica abbia probabilità nota e non nulla di essere coinvolta nel campione.

Come accennato poc'anzi, esistono differenti strutture di campionamento che selezionano in maniera casuale gli individui che andranno a comporre il campione, ognuna delle quali possiede caratteristiche che determinano vantaggi e svantaggi rispetto al fine ultimo di formare un campione che sia rappresentativo della popolazione. In particolare, il campionamento probabilistico si può ulteriormente

classificare come: campionamento casuale semplice, campionamento casuale sistematico, campionamento stratificato, campionamento a grappolo, campionamento multifase e campionamento multistadio²⁹.

Nel campionamento casuale semplice, una volta definita la popolazione oggetto d'indagine, si assegna a ciascuno degli individui un numero identificativo. A questo punto si può procedere alla selezione casuale degli elementi che costituiranno il campione, attraverso l'impiego di numeri casuali generati da software oppure dal sorteggio da un'urna. In questo caso sarà opportuno stabilire se il campionamento casuale semplice sia con oppure senza reinserimento. È importante osservare questo aspetto in quanto i due metodi si fondano su due presupposti differenti. Infatti, nella fattispecie che prevede il reinserimento le unità statistiche che sono state selezionate possono essere rimesse nella popolazione, per cui un soggetto potrebbe essere scelto più volte per far parte del campione, mentre nel campionamento casuale semplice senza reinserimento, gli individui non possono essere rimessi nella popolazione dopo essere stati selezionati, motivo per il quale gli individui possono essere selezionati una volta sola per comporre il campione. Dopo aver effettuato la selezione casuale delle unità statistiche che comporranno il campione, si potrà procedere alla raccolta dei dati d'interesse per lo studio che si sta conducendo. Questo metodo ha come grande vantaggio quello di richiedere una conoscenza della popolazione minima, oltre ad avere un'elevata validità interna ed esterna. A fronte di questi aspetti positivi sussistono alcuni elementi negativi, quali i costi elevati, una minor precisione rispetto ai campioni stratificati della stessa numerosità e la presenza di grandi errori di campionamento.

Il campionamento casuale sistematico prevede che la selezione del primo soggetto avvenga in modo casuale, mentre i soggetti successivi vengono selezionati attraverso un criterio predefinito. Gli individui della popolazione dovranno essere numerati e ordinati e verrà scelta un'unità campionaria ogni K unità della popolazione. Il numero che viene indicato con la K viene definito passo di campionamento ed è determinato dividendo il numero di elementi della struttura di campionamento per la dimensione del campione desiderata. Nell'impiego di questa metodologia di campionamento è necessario adottare alcune accortezze per diminuire il rischio di ottenere risultati distorti. Una di queste

²⁹ Ciascuno dei metodi riportati impiega un processo di selezione casuale.

accortezze prevede di non ordinare gli elementi della lista in funzione della variabile che si intende osservare ed è opportuno che il passo di campionamento ed eventuali andamenti stagionali non coincidano. I vantaggi di questo campionamento sono una maggior facilità e rapidità nel controllo, nella rilevazione e nell'estrazione di un campione, il costo moderato, la validità interna ed esterna e la semplicità della realizzazione. Lo svantaggio è che tecnicamente solo la selezione del primo soggetto è una selezione probabilistica, poiché per le selezioni successive ci saranno alcuni soggetti che avranno zero possibilità di essere selezionati.

Per quanto riguarda, invece, il campionamento casuale stratificato, i dati vengono suddivisi in vari sottogruppi, strati, che condividono caratteristiche comuni come età, sesso, razza, reddito, istruzione ed etnia e da ogni strato viene prelevato un campione casuale. In particolare, la stratificazione si può definire come il processo che consente di aumentare l'efficienza del piano di campionamento, dal momento che permette di ridurre l'ordine di grandezza dell'errore di campionamento senza incrementare la numerosità del campione. Un'altra considerazione da effettuare a tal proposito è rappresentata dal fatto che gli strati, quanto più sono omogenei con riguardo al fenomeno oggetto di studio e tanto più il campionamento stratificato è efficiente rispetto ad altri disegni di campionamento. Il campionamento casuale stratificato ha come vantaggi quelli di assicurare la rappresentazione di tutti i gruppi della popolazione richiesta, di stimare le caratteristiche di ogni strato e di fare confronti. Inoltre, riduce la variabilità del campionamento sistematico. I limiti sono che richiede informazioni accurate sulle proporzioni di ogni strato e che le liste stratificate sono costose da preparare.

Il campionamento casuale a grappoli prevede una suddivisione della popolazione da selezionare in una serie di gruppi, che devono essere esaustivi ed esclusivi, ossia ciascuna unità deve appartenere solo ed esclusivamente ad un solo gruppo. Una volta suddivisa la popolazione in gruppi, si procede alla selezione casuale dei grappoli che comporranno il campione e tutti gli elementi all'interno di ciascun grappolo verranno compresi nel campione, per cui ogni unità ha la possibilità di essere rappresentata all'interno del campione. Tra i vantaggi di questo disegno di campionamento si possono annoverare la maggior efficienza, la facilità di implementazione e il costo contenuto, mentre tra gli svantaggi si possono considerare il rischio di non rappresentatività,

specialmente se i grappoli non sono né omogenei né rappresentativi della popolazione, la perdita di precisione e la difficoltà nel trattare grappoli non omogenei.

Il campionamento multifase è una forma complessa di campionamento a grappolo. In questo caso la popolazione viene organizzata in gruppi, i quali successivamente verranno selezionati in modo casuale e i membri di tali gruppi saranno selezionati in modo casuale³⁰. Una parte delle informazioni viene raccolta dal campione intero e una parte dal sottocampione. Questo metodo di campionamento viene per aumentare la precisione, ridurre i costi e ridurre le mancate risposte.

Infine, il campionamento multistadio è una forma complessa di struttura di campionamento in cui due o più livelli di unità sono inseriti l'uno nell'altro e comporta la ripetizione di due fasi fondamentali, ovvero l'elencazione e il campionamento. Ad ogni fase il cluster diventa di dimensioni sempre minori e, verso la fine, si procede al campionamento dei soggetti. Quindi, tale tecnica può essere un'opzione per campionare popolazioni molto ampie, ma è fondamentale osservare la progettazione e l'implementazione di ogni singolo livello di campionamento, in quanto errori all'interno delle singole fasi potrebbero compromettere la rappresentatività e la precisione del campione.

³⁰ Di solito si estrae un numero di individui uguale per ciascun gruppo.

2.3 Il campionamento non probabilistico

Per quanto riguarda il campionamento non probabilistico, la selezione degli individui che entreranno a far parte del campione avviene su basi soggettive e non è conosciuta la probabilità che ogni unità ha di essere estratta. Rispetto al campionamento probabilistico, non si conosce la probabilità di selezione di un'unità statistica facente parte della popolazione e la selezione può avvenire in modo non casuale, a piena discrezione dell'intervistatore. La soggettività che caratterizza il campionamento non probabilistico può influenzare negativamente le stime e la rappresentatività del campione, in quanto le stime possono più imprecise e maggiormente soggette a distorsioni e la rappresentatività può essere condizionata da fattori non casuali o non noti. Le procedure di campionamento non probabilistico prevedono il reclutamento di partecipanti in modo non casuale per uno studio di ricerca. In questo modo, gli individui della popolazione accessibile non hanno le stesse opportunità di essere selezionati, ciò crea un potenziale bias di selezione, limitando la generalizzabilità dei risultati dello studio. Il bias di selezione si verificherà quando il campione dello studio non è realmente rappresentativo della popolazione di interesse, ma quando coloro che partecipano allo studio condividono uno o più attributi che possono essere sistematicamente diversi dagli attributi di coloro che non vi partecipano. Come il campionamento probabilistico, anche il campionamento non probabilistico è una procedura che si può scomporre in tre fasi. La prima fase consiste nell'identificare la popolazione *target*, mentre nella seconda fase si andrà ad identificare la struttura di campionamento, ovvero la porzione della popolazione *target* accessibile ai ricercatori per l'inclusione nello studio. Infine, nella terza ed ultima fase si procederà al reclutamento degli individui che andranno a formare il campione richiesto dalla struttura di campionamento. Tuttavia, il campionamento non probabilistico può rappresentare un'alternativa migliore rispetto a quello probabilistico in alcune situazioni, come quando i ricercatori non dispongono di una lista della popolazione. Il metodo analizzato porta inevitabilmente con sé alcuni vantaggi e svantaggi. Per quanto riguarda i vantaggi, questi si possono notare in termini di risparmio di tempi, poiché tale metodologia non richiede il calcolo delle probabilità di selezione degli individui, di flessibilità, siccome il ricercatore potrà scegliere i partecipanti a propria discrezione o

sulla base di criteri specifici, permettendogli di concentrarsi su gruppi specifici di particolare interesse ai fini dell'indagine, e di costi. Inoltre, se il criterio di selezione delle unità statistiche è indipendente dal fenomeno di interesse ai fini dell'indagine, un campione non probabilistico ha la stessa valenza di un campione casuale per quanto riguarda l'estensione dei risultati alla popolazione di riferimento. In altri termini, se la selezione degli individui non è influenzata dal fenomeno di interesse ma avviene secondo criteri che non sono correlati direttamente ad esso, i risultati ricavati dal campione non probabilistico potranno essere considerati, in termini di validità, alla stregua di quelli che si potrebbero ottenere da un campione probabilistico quando si tenta di estenderli all'intera popolazione di riferimento. Tra gli svantaggi, invece, oltre a quelli trattati in precedenza, come i potenziali bias di selezione e le difficoltà nel generalizzare i risultati ottenuti, si possono annoverare l'impossibilità di stimare gli errori campionari, dovuta alla mancanza di un processo di selezione casuale delle unità statistiche, e la difficoltà nel replicare l'indagine, siccome i partecipanti non sono selezionati casualmente, per cui la popolazione potrebbe variare di molto da un caso all'altro. Sempre con riguardo alla difficoltà nel replicare l'indagine, un altro punto debole è rappresentato dall'influenza che i partecipanti possono esercitare auto-selezionandosi o selezionando altri individui secondo criteri, come l'interesse per l'argomento oggetto di indagine, in quanto le risposte fornite potrebbero non essere rappresentative della popolazione in generale.

Quando si parla di campionamento non probabilistico non si può non considerare la ponderazione, in quanto risulta essere oggetto di grandi aspettative per la soluzione dei problemi derivanti dai campioni non probabilistici. Prima di entrare nel dettaglio della ponderazione, è necessario specificare che le metodologie che saranno trattate dall'elaborato fanno riferimento al campionamento probabilistico. Tuttavia, nella letteratura e nella pratica, le stesse procedure sono fatte afferire anche ai campionamenti di tipo non probabilistico, con alcune differenze che verranno osservate. Tanto premesso, nei campioni probabilistici i cosiddetti pesi base vengono utilizzati per compensare le probabilità di inclusione. In un secondo momento, si applicano pesi specifici per le mancate risposte per ridurre la distorsione, ovvero una differenza tra la stima e il valore reale, derivante dalle mancate risposte. Un requisito fondamentale per i pesi è che le variabili utilizzate devono essere misurate nel campione e anche nella

popolazione. Inoltre, affinché la ponderazione sia efficace, è necessario che le variabili impiegate siano correlate con le variabili target. In aggiunta a ciò, con la cosiddetta ponderazione della popolazione, di solito si correggono le discrepanze nelle variabili ausiliarie disponibili³¹. Esistono diversi approcci di ponderazione, come quello basato sulla stima di regressione generalizzata, noto come ponderazione lineare, in cui si utilizzano modelli di regressione lineare. Un altro esempio, nonché il più usato, è rappresentato dalla post-stratificazione, in cui le variabili ausiliarie dividono semplicemente la popolazione in diversi strati e a tutte le unità appartenenti ad uno stesso strato viene assegnato lo stesso peso. In letteratura viene delineato anche un secondo filone, basato sulla stima del *ranking ratio*³². Tale approccio è strettamente correlato all'adattamento proporzionale iterativo³³, in cui si aggiusta in sequenza per ogni variabile ausiliaria separatamente, come l'età e il sesso, e si ripete il processo fino a raggiungere la convergenza. Nella maggior parte delle situazioni la ponderazione lineare e il ranking forniscono risultati asintoticamente simili, anche se possono emergere differenze, a seconda che la relazione sottostante tra le variabili ausiliarie, o di controllo, e quelle target sia lineare o loglineare, come ipotizzato nel ranking. Vanno, inoltre, menzionati gli approcci di calibrazione generale, che consentono di includere simultaneamente le informazioni ausiliarie, nonché le potenziali restrizioni per l'intervallo dei pesi, e producono le stime insieme ai corrispondenti calcoli della varianza³⁴. È opportuno ricordare che i pesi possono eliminare le distorsioni nelle variabili target, ma aumentano la varianza di campionamento. L'aumento della varianza si verifica in quanto l'introduzione dei pesi aggiunge complessità alle fasi di campionamento e dell'analisi dei dati, dal momento che ad ogni unità statistica sarà assegnato un peso differente e sarà, pertanto, necessario tenere in considerazione questo aspetto nella selezione del campione e nelle analisi statistiche. Non solo, la varianza di campionamento potrà aumentare anche in ragione del fatto che le unità con un peso

³¹ In genere, si tratta di variabili sociodemografiche, come età, sesso, regione, ma quando sono disponibili, si considerano anche altre variabili, ad esempio il consumo di media.

³² Si tratta di un metodo adottato dall'analisi statistica per stimare una popolazione o un valore non noto, che si basa sull'ordine degli elementi o delle unità che rientrano nel campionamento rispetto ad una certa variabile di interesse

³³ (Bethlehem and Biffignandi, Handbook of web surveys, 2012)

³⁴ La ponderazione lineare e la classificazione possono essere trattate come casi speciali di questi approcci generali di calibrazione (The SAGE Handbook of Survey Methodology, Vesja Vehovar, Vera Toepoel e Stephanie Steinmetz, Londra, 2016.)

elevato incideranno in misura proporzionalmente maggiore sui risultati finali rispetto all'effetto esercitato dalle unità con pesi inferiori. Per questo motivo maggiore è la discrepanza tra i pesi associati alle singole unità, maggiore sarà la variabilità del campione. Infine, un'ultima considerazione in merito all'aumento della variabilità può essere svolta riferendosi all'instabilità delle stime dei parametri, le quali saranno maggiormente sensibili alle variazioni nei dati o nei pesi, aumentando quindi la varianza. Tuttavia, l'aspettativa di fondo è che i guadagni nel ridurre i bias superino la corrispondente perdita dovuta all'aumento della varianza di campionamento³⁵. Inoltre, la varianza di campionamento e potenzialmente anche il bias potrebbero aumentare ulteriormente quando le informazioni ausiliarie non possono essere ottenute dalla popolazione, ma solo da un'indagine di riferimento. Come sottolineato inizialmente, tali procedure fanno sì riferimento ai campionamenti probabilistici, ma vengono impiegate anche per quelli non probabilistici. Con i campioni non probabilistici abbiamo spesso un solo passaggio, perché non possiamo separare i meccanismi di campionamento e di non risposta, a causa dell'assenza della randomizzazione e di informazioni sulla popolazione di riferimento per l'indagine.

2.3.1 Tecniche di campionamento non probabilistico: vantaggi e svantaggi

Come analogamente osservati per i campioni probabilistici, anche nel campionamento non probabilistico si osservano differenti tecniche, ognuna delle quali possiede i propri aspetti positivi e negativi. Tra le principali metodologie di campionamento non probabilistico si possono includere i panel online non probabilistici, il campionamento per quote, il campionamento per convenienza, il campionamento a scelta ragionata e il campionamento a valanga.

I panel³⁶ online non probabilistici sono forse l'esempio più frequente e più avanzato di campionamento non probabilistico contemporaneo e possono essere longitudinali oppure trasversali. I panel longitudinali raccolgono le informazioni di un certo gruppo di partecipanti nel corso del tempo, i quali vengono interrogati

³⁵ Questo non è necessariamente vero, soprattutto nel caso di correlazioni deboli tra variabili ausiliarie e target, dove i pesi possono non avere alcun impatto sulla rimozione del bias.

³⁶ Un panel è una tipologia di dataset che segue le stesse unità statistiche, che possono essere individui, famiglie, imprese o altre entità, nel corso del tempo.

ripetutamente in periodi di tempo successivi. Tale metodologia è molto utile per osservare nel dettaglio le dinamiche nel tempo e i cambiamenti nelle risposte da parte dei partecipanti, consente di aumentare l'efficienza nel reclutamento e riduce la variazione di campionamento³⁷, dal momento che è possibile osservare i cambiamenti a livello individuale. Dall'altra parte, i panel longitudinali possono far aumentare il rischio di bias di selezione, dal momento che gli individui che lo comporranno potrebbero auto-selezionarsi compromettendo la rappresentatività dei risultati, e il tasso di abbandono, visto che i partecipanti potrebbero abbandonare nel corso del tempo il panel portando ad un'erosione della rappresentatività per quanto riguarda il lungo termine. Quelli trasversali, invece, raccolgono le informazioni da un gruppo di individui osservati in un unico momento, per cui tale metodo non prevede un follow-up nel corso del tempo con gli stessi soggetti che sono stati osservati, ossia i partecipanti sono reclutati e interrogati una sola volta. I panel online non probabilistici trasversali, così come quelli longitudinali, permettono di raccogliere rapidamente i dati ad un costo contenuto, anche se si tratta di una tecnica che ha diversi limiti, specialmente nell'analisi delle tendenze nel corso del tempo, oltre a quelli già esaminati nell'ambito dei panel longitudinali. A fronte di tale distinzione, ai fini della trattazione si definisce il panel come un grande database di unità che inizialmente hanno accettato di collaborare, hanno fornito informazioni di base e sono state reclutate per essere selezionate occasionalmente o regolarmente in indagini campionarie specifiche su argomenti correlati o non correlati. Più precisamente, con panel online non probabilistico si fa riferimento ad un gruppo di individui che sono stati coinvolti nell'ambito di un sondaggio, di uno studio o di una ricerca online, la cui selezione non avviene in modo casuale. Nello specifico, la fase di reclutamento avviene in primo luogo con il reclutamento di unità nel panel dalla popolazione generale sulla quale occorre concentrarsi per conseguire gli obiettivi dell'indagine. In un secondo momento, quando le unità vengono reclutate dal panel al campione specifico, è necessario concentrarsi sulla selezione di una struttura ottimale utilizzando quote, stratificazione, *matching* e così via. È necessaria un'attenta gestione per dimostrare la rappresentazione dei sottogruppi essenziali in ogni campione, ma anche per controllare il carico di lavoro di

³⁷ Fa riferimento alla variabilità dei risultati che si ottengono selezionando un campione casuale dalla popolazione

ogni unità. Un sistema di gestione del panel deve quindi registrare e ottimizzare quando e come vengono reclutate le unità, quanto spesso partecipano ai sondaggi e quanto tempo impiegano per completare un sondaggio. Deve inoltre rilevare vari tipi di "imbroglianti", che possono essere tra l'altro reclutati da 'rispondenti professionali', che possono mantenere profili falsi per partecipare ai panel online. Dopo la raccolta dei dati, si procede a un'altra fase di ottimizzazione attraverso l'*editing*, l'imputazione e la ponderazione, che possono essere adattati in modo particolare a determinate circostanze. Tuttavia, dal momento in cui esistono differenti modalità di reclutamento, tra le quali le adesioni volontarie e le selezioni basate su criteri demografici o geografiche, possono sussistere differenze nelle stime. In genere, i metodi classici di campionamento probabilistico risultano essere maggiormente affidabili in termini di accuratezza e attendibilità delle stime e raramente fanno registrare una discrepanza significativa tra le stime e i valori di riferimento corrispondenti. Nonostante ciò, i metodi di campionamento non probabilistico vengono scelti di frequente, a scapito di quelli probabilistici, per via del loro costo contenuto e, soprattutto, perché non sempre la struttura probabilistica è preferibile ad una non probabilistica, come nel caso di popolazioni estremamente disperse e quindi difficili da raggiungere, per le quali potrebbe essere molto difficile da realizzare un campionamento probabilistico e per questo potrebbe essere più pratico impiegare una metodologia non probabilistica. È opportuno sottolineare l'importanza di controllare la qualità di un determinato panel, che può essere verificata monitorando la sua documentazione e i risultati ottenuti in passato.

Nel campionamento per quote i partecipanti sono selezionati in modo non casuale in base a una quota fissa o a una percentuale della popolazione basata su una o più caratteristiche. La quota selezionata può essere proporzionale o non proporzionale alla distribuzione effettiva della popolazione. Per ottenere un campione rappresentativo, un disegno di campionamento a quota proporzionale richiede una conoscenza a priori delle caratteristiche della popolazione. Il campionamento per quote non proporzionale è meno restrittivo, in quanto il ricercatore specifica la percentuale di dati da campionare da ciascun sottogruppo, indipendentemente dalle caratteristiche della popolazione. In prima istanza, il campionamento per quote prevede una divisione della popolazione di riferimento in sottogruppi esaustivi e mutuamente esclusivi. In secondo luogo, sarà

necessario scomporre la popolazione in strati sulla base di differenti criteri³⁸ e determinare la proporzione in ciascuna classe. Successivamente, verrà scelta la dimensione del campione e sarà selezionata una quota per ciascun sottogruppo, che può essere proporzionale o non proporzionale alla popolazione. Infine, si procederà alla raccolta dei dati fino a quando le quote, precedentemente definite, non saranno saturate. In tale metodologia, i ricercatori si dovranno assicurare che certi attributi, ritenuti importanti per la loro ricerca, siano rappresentati in modo proporzionale nel campione. Agli intervistatori dovranno essere indicate le quote del campione che dovranno essere in possesso delle caratteristiche predefinite, affinché il campione possa riproporre le proporzioni della popolazione oggetto d'esame. Le quote possono essere indipendenti, se gli intervistatori devono svolgere un determinato numero di interviste all'interno di un certo strato, oppure correlate, quando agli intervistatori sono assegnati dei campioni sistematici distribuiti in misura sistematica in base a differenti fattori di stratificazione. Viene effettuato ampio ricorso a tale metodologia nelle ricerche di mercato, in caso di assenza di una lista della popolazione, per via della sua facilità di somministrazione e della sua velocità nell'ottenere le informazioni. Il campionamento per quote offre diversi vantaggi, specialmente per la raccolta di dati primari, in quanto si tratta di una tecnica relativamente economica, rapida e semplice da mettere in pratica. Inoltre, i ricercatori possono anche assicurarsi che i dati siano raccolti da tutti i sottogruppi per un determinato insieme di caratteristiche, garantendo così che anche i gruppi più piccoli siano rappresentati nel campione. Un ulteriore aspetto positivo della tecnica per quote è rappresentato dalla sostituibilità delle unità statistiche non rispondenti, per cui un intervistatore potrà inserire all'interno della classe un altro individuo, con caratteristica analoghe, in caso di mancata risposta da parte di un partecipante precedentemente selezionato al fine di rispettare le quote prestabilite. Anche per queste ragioni, il campionamento per quote si configura come la procedura che concede all'intervistatore ampia libertà e flessibilità, nonché come quella più veloce da implementare e più economica in termini di denaro. Tuttavia, dal momento che il processo di selezione non è casuale e soggettivo, uno dei principali svantaggi del campionamento per quote consiste nella difficoltà nel determinare l'errore campionario e nell'effettuare inferenza

³⁸ Generalmente la suddivisione avviene per caratteristiche anagrafiche.

sulla popolazione generale a partire dal campione³⁹. Inoltre, la discrezionalità dell'intervistatore nella scelta degli intervistati aumenta il rischio di distorsioni, in quanto gli individui che non rispondono o non intendono fornire una risposta vengono trascurati e questo rappresenta un problema poiché, in molti casi, hanno comportamenti analoghi diversi da quelli registrati nella media della popolazione.

Il campionamento per convenienza, noto anche come campionamento per disponibilità, è un metodo in cui la selezione dei partecipanti o di altre unità di analisi si basa sulla loro disponibilità, la quale di solito è da intendersi in termini di vicinanza geografica, ma può coinvolgere altri tipi di accessibilità, come contatti conosciuti. Il campionamento per convenienza si differenzia dal campionamento per quote per il fatto che non cerca specificamente la rappresentatività. Anche questa tecnica presenta alcuni vantaggi pratici, come non richiedere un elenco esaustivo della popolazione oggetto di indagine e presentare chiari vantaggi logistici e di risorse in termini di viaggi, costi e tempo. Nonostante questi vantaggi, non si può non constatare la presenza di alcuni aspetti negativi come l'errore di campionamento e la sotto copertura. Errore di campionamento significa che il metodo di campionamento fornisce un campione le cui caratteristiche, ad esempio l'età o il livello di istruzione, differiscono sistematicamente da quelle della popolazione di interesse, mentre la sotto copertura implica che alcuni individui della popolazione di interesse rimangano esclusi dal metodo di campionamento. Sebbene il campionamento per convenienza presenti delle carenze metodologiche, queste possono essere attenuate descrivendo dettagliatamente le caratteristiche demografiche e di altro tipo del campione e, se possibile, confrontandole con quelle della popolazione di riferimento, in modo che i lettori dello studio possano valutarne la rappresentatività, ma anche sforzandosi di ottenere la partecipazione di tutti i partecipanti previsti, in modo che la distorsione della risposta o l'autoselezione non aggravino la mancanza di rappresentatività. Infine, si possono attenuare anche assicurandosi che i partecipanti reclutati siano teoricamente rilevanti per lo studio, in modo che la selezione non sia basata esclusivamente sulla convenienza.

³⁹ Qualora le difficoltà fossero tali da impedire una corretta analisi dei dati, si può passare al campionamento casuale stratificato, in quanto rappresenta l'alternativa basata sulla probabilità che introduce un elemento di casualità.

Per quanto riguarda le ultime due tipologie di campionamento, quello a scelta ragionata prevede che il *team* di ricercatori selezioni gli elementi sulla base della loro esperienza, in questo modo saranno in grado di ottenere un campione rappresentativo della popolazione, e viene applicata su campioni di piccole dimensioni, generalmente selezionati con l'aiuto di esperti del settore. Mentre il campionamento a valanga consiste nello scegliere una prima unità in modo casuale e nel selezionare quelle successive sulla base delle informazioni che vengono fornite dalla prima unità. Sebbene la prima unità venga selezionata mediante un procedimento casuale, il campionamento a valanga rientra nelle tecniche non probabilistiche poiché le unità successivamente selezionate saranno simili a coloro che hanno svolto il ruolo dei referenti dal punto di vista demografico e di altre caratteristiche. Si tratta di un metodo particolarmente adatto per popolazioni di piccole dimensioni, specializzate e difficili da reperire e, in particolare, permette di individuare all'interno della popolazione quelle unità che possiedono le caratteristiche di interesse ai fini dell'indagine che si sta svolgendo.

CAPITOLO III. Gli errori

3.1 L'errore complessivo del questionario.

Prima di procedere ad effettuare una classificazione delle tipologie di errore, occorre specificare cosa si intende per errore. Con errore si fa riferimento a quella discrepanza o deviazione dalla qualità ottimale ricercata nell'ambito delle fasi di progettazione, di amministrazione e di interpretazione dei dati raccolti per mezzo di un questionario. Tali errori possono derivare da fonti differenti, come i problemi legati alla formulazione delle domande, gli errori di struttura del questionario, bias ed errori nella misurazione dei dati ottenuti. Si può, dunque, intuire come questi errori possano compromettere la validità e l'affidabilità dei dati raccolti, condizionando le conclusioni e le decisioni che si fondano sui dati analizzati. Quindi, al fine di garantire la qualità e l'affidabilità, è necessario riconoscere, comprendere e cercare di ridurre l'effetto degli errori nei sondaggi nell'ambito delle fasi di progettazione, di somministrazione e di analisi dei dati. È doveroso, però, riconoscere che questa nozione di errore rappresenta solamente una delle numerose fonti che possono influenzare negativamente la qualità e l'affidabilità dei dati raccolti. Per questa ragione si introduce il concetto di errore totale del sondaggio, *total survey error* (TSE), che può essere definito come l'insieme degli errori che possono verificarsi nelle fasi di ricerca, di identificazione del campione e dell'analisi dati, descrive le proprietà di errore statistico delle statistiche di indagine campionaria. In altri termini, l'errore totale del sondaggio è un concetto che pretende di descrivere le proprietà statistiche delle stime dell'indagine, incorporando una varietà di fonti di errore. Gli errori del sondaggio possono derivare dalle carenze della struttura del sondaggio, dal processo di campionamento, dalle interviste e dagli intervistatori, dagli intervistati, dai dati mancanti, dai processi di codifica⁴⁰ e di digitazione e dalla fase di

⁴⁰ Rispetto all'accezione con il quale era stato osservato precedentemente il processo di codifica nell'ambito degli aspetti cognitivi delle risposte, in questo caso la codifica è quella fase in cui viene assegnato un codice numerico, oppure simbolico, alle risposte per facilitare l'analisi dei dati. Questa fase assume particolare rilevanza quando si ha a che fare con dati qualitativi o con risposte aperte, dal momento che rende possibile la trasformazione di questi dati in dati quantitativi che possono essere elaborati con tecniche statistiche. Inoltre, permette di standardizzare le risposte e agevola l'analisi dei dati.

*editing*⁴¹. La presenza di un qualsiasi errore all'interno del questionario è problematica, in quanto diminuisce l'accuratezza⁴² delle inferenze derivate dai dati dell'indagine, per cui uno stimatore di indagine sarà accurato se avrà un bias e una varianza ridotti, il che si verificherà solamente quando il *total surveys error* ha un'influenza ridotta sulla stima. Il paradigma dell'errore totale del questionario fa parte del concetto molto più ampio di qualità totale dell'indagine, che considera l'idoneità all'uso di una stima. Il concetto di idoneità all'uso, introdotto da Juran e Gryna nel 1980, riconosce che i produttori e gli utilizzatori dei dati delle indagini spesso percepiscono la qualità dell'indagine da prospettive molto diverse. Infatti, i produttori attribuiscono un'alta priorità alla qualità dei dati, ad esempio, ampie dimensioni del campione, alto tasso di risposta, risposte coerenti, buona copertura del territorio, risposte coerenti tra loro, buona copertura della popolazione target, e possono destinare un'ampia parte del budget dell'indagine per raggiungere un elevato livello di accuratezza per alcune stime chiave. Mentre gli utenti dei dati spesso danno per scontata l'accuratezza e attribuiscono una priorità maggiore ad attributi quali la tempestività, l'accessibilità e la fruibilità dei dati, oltre al contenuto del questionario che è altamente rilevante per i loro obiettivi di ricerca. Queste due prospettive suggeriscono che la qualità dell'indagine è un concetto complesso e multidimensionale che va oltre al *total surveys error*. Si identificano due aspetti distinti del concetto generale di qualità, ovvero l'assenza di carenze e la capacità di rispondere alle esigenze di coloro che commissionano l'indagine. Per la maggior parte delle indagini, l'assenza di carenze è coerente con il paradigma dell'errore totale del sondaggio, mentre la capacità di rispondere alle esigenze di coloro che commissionano la ricerca può essere raggiunta solo dando un'adeguata enfasi nella progettazione dell'indagine agli attributi che porteranno ad un'elevata soddisfazione del soggetto che ha commissionato l'indagine. Da qui si può intuire l'importanza di una progettazione ottimale del questionario, che consenta di conseguire gli obiettivi di accuratezza dei dati e di credibilità metodologica che il produttore intende perseguire. Per poter ottenere una progettazione ottimale del questionario è necessario ottimizzare la qualità totale

⁴¹ Con questo termine si fa riferimento a quella fase di revisione dei dati raccolti, nella quale si cerca prima di individuare e poi di correggere eventuali errori, incoerenze o ambiguità. L'obiettivo di questo procedimento è quello di garantire la qualità e l'integrità dei dati raccolti, in modo che siano pronti per essere passati al vaglio nelle fasi di codifica e analisi.

⁴² Con il termine accuratezza si può definire lo scostamento di una stima del questionario dal valore reale del parametro sottostante.

dell'indagine, ovvero la capacità complessiva di quanto un questionario sia valido, affidabile e accurato nel restituire al ricercatore informazioni sulla popolazione di riferimento, e per raggiungerla occorre che gli obiettivi di qualità per ogni dimensione⁴³ siano specificati e che gli approcci per raggiungere tali obiettivi siano progettati e preventivati in modo ottimale. Quindi, l'obiettivo di una progettazione ottimale dell'indagine può essere definito semplicemente come la minimizzazione del *total surveys error*, soggetta a vincoli di costo e di tempestività che siano coerenti con le altre dimensioni. È necessaria un'attenta pianificazione per allocare le risorse alle varie fasi del processo di indagine in modo da controllare le principali fonti di errore a livelli accettabili, per cui l'obiettivo non è quello di condurre ogni fase del processo di indagine il più possibile esente da errori, perché ciò comporterebbe il superamento del budget e del calendario dell'indagine di un margine considerevole. Siccome anche nelle migliori circostanze, con un budget e un tempo illimitati, in alcune operazioni rimarrà sempre la possibilità di errori di indagine, allora l'obiettivo sarà quello di evitare gli errori più gravi e di controllare gli altri errori nella misura in cui gli errori residui siano per lo più irrilevanti e tollerabili. In ogni caso, si può affermare che la risoluzione degli errori più gravi attraverso una progettazione oculata dell'indagine è facilitata dalle suddivisioni dell'errore in componenti sempre più piccole, come quella che vede la scomposizione dell'errore totale del questionario in errore campionario ed errore non campionario. A sua volta, l'errore di campionamento può essere suddiviso in fonti di errore attribuibili allo schema di campionamento, come il campionamento a più stadi o a più fasi, alla dimensione del campione e alla scelta dello stimatore, ad esempio nel caso di uno stimatore di rapporto o di regressione. Mentre l'errore non campionario può essere ulteriormente scomposto in errore di specificazione, errore di struttura, errore di non risposta, errore di misurazione ed errore di elaborazione. Va sottolineato come per individuare meglio le fonti di errore specifiche si potrebbe rendere necessario procedere ad una scomposizione ulteriore di entrambe le tipologie di errore dell'indagine. Tuttavia, gli errori campionari e quelli non campionari, seppur definiti brevemente,

⁴³ Nell'ambito del design di un questionario esistono differenti dimensioni che vengono considerate dal ricercatore, come l'accuratezza, l'accessibilità, la credibilità, la rilevanza e così via. A queste dimensioni corrispondono specifici obiettivi, che se conseguiti consentono di ottimizzare la qualità del sondaggio. Tra gli obiettivi sopracitati, si possono trovare la minimizzazione dell'errore totale di un questionario, la necessità di soddisfare le esigenze dei clienti e la possibilità di facilitare agli utenti l'accesso ai dati.

saranno approfonditi nel prosieguo dell'elaborato di tesi. Premesso ciò, va detto che nella realizzazione di un questionario ci si trova quasi sempre in presenza di vincoli pratici, come la mancanza di risorse, per cui la progettazione delle indagini non è vincolata solo da competenze sostanziali e metodologiche, ma si assiste ad un compromesso tra errori e costi. Maggiori risorse, cioè costi più elevati, riducono in genere la maggior parte degli errori, per cui con un determinato livello di risorse, si vogliono fare le scelte progettuali ottimali per ridurre al minimo l'errore totale del questionario. Quindi, è necessario che coloro che vogliono condurre l'indagine decidano quanto spendere per ogni elemento, come la dimensione del campione, il tasso di risposta, lo sviluppo degli item e il pre-testing, la formazione degli intervistatori e la pulizia dei dati. In generale, l'obiettivo sarebbe quello di spendere le risorse disponibili in modo da ridurre al minimo l'errore totale di un questionario e di fare compromessi tra le diverse allocazioni a seconda di quali siano più efficaci dal punto di vista dei costi per ridurre il *total surveys error*. Purtroppo, di solito si hanno poche informazioni precise sul livello di riduzione dell'errore che una determinata spesa potrebbe ottenere o sulla riduzione netta dell'errore che si otterrebbe scambiando una componente con un'altra. Inoltre, le informazioni sui costi dei vari disegni sono minime e l'impatto dei costi dei vari disegni di studio può essere stimato solo in modo approssimativo. Mentre molti cambiamenti nel disegno dell'indagine possono essere preventivati in modo ragionevolmente accurato, ad esempio il costo di una maggiore dimensione del campione piuttosto che la spesa sostenuta per una formazione più lunga per gli intervistatori, raramente si conoscono le variazioni di errore che ne derivano. La minimizzazione dell'errore non dipende solo dalla progettazione, ma anche dall'esecuzione. Infatti, aldilà di alcune indagini che possono essere progettate male, altre sono progettate in modo ragionevole, ma non vengono condotte in modo adeguato. Pertanto, si rendono necessarie procedure di controllo della qualità per garantire che le indagini siano effettivamente condotte come previsto. Sebbene ognuna delle componenti sia una fonte di errore, alcune componenti dovrebbero portare a una riduzione netta degli errori, ad esempio, mentre l'editing e la pulizia dei dati possono creare nuovi errori a causa di una correzione errata e possono non correggere un errore esistente, se eseguiti correttamente l'editing e la codifica dovrebbero ridurre l'errore totale del questionario. Anche la convalida dell'intervistatore e lo sviluppo dello

strumento tramite test preliminari sono altre componenti che dovrebbero ridurre l'errore complessivo di un sondaggio. Infatti, nella fase di convalida dell'intervistatore, quest'ultimo viene formato per formulare correttamente le domande, per spiegare agli intervistati il funzionamento del questionario, per registrare le risposte in modo accurato e per mantenere un comportamento idoneo, ovverosia neutrale e professionale, nel corso dell'intervista. Naturalmente, il fatto di avere intervistatori che vengano addestrati verso queste finalità riduce la probabilità di errori riconducibili alla loro condotta. Invece, per quanto riguarda lo sviluppo dello strumento tramite test preliminari, questo procedimento implica la somministrazione di un questionario preliminare o pilota ad un campione rappresentativo della popolazione prima del lancio del sondaggio definitivo. Così facendo, i ricercatori hanno la possibilità di identificare e correggere problematiche o ambiguità all'interno del questionario prima della sua distribuzione su larga scala. Nonostante ciò, la complessità delle indagini e la mancanza di informazioni su molte componenti dell'errore fanno sì che ridurre al minimo l'errore complessivo di un questionario non sia facile e che spesso le decisioni di progettazione debbano basarsi su dati imperfetti e incompleti.

In ultima analisi, occorre accennare che in taluni casi i progettisti dei sondaggi possono utilizzare l'errore totale dell'indagine come criterio di pianificazione, per cui tra un insieme di progetti alternativi il progetto che fornisce l'errore totale di indagine minore, per un dato costo fisso, dovrebbe essere scelto. In questo modo, si riesce a massimizzare l'efficienza del questionario e a migliorare la qualità dei dati, poiché una riduzione dell'errore totale di un questionario comporta un miglioramento della qualità complessiva, aumentando la validità e l'affidabilità dei dati raccolti. Non solo, la selezione di un'alternativa di sondaggio con un *total surveys error* minore consente di ottimizzare l'utilizzo delle risorse a disposizione e di aumentare la precisione delle stime.

3.2 Errori Campionari

Una prima tipologia di errore che si può registrare all'interno di un questionario è rappresentata dall'errore campionario, che si verifica nel corso della fase di campionamento di una popolazione ed indica la discrepanza tra la struttura di campionamento prevista e quella effettivamente realizzata durante il processo di selezione del campione. Quanto descritto si verifica dal momento che si decide di prendere in considerazione una parte della popolazione di riferimento e non la totalità di quest'ultima. In altri termini, questa deliberata non osservazione delle unità statistiche introduce una deviazione dalle statistiche del campione ottenuto e dalla stessa statistica sull'intera struttura di campionamento, per cui quando si parla di errore campionario si fa riferimento alla differenza tra la popolazione che si intende studiare nell'indagine e il campione che viene studiato nella pratica. Il particolare campione selezionato per un'indagine è solo uno dei tanti possibili campioni che avrebbero potuto essere selezionati, per cui le stime, ad esempio le medie o le proporzioni, di ciascun campione possono quindi variare da un campione all'altro solo casualmente. Quindi, si tratta di una situazione in cui viene estratto un campione dalla popolazione per realizzare delle inferenze sulla popolazione, in cui il campione in questione è solamente uno dei tanti campioni possibili che possono essere estratti dalla popolazione e in cui le caratteristiche demografiche e le altre caratteristiche che possono influenzare la partecipazione o il ritiro da uno studio non sono probabilmente le stesse per il campione e per la popolazione. Esistono due tipologie di errore di campionamento, ossia la distorsione del campionamento e la varianza di campionamento.

La distorsione campionaria si verifica quando ad alcuni membri del campione non viene data alcuna possibilità, o quanto meno una possibilità ridotta, di essere selezionati e, in un disegno di questo tipo, ogni possibile serie di selezioni li esclude sistematicamente. Questa componente dell'errore di campionamento deriva da una fonte sistematica che fa sì che le stime di campionamento, mediate su tutte le realizzazioni del campione, differiscano in modo consistente dai valori reali della popolazione target. Mentre, come si avrà modo di vedere successivamente, la varianza di campionamento può essere controllata attraverso caratteristiche progettuali come la dimensione del campione, la stratificazione e il raggruppamento, si deve ricorrere ad altri metodi per controllare e

ridurre il più possibile la distorsione. La distorsione campionaria può essere misurata solo se abbiamo accesso ai corrispondenti valori della popolazione. Naturalmente, ci si potrebbe domandare perché effettuare un campionamento se tali informazioni fossero disponibili, siccome non avrebbe senso estrarre un campione e realizzare un sondaggio. Tuttavia, ci sono situazioni in cui si rende necessario approssimare i bias di campionamento confrontando le informazioni sottostanti, come i dati demografici di base del campione, con i dati corrispondenti di un'altra fonte più affidabile, come un censimento o un'ampia indagine nazionale, per identificare le aree dello spazio dati in cui il campione potrebbe essere sottorappresentato o sovra rappresentato. Una delle principali fonti di bias di campionamento è la copertura del *frame*, cioè il *frame* da cui viene estratto il campione è difettoso in quanto non include tutti gli elementi della popolazione. Si tratta di un errore grave perché non può essere rilevato, in alcuni casi il suo impatto non può nemmeno essere misurato e questo problema viene definito sottocopertura, perché il campione manca di elementi che dovrebbe contenere. Il fenomeno opposto, la sovracopertura, è meno grave, dal momento che si verifica quando il *frame* include elementi estranei, ossia elementi che non appartengono alla popolazione target. Tuttavia, questi elementi, se campionati, possono essere identificati durante le operazioni sul campo ed esclusi da ulteriori elaborazioni. Una terza potenziale fonte di distorsione del *frame* è la duplicazione. Se alcuni elementi compaiono più volte nel *frame*, le loro probabilità di selezione sono più alte e quindi potrebbero essere sovrarappresentati nel campione. Inoltre, non sempre si sa quante volte gli elementi compaiano nel *frame*, nel qual caso è impossibile accertare la portata del problema e quindi l'entità del bias. I bias di campionamento possono anche essere il risultato di difetti nel processo di selezione del campione, di errori nell'implementazione del campione e di errori di programmazione durante la fase di elaborazione del campione. Infine, i bias di campionamento o di stima possono verificarsi durante la fase di elaborazione del campione, ad esempio attraverso un calcolo errato dei fattori di aggiustamento della ponderazione, dando un'importanza eccessiva a determinate sottopopolazioni. È molto difficile misurare la distorsione. Infatti, mentre la stima della varianza di campionamento deriva dalla teoria statistica, l'unico modo per misurare la distorsione di campionamento è confrontare il valore empirico risultante con il vero valore della popolazione target. Naturalmente, questo è problematico perché raramente

si è in possesso del valore della popolazione e quindi si rende necessario usare metodi indiretti per stimare il bias, ad esempio un approccio utilizza dati provenienti da altre fonti, come il censimento o grandi campioni nazionali, come surrogati della popolazione campionata. Il problema di questa strategia è che anche il censimento è soggetto a errori, sia in termini di varianza che di bias. Sebbene sia stato appena sottolineato di quanto la ponderazione tenda ad aumentare la varianza di campionamento e a ridurre la precisione, in alcuni approcci proposti dalla letteratura si implementa la ponderazione perché in molti casi può essere utilizzata per ridurre la distorsione portando le distribuzioni di campionamento più in linea con le distribuzioni note della popolazione. La ponderazione potrebbe ridurre le distorsioni in quanto può compensare alcuni segmenti della popolazione sottorappresentati andando a correggere i bias generati dal processo di campionamento, ma anche perché sono in grado di aumentare il grado di rappresentatività del campione e di migliorare la precisione delle stime, garantendo che quest'ultime siano più vicine ai valori reali presenti all'interno della popolazione di riferimento.

La varianza campionaria, invece, deriva dal fatto che, dato il disegno del campione, potrebbero essere estratti molti insiemi diversi di elementi del *frame*, ragion per cui si può affermare che il concetto statistico di varianza campionaria viene influenzato dalle modalità di selezione del campione. Quando si utilizza un campione probabilistico, la variabilità casuale delle stime campionarie può essere misurata dagli errori standard delle stime associate a una particolare indagine. Quest'ultima affermazione è spiegata dal fatto che quando si effettua un'analisi statistica di un campione probabilistico diventa importante capire quanto le stime ottenute da quel determinato schema di campionamento varino intorno al valore reale del parametro della popolazione che si intende misurare e gli errori standard restituiscono informazioni circa la dispersione delle stime attorno al vero valore del parametro. Si potrebbe, dunque, definire la varianza di campionamento come la parte che può essere controllata da fattori di progettazione campionaria come la dimensione del campione, le strategie di raggruppamento, la stratificazione e le procedure di stima. La varianza di campionamento è la componente casuale dell'errore di campionamento, poiché deriva dalla "fortuna del sorteggio" e dagli elementi specifici della popolazione inclusi in ciascun campione. È importante sottolineare come disegni inefficienti che non

impiegano alcuna stratificazione o ne impiegano una debole daranno luogo a campioni e stime che fluttuano ampiamente, mentre se il disegno incorpora strategie di stratificazione efficaci e una clusterizzazione minima, è possibile avere campioni le cui stime sono molto simili, generando così una bassa varianza tra le stime e raggiungendo così alti livelli di precisione di campionamento. In generale, i principali fattori che determinano la varianza di campionamento sono la stratificazione e il clustering. La stratificazione di solito determina una varianza di campionamento più bassa perché il numero di campioni possibili è ridotto rispetto a un campione casuale semplice non ristretto. Non solo si riduce il numero di campioni possibili, ma si eliminano anche i potenziali outlier. Nella misura in cui la variabile di interesse è correlata alle variabili di stratificazione, quest'ultima ridurrà la varianza complessiva del campionamento. Nell'impostazione del disegno, quindi, è importante cercare di definire strati relativamente omogenei rispetto alle variabili di interesse. Il clustering, invece, funziona in modo molto diverso. La clusterizzazione svolge un ruolo importante nei disegni campionari utilizzati per le indagini in cui i dati vengono raccolti di persona e viene utilizzata per controllare i costi sul campo, in particolare quelli relativi ai viaggi, che spesso rappresentano una parte significativa del budget complessivo dell'indagine. Tuttavia, ciò comporta un minor numero di gradi di libertà, nel senso che il campione si concentra su un numero inferiore di unità di campionamento, ovvero i cluster del primo stadio, spesso definiti unità di campionamento primarie. Se si limita il campione ad un numero di cluster inferiori e prendendo poi un numero piuttosto contenuto all'interno di ciascun cluster, si riducono i costi di viaggio perché magari i ricercatori concentrano i propri sforzi in un'area geografica circoscritta, ma si riduce la capacità di distribuire efficacemente il campione sull'intera popolazione. Questa riduzione dell'efficienza è ulteriormente aggravata dal fatto che all'interno di ogni cluster le unità tendono ad assomigliarsi di più rispetto agli altri membri della popolazione. Questo fenomeno, chiamato omogeneità intraclasse, tende a far aumentare la varianza di campionamento. Quindi, in sintesi, per quanto riguarda l'ottimizzazione del disegno campionario, la stratificazione è vantaggiosa in quanto riduce la varianza di campionamento, mentre il raggruppamento deve essere evitato quando possibile o almeno ridotto al minimo in quanto il suo effetto è quello di aumentare la varianza di campionamento, anche se in molte situazioni il raggruppamento è necessario per motivi di costo. Di solito l'effetto

del raggruppamento è più marcato di quello della stratificazione. Pertanto, la migliore strategia di progettazione a raggruppamento consiste nel trovare un compromesso tra il risparmio sui costi e un minor livello di precisione. Un altro fattore importante che influenza la varianza del campionamento è la ponderazione, che si riferisce ai fattori di aggiustamento che tengono conto delle deviazioni del progetto, come le probabilità diseguali di selezione, i tassi variabili di non risposta e l'inevitabile introduzione di distorsioni in varie fasi del processo di indagine, che vengono corrette attraverso un processo chiamato post-stratificazione. L'effetto della ponderazione è quello di aumentare la varianza del campionamento e l'entità di questo aumento è proporzionale alla varianza tra i pesi. Un concetto utile per quantificare e riassumere l'impatto della stratificazione, del raggruppamento e della ponderazione sulla varianza del campionamento è l'effetto del disegno. Si tratta del rapporto tra la vera varianza del campionamento, tenendo conto di tutte le complessità del disegno, e la varianza che si sarebbe ottenuta se il campione fosse stato estratto con un semplice campione casuale, senza stratificazione, raggruppamento o ponderazione.

In sintesi, l'errore di campionamento complessivo deve essere visto in termini di combinazione di varianza e distorsione del campione. Se nel caso della varianza campionaria i ricercatori hanno la possibilità di misurarla e, di conseguenza, di ridurla, la distorsione campionaria rappresenta una sfida maggiore, in quanto è spesso difficile da misurare e, anche se è misurabile, la riduzione della distorsione è spesso costosa e problematica da ottenere. L'obiettivo finale, comunque, sarà quello di ridurre al minimo l'errore quadratico medio⁴⁴.

⁴⁴ Con errore quadratico medio si intende fare riferimento a quella misura che descrive la differenza tra i valori osservati e quelli previsti o stimati da un modello.

3.3 Errori non campionari

L'errore non campionario fa riferimento agli errori che si possono verificare nell'ambito delle fasi di raccolta, elaborazione o interpretazione delle informazioni, non attribuibili alla casualità del campionamento. Si potrebbe semplificare il ragionamento affermando che l'errore non campionario spiega tutte le altre fonti di errore, tra cui le mancate risposte, i bias di selezione, l'errore di misura, o di risposta imprecisa, e così via, che sono correlate alla somministrazione del questionario. Inoltre, può derivare dalle caratteristiche del compito di indagine stesso, dalle prestazioni dei rispondenti e dalle persone coinvolte nel processo di intervista. A differenza della categoria precedentemente osservata, si fanno afferire all'errore non campionario numerose tipologie di errore, tra cui gli errori di mancata risposta, quelli di misurazione e altri che verranno trattati successivamente nel corso della trattazione. Indipendentemente da quanto riportato, è bene prestare la massima accortezza riguardo il fatto che nei dati delle indagini ci sarà sempre un certo grado di errore non campionario e che non è mai possibile eliminare tutte le fonti di errore. Inoltre, misurare l'effetto degli errori non campionari è estremamente difficile, in quanto essi possono avere un effetto enorme ma spesso sconosciuto su stime importanti, ad esempio come le percentuali di disoccupati, i tassi di crescita dell'economia e molte altre misure. Un'altra peculiarità propria dell'errore non campionario è costituita dal fatto che l'errore non campionario può verificarsi in qualsiasi fase del processo di indagine e ha molte fonti, ad esempio dalla mancata volontà degli intervistati di rispondere a una o a tutte le domande del sondaggio, da errori che si verificano durante l'elaborazione dei dati o da risposte errate fornite dagli intervistati o registrate dagli intervistatori. Un'altra difficoltà fondamentale dell'errore non campionario è data dalla quasi impossibilità di calcolare un valore numerico dell'effetto di quello che spesso è un errore non campionario molto grande, poiché il valore "vero" di una stima non è quasi mai noto. È opportuno specificare che le indagini sono soggette sia ad errori campionari sia ad errori non campionari, il che è molto significativo perché influenzano l'accuratezza delle stime, ma mentre per i primi si riesce a misurarne l'entità, il valore dei secondi è sempre sconosciuto. Questo rappresenta un punto importante perché per i ricercatori diventa di fondamentale importanza comprendere la presenza di eventuali errori non campionari, come gli errori di segnalazione, la varianza delle risposte, la distorsione dell'intervistatore e del

rispondente, le mancate risposte, gli errori di imputazione e gli errori nell'elaborazione dei dati devono essere considerati con attenzione. Rispetto agli errori campionari, questi errori possono aumentare con l'incremento della dimensione del campione, dal momento che aumenteranno le complessità legate alla raccolta dati e dei modelli statistici oppure possono verificarsi situazioni particolare come l'effetto scala⁴⁵. Pertanto, gli errori non campionari possono risultare più dannosi dell'errore di campionamento per le indagini su grandi campioni. Come accennato in precedenza, l'errore non campionario comprende tutti gli errori di ricerca diversi dall'errore di campionamento ed esistono diversi modi per classificare i vari tipi di errore non campionario, come l'approccio a due classi generali, che individua due tipologie di errori, ovvero gli errori di specificazione del modello ed errori di misurazione. Per quanto riguarda gli errori di specificazione questi derivano dall'incapacità del modello di descrivere in maniera corretta i legami tra le differenti variabili, mentre gli errori di misurazione possono dipendere da errori casuali o sistematici nel corso della fase di misurazione e si verificano quando si registrano problematiche con la precisione. Un'altra classificazione proposta dalla letteratura è quella che suddivide l'errore non campionario in tre classi: errore di cornice, errore di misura ed errore di risposta. L'errore di cornice è considerato un errore casuale, mentre l'errore di misurazione e quello di risposta possono essere errori casuali o sistematici. L'errore di cornice è associato solo alla ricerca campionaria, fa parte del processo di selezione e mantenimento di una cornice, o lista, da cui vengono estratte le unità campionarie. L'errore di misura e di risposta si verificano sia nelle indagini censuarie sia nelle indagini campionarie e nei panel. L'errore dell'intervistatore si verifica solo nelle indagini campionarie. L'errore di cornice si verifica nei disegni campionari ed è più spesso considerato un errore sistematico; comprende errori nella specificazione della popolazione, nella selezione del campione e nella selezione e manutenzione della cornice. Questi producono errori causati dall'utilizzo di un elenco o di una fonte imprecisa o incompleta da cui viene estratto il campione. L'errore di misurazione si verifica quando viene registrato un valore o una risposta errata invece di un valore vero e rappresenta la differenza tra un valore vero e un valore misurato, indipendentemente

⁴⁵ L'effetto scala è quel fenomeno per cui in relazione ad un aumento della numerosità campionaria possono diventare più evidenti alcune distorsioni nel contesto della raccolta dei dati.

dal modo o dal momento in cui si verifica nel processo di indagine. L'errore di non risposta è dovuto alla mancanza di un'unità di campionamento o all'omissione di risposte. Oltre a queste due categorie, si può anche descrivere l'errore dell'intervistatore, che si verifica quando un intervistatore, per caso o accidentalmente, registra un valore falso per una risposta ed è costituito da dati errati raccolti involontariamente o intenzionalmente da intervistatori personali durante il processo di indagine. L'errore risultante può essere casuale o sistematico. Gli errori casuali sono differenze nei dati misurati dovute al processo di rilevamento, possono verificarsi in qualsiasi direzione e possono essere valutati attraverso l'analisi statistica e ridotti, ma non eliminati, calcolando la media su un numero elevato di osservazioni. Gli errori di misurazione sistematici sono coerenti nella stessa direzione, producono una misurazione distorta e non possono essere rilevati o ridotti aumentando la dimensione del campione. L'errore di risposta è la differenza tra un valore vero e un valore errato fornito da un intervistato, si verifica quando gli intervistati non danno la loro opinione onesta su un argomento, ma quando frainendono la domanda e danno una risposta irrilevante, oppure quando mentono per qualsiasi motivo, quando cercano di compiacere l'intervistatore o quando cercano di fare bella figura.

Finora si è proceduto a fornire una visione d'insieme dell'errore non campionario, volta a dare una prima descrizione delle sue caratteristiche e, in particolare, delle sue componenti. Data la rilevanza che le singole metodologie di errore possiedono nell'economia di un questionario e in ragione del fatto che per i ricercatori, o più in generale per chiunque voglia condurre un'indagine campionaria, diventa di cruciale importanza saper riconoscere i singoli errori per poterne controllare in parte gli effetti, occorre fornire una descrizione maggiormente dettagliata delle singole fattispecie accennate in precedenza. Prima di procedere con l'esame delle singole tipologie di errore non campionario, bisogna specificare che l'elaborato di tesi seguirà uno degli approcci proposti dalla letteratura (Non sampling error in social surveys, Davide E. McNabb, Los Angeles, 2014) basato sulla suddivisione delle quattro categorie individuate, ovverosia l'errore dell'intervistatore, l'errore di cornice, l'errore di misura e l'errore di risposta, in ulteriori sottoclassi. Tanto premesso, durante il processo di indagine campionaria possono verificarsi tre classi di errori di cornice: errore di specificazione della popolazione, errore di selezione dell'unità campionaria ed errore di

copertura. L'errore di specificazione prevede la selezione di un campione di soggetti dalla popolazione sbagliata. Proprio per questo motivo può essere il più dannoso, dal momento che la presenza di tale errore rischia di portare i ricercatori ad un'interpretazione distorta dei risultati, senza considerare la totale mancanza di affidabilità del modello. Di conseguenza, per quanto le restanti parti del processo di indagine siano prive di errori, i dati raccolti sono privi di significato poiché sono stati intervistati i soggetti sbagliati. Questo accade spesso perché non si riconosce che gli elenchi di riferimento, così come i soggetti che includono, cambiano nel tempo e quindi devono essere gestiti e rivisti regolarmente. L'errore di selezione può derivare da diversi errori di gestione. In primo luogo, è necessario assicurarsi che vengano scelti soggetti che conoscano l'oggetto dell'indagine, poi sarà necessario assicurarsi che le caratteristiche etniche e le differenze linguistiche note siano prese in considerazione quando si specificano i soggetti. Invece, gli errori di copertura comprendono la sotto-copertura, la sovra-copertura e la copertura multipla o duplicata. La sotto-copertura si verifica quando alcune unità della popolazione non sono incluse nella popolazione campione, mentre il problema di riconoscere la sotto-copertura è che non può essere identificata dal campione o dal frame, ma sono necessarie fonti esterne per effettuare l'identificazione e la mancata correzione della sotto-copertura porta a una sottostima del totale. La sovra-copertura può comportare l'inclusione di unità non eleggibili o la duplicazione di unità nel campione. Nello specifico, la letteratura (Rao, 2005) suggerisce che la sotto-copertura è un problema più grave della sovra-copertura, dal momento che diventa difficile aggiungere alcuni dati mancanti e che le inferenze svolte sulla base dei dati raccolti dal campione non sono corrette.

L'errore di misurazione si verifica in diverse fasi del processo di indagine e si riferisce ai molti modi in cui i risultati effettivi della ricerca possono variare dai risultati reali. Questo tipo di errore si verifica quando viene acquisito un valore errato o una risposta incompleta rispetto al valore vero attribuito o assegnato a un determinato item, rispondente, unità campionaria o gruppo. Gli errori di misurazione sono modelli comportamentali idiosincratici che gli intervistati seguono durante le interviste. Quest'ultimi possono derivare da tre fonti, che sono il questionario, il rispondente e il processo di raccolta dati. Facendo riferimento al questionario, gli errori di misura possono derivare da parole incomprensibili che formano le domande, dalla lunghezza e

dalla forma dei quesiti, dall'ordine delle risposte e dall'argomento delle domande. Gli effetti sulle informazioni raccolte dipendono dalle tecniche di intervista, per cui ci saranno conseguenze differenti sui risultati a seconda che la rilevazione sia stata condotta con una modalità di intervista personale, con una modalità telefonica oppure con un questionario autosomministrato. Spesso vengono identificati durante un'analisi a posteriori dei dati e confrontando le prime indagini con i risultati delle re-interviste. I modelli di risposta improbabili sono quelle combinazioni di variabili, o la loro mancanza, che hanno una probabilità sproporzionata o improbabile di essere riportate in una popolazione, ma che possono verificarsi in dati di sondaggi falsificati o errati. È importante notare che le variazioni rispetto alla norma di questi risultati possono servire come indicatori di un potenziale errore del sondaggio. Due strumenti utilizzati per individuare la potenziale fonte di errore sono le re-interviste e il data mining⁴⁶.

Per quanto concerne l'errore di risposta, la letteratura prevalente identifica tre sottoclassi, che sono l'errore del rispondente, l'errore di non risposta e l'errore dell'intervistatore. Le tre sottoclassi di errore di risposta sono l'errore del rispondente, l'errore di non risposta e l'errore dell'intervistatore. L'errore del rispondente si verifica quando gli intervistati forniscono dati errati o incompleti, mentre le mancate risposte si verificano quando gli intervistatori non riescono ad ottenere una risposta a uno o più item di un sondaggio o quando gli intervistati rifiutano di partecipare. L'errore dell'intervistatore, invece, si verifica quando un intervistatore non segue il protocollo del sondaggio, registra risposte false, risponde falsamente a sondaggi completi, recluta consapevolmente rispondenti al sondaggio come sostituti o pone domande guida che influenzano le risposte nei modi preferiti dall'intervistatore. Dei tre, l'errore di non risposta è considerato la principale fonte di dati mancanti nella ricerca sui sondaggi ed è quindi una delle principali cause di distorsione nei risultati dei sondaggi. I ricercatori considerano l'errore di non risposta abbastanza importante da essere considerato singolarmente come una delle tre principali classi di errori non campionari. In particolare, l'errore di non risposta si riferisce all'errore che si verifica quando le persone o altri elementi inclusi in un campione non rispondono a un sondaggio. Esistono due tipi di non risposta alle indagini: la non risposta unitaria e la non risposta

⁴⁶ Con questo termine si intende quell'insieme di tecniche e metodi volti ad estrarre informazioni da grandi quantità di dati, come banche dati, *data warehouse*, *software* e così via.

per tematiche. Le mancate risposte per tematiche, ovvero i dati mancanti, si verificano quando un intervistato che completa un sondaggio non rispondendo ad una domanda, mentre le mancate risposte unitarie, invece, si verificano quando gli intervistati non rispondono affatto a un sondaggio. Quest'ultima fattispecie si può verificare per una serie di motivi, come il caso di alcuni intervistati che non sono in grado di completare un sondaggio per ragioni collegate alle condizioni di salute o perché parlano una lingua diversa da quella in cui viene somministrato il sondaggio oppure perché non sono disponibili a completare un sondaggio, ad esempio perché non sono in casa quando l'intervistatore chiama. Il pericolo principale legato alle mancate risposte è la potenziale distorsione dei risultati. Analogamente all'errore di copertura, l'entità dell'errore di non risposta dipende sia dalla percentuale di popolazione che non risponde a una determinata domanda, sia dalle differenze tra rispondenti e non rispondenti su qualsiasi statistica del sondaggio. In sintesi, le ragioni degli errori di non risposta sono molteplici e tra queste sono incluse, tra l'altro, l'incomprensione di una domanda, il rifiuto di rispondere a una domanda imbarazzante o incriminante, la mancata restituzione di un questionario spedito per posta, la mancata presenza in casa degli intervistati al momento della telefonata o della visita, l'impossibilità di raggiungere un intervistato qualificato, l'esaurimento del tempo a disposizione, gli elementi inavvertitamente mancati o saltati e la confusione sulle indicazioni di diramazione. I responsabili delle indagini incontrano sempre maggiori difficoltà nell'assicurarsi la partecipazione allo studio da parte degli intervistati e si sono concentrati maggiormente sulla ricerca di modi per gestire le mancate risposte e i ricercatori di sondaggi dedicano grandi sforzi per migliorare l'impatto dei bias causati dalle mancate risposte, indipendentemente dalla causa.

Infine, per quanto riguarda gli effetti dell'errore dell'intervistatore, questo può avere un impatto significativo sulla volontà degli intervistati di partecipare a un sondaggio e sul modo in cui gli intervistati rispondono alle domande, fornendo risposte false a sondaggi completi o a singoli item, l'errore dell'intervistatore può contribuire sia al bias che alle componenti di varianza dei sondaggi. Le stime possono essere distorte quando gli intervistatori influenzano i loro intervistati a scegliere un'alternativa di risposta che non è corretta, quando nascondono informazioni rilevanti quando pongono una domanda o quando non registrano alcune delle informazioni rilevanti fornite da un intervistato. Gli intervistatori hanno la capacità di modellare la varianza attraverso interviste non

standard, come quelle che si riscontrano nell'uso di domande aperte o quando gli intervistatori elaborano o abbreviano le risposte degli intervistati in reazione ai sondaggi dell'intervistatore. È stato dimostrato che un sondaggio errato, o l'incapacità di sondare quando è appropriato, determina tassi più elevati di errori nel sondaggio. Nel corso del tempo i ricercatori di sondaggi hanno troppo spesso considerato l'intervistatore come poco più di un prolungamento neutrale del questionario, utile solamente per compilare i questionari. Tuttavia, al giorno d'oggi un numero crescente di ricercatori riconosce ora che gli effetti dell'intervistatore sono una parte importante dell'errore di misurazione, che a sua volta è una componente importante dell'errore totale del questionario e richiede un maggiore sforzo di controllo. A volte l'influenza negativa dell'intervistatore sulla qualità dei dati è molto più forte, in particolare nel caso in cui alcuni intervistatori saltino parti del questionario o le compilino da soli. In questo modo, possono accorciare notevolmente il tempo di intervista, ma le risposte inventate da questi influiranno negativamente sulla qualità dei dati. Uno dei metodi per correggere, quantomeno per ridurre, l'effetto dell'errore dell'intervistatore è rappresentato dai sondaggi assistiti da computer oppure dai sondaggi autosomministrati via Internet, perché vengono, di fatto, eliminati gli errori di registrazione e perché viene eliminata ogni possibile forma di influenza interpersonale. L'adozione diffusa di strumenti computerizzati per la raccolta dei dati e il caricamento quotidiano dei dati delle interviste sul campo tramite modem o altri mezzi rende possibile l'analisi dei dati mentre vengono raccolti e questo è un aspetto importante in quanto prima inizia l'analisi dei dati, tanto prima si possono individuare indizi di potenziale falsificazione.

Due dei metodi più utilizzati dai ricercatori per identificare e controllare gli errori non campionari sono le re-interviste e il *data mining*. Quindi, uno dei metodi più efficaci per rilevare e controllare gli errori non campionari è quello di seguire un approccio di re-intervista in due parti, dove la prima parte prevede una re-intervista di intervistatori selezionati a caso e di alcuni dei soggetti che hanno risposto a tali intervistatori, mentre la seconda parte consiste in una re-intervista supplementare condotta dai supervisori degli uffici sul campo su specifici intervistati. Lo scopo della re-intervista mirata è quello di identificare in modo più efficace gli intervistatori che falsificano i dati utilizzando metodi e strumenti, come il controllo statistico dei processi, che focalizzano la re-intervista sugli intervistatori che fanno eccezione. La re-intervista mirata utilizza

carte di controllo simili a quelle utilizzate nella gestione della qualità totale, queste permettono di identificare gli intervistatori i cui dati si distinguono per una determinata variabile o i cui dati mostrano uno schema di risposte insolite per più di una variabile. La re-intervista è stata ritenuta necessaria perché l'inserimento dei dati nell'intervista personale assistita da computer rende impossibile per i supervisori vedere il lavoro dell'intervistatore. Le re-interviste per il controllo di qualità sono utilizzate sia nelle indagini che nei disegni di raccolta dati panel, possono svolgersi più o meno in concomitanza con il programma di indagine o dopo il completamento della rilevazione. In entrambi i casi, la re-intervista ha due obiettivi chiave, ossia correggere i dati già raccolti che sono falsi o errati e aggiungere altri dati a quelli raccolti durante l'intervista originale. Per la re-intervista vengono selezionati campioni casuali delle indagini completate da ogni intervistatore, i casi selezionati sono solitamente basati sul livello di esperienza dell'intervistatore e, infatti, le re-interviste sono solitamente condotte da supervisori o da intervistatori esperti. Per quanto riguarda, invece, il *data mining* è definito come l'esplorazione e l'analisi di insiemi di dati molto grandi per scoprire modelli e regole e comporta l'uso di modelli statistici esplorativi e predittivi. Durante il *data mining* possono essere scoperti gli effetti delle preferenze delle cifre, dell'accumulo dei dati, delle anomalie di combinazione, dei modelli di risposta improbabili e di altre potenziali fonti di errore non campionario. Le tecniche di *data mining* sono state utilizzate in un modo e circostanza limitate per l'analisi di insiemi molto ampi di dati raccolti tramite sondaggio. Poiché l'obiettivo di gran parte del data mining è la previsione, il data mining predittivo è la forma più comune ed il processo si articola in tre fasi: raccolta dei dati, identificazione dei modelli e verifica e applicazione del modello a nuovi dati per generare previsioni. Il data mining di un grande insieme di dati con l'analisi dei cluster o dei fattori è un metodo per identificare modelli inizialmente sconosciuti nelle risposte. Questi schemi possono poi essere confrontati con un'analisi simile di un sottoinsieme di risposte, come le risposte raccolte da un singolo intervistatore o da un sottogruppo distinto. Le anomalie identificate durante l'analisi possono essere utilizzate come indicatori di dati potenzialmente falsificati o errati. Il data mining, infatti, prevede tecniche come l'analisi dei cluster di tutte o di un campione molto ampio di interviste completate per identificare modelli insoliti e anomalie di associazione. Le anomalie di associazione sono discontinuità nei valori di correlazione,

cioè indicano un'associazione quando intuitivamente non dovrebbe esistere, o dovrebbe esistere poco, o il contrario. Lo studio su un ampio set di dati è stato suggerito come un modo per evitare i problemi associati ai tradizionali approcci di pre-testing e di intervista cognitiva su piccoli campioni utilizzati per identificare potenziali errori di segnalazione e soluzioni. L'obiettivo del data mining è quello di elaborare i dati, trovare modelli e relazioni sconosciuti riassumendo o comprimendo i dati in modo che siano informativi e utili. La successiva analisi di questi schemi e connessioni può portare a una maggiore comprensione dei fattori e delle conseguenze sottostanti, l'estrazione di informazioni dai dati originali può anche portare a una rappresentazione accurata della popolazione di interesse. Il data mining permette di riassumere i dati per fare inferenze statistiche o affermazioni sulla popolazione da cui i dati sono stati estratti e può anche rivelare modelli che presentano interessanti opportunità di analisi. L'analisi dei pattern può portare alla scoperta di scostamenti non normali dalla distribuzione generale o di tendenze nei dati che altrimenti non sarebbero evidenti. I risultati del data mining possono includere la scoperta di associazioni tra due o più variabili separate con una relazione lineare insolitamente forte, una combinazione di variabili che hanno una correlazione estremamente elevata con una certa variabile, o un raggruppamento nei dati che aiuta a identificare caratteristiche importanti nelle variabili. Descrivendo un programma commerciale di data mining, si può intuire perché e come il data mining può essere utilizzato per identificare potenziali variabili predittive. Infatti, nella modellazione predittiva, è importante identificare le variabili per determinare determinate relazioni distributive nell'insieme di dati al fine di generare osservazioni future o persino scoprire modelli insoliti e identificare osservazioni insolite nell'insieme di dati che sono ben al di là della tendenza generale del resto degli altri punti di dati. Questi dati vengono poi analizzati con software statistici di data mining, utilizzando strumenti come l'analisi dei cluster o l'analisi dei fattori. Si formano quindi dei modelli per l'intero set di dati, le risposte di ogni singolo intervistatore possono poi essere confrontate con l'insieme dei dati per identificare gli indicatori più forti e le differenze significative. Infine, è possibile utilizzare un'analisi discriminante per misurare il potere predittivo degli indicatori selezionati per prevedere l'appartenenza a gruppi di errore o di non errore.

In estrema sintesi, gli errori non campionari sono tutti gli altri errori che possono confondere i risultati del sondaggio e insieme agli errori campionari costituiscono l'errore totale del questionario. Entrambe le classi di errore possono ridurre significativamente l'utilità dei risultati di un sondaggio e la gestione dell'errore totale del sondaggio consiste nel gestire tutti gli errori che possono influire sui sondaggi. L'errore di campionamento esiste sempre in una certa misura quando si utilizza un campione anziché un censimento e l'aumento della dimensione del campione è il metodo per ridurre l'errore di campionamento, anche se non può essere eliminato del tutto. Invece, una delle principali fonti di errore non campionario è rappresentata dall'intervistato e dall'intervistatore ed è definito in molti modi diversi, ma in genere è causato da uno o da entrambi i fattori: errore di non risposta ed errore di risposta. L'errore non campionario può essere parte del compito dell'indagine o derivare dalle prestazioni degli intervistati o degli intervistatori e il compito delle indagini statistiche è la raccolta di dati accurati, affidabili, validi e precisi. A differenza degli errori di campionamento, essi possono aumentare con gli errori di dimensione del campione ed essere più dannosi degli errori di campionamento per le indagini su famiglie di grandi dimensioni. I ricercatori di sondaggi si concentrano in generale su quattro grandi classi di errori non campionari: errore di cornice, errore di misurazione, errore di non risposta ed errore dell'intervistatore. L'errore di cornice si verifica nella progettazione del campione ed è solitamente considerato un errore sistematico; comprende gli errori di specificazione della popolazione e di selezione del campione, nonché gli errori generali di cornice, ovvero gli errori causati dall'utilizzo di un elenco o di una fonte imprecisa o incompleta da cui viene estratto il campione. L'errore di misurazione si verifica quando viene registrato un valore o una risposta errata invece di un valore vero, mentre l'errore di non risposta deriva da un'unità di campionamento mancante o da omissioni di risposta. Infine, l'errore dell'intervistatore si verifica quando un intervistatore, per caso o accidentalmente, registra un valore falso per una risposta e delle quattro grandi classi, l'errore dell'intervistatore è quello che sembra ricevere le minori preoccupazioni.

CAPITOLO IV. Analisi dei dati

4.1 Metodi di analisi dati

L'analisi dei dati è parte integrante della ricerca qualitativa e costituisce un passaggio essenziale per la raccolta dei dati. Esistono molte varianti della ricerca qualitativa che comportano molte forme di analisi dei dati, tra cui le trascrizioni delle interviste, le note sul campo, l'analisi delle conversazioni e i dati visivi, che si tratti di fotografie, filmati o osservazioni di eventi su Internet. In primo luogo, la raccolta dei dati e la loro analisi sono processi iterativi, non a caso, nella sua forma ideale, l'analisi precoce dei dati fornisce una visione sufficiente per modellare la raccolta di altri dati. Infatti, possono essere svolte delle analisi preliminari per rivedere le linee guida da osservare durante il prosieguo dell'intervista o per porre maggior accortezza sulle interviste future. In secondo luogo, sia durante che dopo la raccolta dei dati, i ricercatori si dedicano alla fase di memorizzazione, anche nota come *memoing*. La memorizzazione avviene quando i ricercatori prendono nota di idee o riflessioni personali, concettuali o teoriche che vengono loro in mente durante la raccolta e l'analisi dei dati. Le prime annotazioni possono avvenire mentre i ricercatori scrivono le note sul campo o trascrivono le interviste, in questo caso i ricercatori inseriscono i promemoria all'interno del testo stesso. Questi promemoria possono contenere ipotesi dei ricercatori su ciò che sta accadendo, domande sollevate dai dati o collegamenti alla letteratura che possono essere utili per aiutare i ricercatori a interpretare i dati e a metterli in relazione con altre aree sostanziali che potrebbero coinvolgere processi sociali generici simili. Va detto che alcuni ricercatori trovano utile incorporare i promemoria nelle note di campo o nelle trascrizioni delle interviste, mentre altri trovano più proficuo creare i promemoria separatamente oppure utilizzare i promemoria stessi come materiale per la codifica. Volendosi soffermare sulla codifica, è opportuno dire che qualsiasi analisi dei dati comporta una forma di codifica. Quest'ultima riflette sia le abitudini analitiche personali dei ricercatori sia i principi generali che derivano da particolari metodologie di ricerca qualitativa e prospettive teoriche. Nella sua forma più rigorosa, l'analisi dei dati può comportare la codifica riga per riga del testo, in cui i ricercatori catturano ogni occorrenza empirica e concettuale. Nell'analisi delle interviste personali, persino la durata delle pause viene misurata e utilizzata come dato. Per ragioni legate

principalmente al risparmio di tempo e di risorse, non viene messa in pratica una forma di codifica rigorosa, per cui è molto più frequente osservare come i ricercatori adottino un approccio più flessibile, magari codificando interi paragrafi o gruppi di frasi alla volta. Si potrebbe concludere che la codifica a livelli successivi consenta di ottenere l'analisi migliore e, in questo caso, i ricercatori indicano innanzitutto le istanze empiriche che sembrano essere correlate all'oggetto dello studio. Questa scansione del testo da parte dei ricercatori costituisce il primo e il secondo processo di familiarizzazione con le istanze empiriche. L'analisi, tuttavia, rimarrebbe un esercizio banale se il *team* che sta volgendo l'indagine, alla ricerca di concetti di ordine superiore, si fermassero al livello della mera descrizione, invece di effettuare scansioni successive. Infatti, a prima vista, i dati sembrano spesso opachi. Tuttavia, attraverso il processo di lettura multipla del testo, i ricercatori acquistano fiducia nel fatto che esso contenga abbastanza materiale da giustificare la scoperta e l'analisi, passando dalle osservazioni empiriche. Spesso, durante le prime letture del materiale, i ricercatori sviluppano un elenco di codici preliminari da provare, mentre man mano che la codifica prosegue, i ricercatori possono affinare i codici per includere categorie più rilevanti. Dopodiché, tutta l'analisi dei dati deve essere finalizzata allo sviluppo di concetti o al collegamento con concetti già esistenti.

Tanto premesso, occorre soffermarsi sull'analisi vera e propria dei dati, osservandone anche le metodologie principali. Le variabili quantitative che si distribuiscono normalmente vengono analizzate con test parametrici. I due prerequisiti fondamentali per l'analisi statistica parametrica sono rappresentati dall'ipotesi di normalità, che specifica che le medie del gruppo campione sono normalmente distribuite, e dall'ipotesi di varianza uguale, che prevede che le varianze dei campioni e della popolazione corrispondente siano uguali. Quindi, i test parametrici presuppongono che i dati siano su scala quantitativa, detta anche numerica, con una distribuzione normale della popolazione sottostante e che i campioni abbiano la stessa varianza, per cui si sia in presenza di omogeneità delle varianze. I campioni sono estratti casualmente dalla popolazione e le osservazioni all'interno di un gruppo sono indipendenti l'una dall'altra. I test parametrici comunemente utilizzati sono il test t di Student, l'analisi della varianza (ANOVA) e l'ANOVA a misure ripetute. Il t-test di Student viene utilizzato per verificare l'ipotesi nulla che prevede non vi siano differenze tra le medie

dei due gruppi. Viene utilizzato in tre circostanze: per verificare se la media di un campione⁴⁷ differisce significativamente da una determinata media della popolazione, in questo caso si parlerà di *one-sample* t-test; per verificare se le medie della popolazione stimate da due campioni indipendenti differiscono in modo significativo, per cui si può definire come t-test non accoppiato; e per verificare se le medie della popolazione stimate da due campioni dipendenti differiscono in modo significativo, qui si farà riferimento al t-test a coppie. Il test t, però, non può essere utilizzato per il confronto di tre o più gruppi, per cui per osservare se esiste una differenza significativa tra le medie di due o più gruppi si impiega il metodo ANOVA. L'ANOVA studia due varianze, la varianza tra i gruppi, che rappresenta le differenze tra le medie dei gruppi, e la varianza all'interno dei gruppi, che descrive la variabilità dei dati presenti all'interno di ogni singolo gruppo. Dunque, per verificare se le differenze tra le medie dei gruppi presi in considerazione sono significative, queste due stime di varianza vengono confrontate con il test F di Fisher, nel seguente modo: $F = \frac{MS_b}{MS_w}$, dove MS_b è la media dei quadrati tra i gruppi e MS_w è la media dei quadrati all'interno dei gruppi. Per osservare se l'ipotesi nulla, ossia che non esiste differenza tra le medie dei gruppi considerati, è verificata, sarà necessario che il test F di Fisher restituisca un valore inferiore al *p-value* calcolato. In caso contrario, un risultato del test F più elevato suggerisce una probabilità più elevate che le medie dei gruppi presi in considerazione differiscano tra loro, dando maggior evidenza statistica all'ipotesi alternativa, secondo la quale almeno una delle medie è diversa. Invece, l'ANOVA a misure ripetute analizza l'uguaglianza delle medie di tre o più gruppi e viene utilizzata quando tutte le variabili di un campione sono misurate in condizioni diverse o in momenti diversi e siccome le variabili sono misurate da un campione in momenti diversi, la misurazione della variabile dipendente verrà ripetuta.

Tuttavia, se la distribuzione del campione è inclinata verso un lato o se la distribuzione è sconosciuta a causa delle ridotte dimensioni del campione, si ricorre a tecniche statistiche non parametriche, le quali vengono impiegate per analizzare dati ordinali e categorici. Quindi, quando le ipotesi di normalità non sono soddisfatte e le medie campionarie non sono normali, i test parametrici distribuiti possono portare a risultati

⁴⁷ intesa come stima della media della popolazione.

errati. Perciò, vengono utilizzati i test non parametrici perché non richiedono l'ipotesi di normalità. Nonostante ciò, i test non parametrici possono non rilevare una differenza significativa rispetto ai test parametrici, per cui si può affermare che di solito hanno una potenza inferiore rispetto a quelli parametrici. Anche in questo caso, come per i test parametrici, la statistica di prova viene confrontata con i valori noti della distribuzione campionaria di quella statistica e l'ipotesi nulla viene accettata o rifiutata. Tra i test non parametrici più diffusi si possono trovare il *sign test*, il test di Wilcoxon, il test di Kolmogorov-Smirnov e il test di Kruskal-Wallis. Il *sign test* esamina l'ipotesi sulla mediana θ_0 di una popolazione, quindi vuole dimostrare l'ipotesi nulla $H_0 = \theta_0$. Quando il valore osservato, X_i , è maggiore del valore di riferimento θ_0 , viene contrassegnato con il segno +. Invece, se il valore osservato è inferiore al valore di riferimento, viene contrassegnato con il segno -, mentre se il valore osservato è uguale al valore di riferimento (θ_0), viene eliminato dal campione. Se l'ipotesi nulla è verificata, ci sarà un numero uguale di segni + e di segni -. Il *sign test* ignora i valori reali dei dati e utilizza solo i segni + o -, per cui è utile quando è difficile misurare i valori. Per queste ragioni, si può affermare come il *sign test* presenti alcune limitazioni, ovvero perde le informazioni quantitative dei dati e si limita ad impiegare i segni + o -. Il test di Wilcoxon non solo esamina i valori osservati rispetto a θ_0 , ma prende anche in considerazione le dimensioni relative, aggiungendo maggiore potenza statistica al test. Come nel test dei segni, se c'è un valore osservato che è uguale al valore di riferimento θ_0 , questo valore osservato viene eliminato dal campione. Il test di Wilcoxon ordina tutti i punti dati, calcola la somma di rango di ogni campione e confronta la differenza delle somme di rango. Invece, il test di Kolmogorov-Smirnov (KS) a due campioni è stato concepito come metodo generico per verificare se due campioni casuali sono estratti dalla stessa distribuzione. L'ipotesi nulla del test KS è che le due distribuzioni siano identiche, la statistica del test è una distanza tra le due distribuzioni empiriche, calcolata come la massima differenza assoluta tra le loro curve cumulative. Infine, il test di Kruskal-Wallis è un test non parametrico per analizzare la varianza, analizza se esiste una differenza tra i valori mediani di tre o più campioni indipendenti. I valori dei dati vengono classificati in ordine crescente e vengono calcolate le somme dei ranghi, seguito dal calcolo della statistica del test.

4.2 Affidabilità e validità dei dati raccolti.

L'affidabilità e la validità sono idee distinte ma complementari che descrivono l'utilità di uno strumento di misurazione, per cui comprendere il significato di questi termini è fondamentale per poter valutare la bontà di un questionario, in particolare delle scale a più voci. In termini generali, l'affidabilità è definita come la misura in cui una rilevazione è priva di errori, ovvero si riferisce al grado di precisione di una misurazione qualora questa fosse ripetuta più volte. In questo modo, l'affidabilità restituisce un'approssimazione della coerenza della misurazione nel caso in cui quest'ultima dovesse essere replicata. L'affidabilità informa i ricercatori sulla quantità relativa di incoerenza casuale o di fluttuazione non sistematica delle risposte individuali a una misura ed è un elemento di prova necessario per sostenere la validità costruttiva della misura, anche se una misura affidabile non è necessariamente una misura valida. Pertanto, l'affidabilità di qualsiasi misura stabilisce un limite superiore alla validità del costruito. Da un punto di vista teorico, l'affidabilità si potrebbe definire partendo dalla relazione della varianza, che viene descritta come $\sigma^2 X = \sigma^2 T + \sigma^2 E$, dove la varianza totale è spiegata dalla somma della variazione legata ai trattamenti⁴⁸ e dalla variazione residua. Quando $\sigma^2 E$ aumenta, le differenze osservate su una misura riflettono sia le vere differenze di punteggio tra gli intervistati sia gli errori casuali. Quando $\sigma^2 X = \sigma^2 E$, tutte le differenze osservate riflettono solo errori casuali di misurazione, il che suggerisce che le risposte alla misura sono completamente casuali e inaffidabili. Se si dividono entrambi i membri dell'equazione precedente per $\sigma^2 X$, la relazione di varianza può essere modificata in $1 = \frac{\sigma^2 T}{\sigma^2 X} + \frac{\sigma^2 E}{\sigma^2 X}$. Da qui, si può definire l'affidabilità come $\frac{\sigma^2 T}{\sigma^2 X}$, e può essere interpretata come la correlazione al quadrato tra i punteggi osservati e quelli veri ($\rho^2 XT$). Concettualmente, l'affidabilità è interpretata come la proporzione della varianza del punteggio osservato che è spiegata o giustificata dalla varianza del punteggio vero. Esistono una serie di coefficienti in grado di stimare l'affidabilità, alcuni, detti "classici", basati sul punteggio reale, tra i quali l'affidabilità test-retest⁴⁹ e

⁴⁸ In altre parole, $\sigma^2 T$ indica le differenze medie tra i gruppi che sono stati presi in considerazione.

⁴⁹ L'affidabilità del test-retest è l'indicatore più comunemente utilizzato per valutare l'affidabilità di uno strumento di indagine. Si tratta di una misura della riproducibilità di una serie di risultati. Nella pratica, i ricercatori di sondaggi testano uno strumento di indagine facendo completare lo strumento agli stessi intervistati in due momenti diversi, per vedere quanto siano stabili le loro risposte. Quindi calcolano i

l'affidabilità di consistenza interna, e altri “alternativi”, come l'affidabilità delle forme alternate, l'affidabilità dei punteggi di differenza e l'affidabilità dei punteggi compositi. Il coefficiente di affidabilità delle forme alternative, detto anche coefficiente di equivalenza, viene stimato calcolando la correlazione tra i punteggi osservati per due forme simili che hanno medie, deviazioni standard e correlazioni con le forme simili. In genere, un gruppo di intervistati riceve due moduli in ordine casuale in un periodo di tempo molto breve, anche se il periodo di tempo tra le somministrazioni può essere più lungo. Va detto che i ricercatori sono spesso interessati a qualsiasi cambiamento nei punteggi dei test pre e post un determinato intervento, che può essere rappresentato, ad esempio, dalla modifica dell'ordine delle domande nell'ambito di un questionario. La differenza tra i punteggi della stessa misura o di una misura comparabile in due momenti diversi viene spesso definita punteggio di cambiamento o punteggio di differenza, indicato con la lettera D. L'affidabilità dei punteggi di differenza in un campione può essere stimata da:

$$r_{DD'} = \frac{S_X^2 r_{XX'} + S_Y^2 r_{YY'} - 2S_X S_Y r_{XY}}{S_X^2 + S_Y^2 - 2S_X S_Y r_{XY}}$$

, dove S_X^2 e S_Y^2 sono le varianze dei test pre e post, $r_{XX'}$ e $r_{YY'}$ sono le affidabilità dei test e r_{XY} è la correlazione tra i test. Se le varianze di entrambe le misure sono uguali, l'equazione precedente può essere riscritta come:

$$r_{DD'} = \frac{\frac{r_{XX'} + r_{YY'}}{2} - r_{XY}}{1 - r_{XY}}$$

Dato che $r_{XX'}$ e $r_{YY'}$ sono costanti, un r_{XY} più forte si traduce in un $r_{DD'}$ più debole. Inoltre, $r_{DD'}$ si avvicina a zero quando r_{XY} si avvicina all'affidabilità media delle misure. In realtà, la dimensione di r_{XY} è spesso forte a causa della somiglianza tra le due misure. Di conseguenza, i punteggi di differenza tendono ad avere una bassa affidabilità, anche quando entrambe le misure sono relativamente affidabili. Nello stimare l'affidabilità di una misura, i ricercatori non possono separare le caratteristiche dei rispondenti da quelle della misura. Le definizioni sopra riportate dimostrano chiaramente che le stime di affidabilità basate sulla teoria classica del punteggio reale cambiano a seconda del

coefficienti di correlazione per confrontare le due serie di risposte. Questi coefficienti di correlazione sono indicati collettivamente come affidabilità test-retest dello strumento d'indagine e, in generale, i coefficienti di correlazione sono considerati buoni se sono pari almeno a 0,70.

gruppo di partecipanti al test e degli *item* del test. In particolare, le procedure tradizionali per stimare l'affidabilità dipendono dal gruppo e dal test, il che suggerisce che l'errore standard di misura, e di conseguenza l'affidabilità, di una misura non è lo stesso per tutti i livelli di rispondenti. L'evidenza empirica ha dimostrato che l'errore standard di misura tende a essere più piccolo per i punteggi estremi rispetto a quelli moderati. Tuttavia, esistono altre caratteristiche delle misure che influenzano l'affidabilità, in primo luogo, quest'ultima tende ad aumentare quando aumenta la lunghezza della misura stessa, a condizione che vengano inclusi gli *item* appropriati:

$$\hat{\rho}_{XX'} = \frac{n * \rho_{XX'}}{1 + (n - 1) \rho_{XX'}}$$

in cui n è il fattore di allungamento del test originale $\rho_{XX'}$, e $\hat{\rho}_{XX'}$ sono le stime di affidabilità rispettivamente per il test allungato e per il test originale. In secondo luogo, i risultati empirici hanno dimostrato che i coefficienti di affidabilità tendono a crescere quando aumentano le categorie di risposta di una misura. Esistono evidenze dalla letteratura che hanno rivelato che le stime di affidabilità complessiva di una misura a 10 *item*, come ad esempio l'affidabilità test-retest, aumentano quando le categorie di risposta passano da una scala a 2 punti a una scala a 5 punti e anche il contenuto del test e il tempo a disposizione possono influire sull'affidabilità. Infatti, quando i limiti di tempo sono brevi, la maggior parte degli intervistati non può rispondere a tutti gli *item* e, di conseguenza, le loro risposte tra gli *item* diventano artificialmente coerenti. Infine, l'affidabilità può essere influenzata dalle procedure che i ricercatori scelgono di utilizzare. Come descritto in precedenza, i coefficienti di affidabilità basati sulla teoria classica del punteggio reale stimano uno o due tipi di errori di misurazione. Per esempio, la coerenza interna stima le variazioni di risposta attribuite all'eterogeneità degli *item* o all'incoerenza del contenuto del test, mentre l'affidabilità test-retest stima le variazioni attribuite ai cambiamenti nel tempo, compresi gli effetti di *carry over*⁵⁰. Al contrario, l'affidabilità a forme alterne con un lungo lasso di tempo stima due tipi di errori casuali, ovvero quelli attribuiti ai cambiamenti nel tempo e all'incoerenza del contenuto del test. Sebbene tutti i coefficienti di affidabilità basati sulla teoria classica del punteggio reale stimino la quantità relativa di errori di misurazione, gli errori stimati

⁵⁰ Con effetti di *carry over* si intende far esplicito riferimento alla situazione in cui l'influenza di una determinata variabile si estende o continua a rimanere costante durante i periodi successivi.

dai diversi coefficienti di affidabilità sono qualitativamente diversi tra loro. Pertanto, non è logico confrontare due diversi tipi di stime di affidabilità. Il fatto che errori casuali diversi siano stimati da coefficienti di affidabilità diversi ha un'implicazione importante quando i ricercatori correggono l'attenuazione dovuta all'inaffidabilità. Derivata dalle ipotesi della teoria classica del punteggio vero, una correlazione di popolazione tra due punteggi veri ($\rho_{T_X T_Y}$) può essere influenzata negativamente dall'inaffidabilità delle misure (X e Y). L'effetto negativo può essere dedotto dalla relazione tra la correlazione della popolazione e l'affidabilità della popolazione delle misure X e Y ($\rho_{XX'}$ e $\rho_{YY'}$), viene definita come: $\rho_{T_X T_Y} = \frac{\rho_{XY}}{\sqrt{\rho_{XX'} \rho_{YY'}}$.

D'altra parte, la validità è definita come la capacità di uno strumento di rappresentare correttamente il fenomeno che vuole misurare, in altre parole, indica se una misura è in grado di descrivere accuratamente ciò che si propone di misurare. In altre parole, secondo questa visione, la validità è una proprietà dei test piuttosto che delle interpretazioni. Infatti, la validità è verificata semplicemente se l'affermazione "il test X misura l'attributo Y" è vera o meno. Per essere vera, richiede che Y esista e che le variazioni in Y causino variazioni in X (Borsboom et al., 2004). Questa definizione può essere definita la visione della validità del test e trova ampi precedenti nei testi di psicometria. È necessario documentare la validità quando si valutano nuovi strumenti di indagine o quando si applicano strumenti di indagine consolidati a nuove popolazioni. La validità è una misura importante dell'accuratezza di uno strumento di indagine. Quindi, l'unico modo in cui un ricercatore può ottenere una misura corretta della validità di una misurazione è convalidare un'interpretazione dei dati derivanti da una procedura specifica. La chiara implicazione è che uno strumento di misurazione può essere relativamente valido per misurare un fenomeno ma non valido per misurare altri fenomeni, per cui non viene validato lo strumento di misura in sé, quanto lo strumento di misura in relazione allo scopo per cui viene utilizzato. Tuttavia, sebbene sia riconosciuta dalla maggior parte degli addetti ai lavori, come psicologi e statistici, la grande importanza che riveste la validità, non esiste una definizione che ne faciliti un'interpretazione univoca. In particolare, esistono due prospettive differenti che si prepongono l'obiettivo di definire il concetto di validità. Il primo approccio è rappresentato dalla definizione di validità che è stata appena analizzata, mentre il

secondo primo approccio definisce la validità non come una proprietà dei test, ma come una proprietà delle interpretazioni dei punteggi dei test e, pertanto, può essere definito come campo dell'interpretazione o validità di costrutto. La validità del costrutto si basa su una sintesi e un'analisi delle prove che supportano una certa interpretazione dei punteggi dei test, quindi, la validità è una proprietà delle inferenze interpretative sui punteggi dei test, in particolare l'interpretazione del significato dei punteggi. Poiché il contesto di misurazione influisce sui punteggi dei test, i risultati di qualsiasi sforzo di validazione sono condizionati dal contesto e dalle caratteristiche del gruppo con cui gli studi sono stati condotti, così come le affermazioni di validità tratte da questi risultati empirici. Contrariamente a quanto successo in passato, la letteratura odierna è piuttosto concorde, nonostante le divergenze riguardo la definizione del concetto di validità, nell'affermare che non esistano differenti tipologie di validità. Infatti, la letteratura prevalente risalente agli ultimi anni del Novecento era orientata sull'individuazione di differenti tipologie di validità, tra le quali la validità predittiva, la validità concorrente e la validità di criterio. Attualmente, la tendenza è quella di considerare una sola validità e le diverse procedure di validazione, mentre i diversi tipi di prove contribuiscono a fare inferenze sul significato dei punteggi. Quanto fatto negli ultimi anni dai ricercatori è andato nella direzione di semplificare l'interpretazione del concetto di validità, dal momento che, all'inizio degli anni duemila, esistevano circa centoventi sottotipi distinti di validità (Newton e Shaw, 2013). Questa visione aveva come conseguenza quella di consolidare l'idea che esistessero tipi indipendenti di validità, che comportassero procedure analitiche diverse per produrre tipi corrispondenti di prove, i quali a loro volta corrispondevano a diverse categorie di inferenza (Messick, 1989). Pertanto, la chiave di lettura proposta dalla letteratura contemporanea è quella di non pensare alla validità discriminante o a tipologie distinte di validità, ma di utilizzare un'ampia varietà di disegni di ricerca e di analisi statistiche per fornire potenzialmente prove che possano o meno supportare una determinata inferenza in esame. Questo dimostra che pensare a tipologie di validità può essere inutilmente restrittivo, perché induce i ricercatori a pensare alla validità come a un concetto frammentato, portando a conseguenze negative nella pratica della validazione.

4.3 Le scale di misurazione e la *cluster analysis*

Uno dei processi più diffusi tra i ricercatori è rappresentato dalla classificazione degli oggetti in categorie. Per classificare un oggetto come membro di una categoria, è necessario avere una comprensione esplicita o implicita di come gli oggetti possano essere simili o diversi tra loro. Quindi, è proprio in questo momento che è necessario introdurre due concetti significativi, ovvero quello di *scaling* e di *cluster analysis*. Con il termine *scaling* si fa riferimento alla fase di assegnazione, o di categorizzazione, di valori numerici a determinati oggetti o fenomeni sulla base di criteri prestabiliti, che ha la finalità di misurare le proprietà di differenti oggetti o fenomeni in modo tale che possano essere compresi, comparati e analizzati tra loro. Lo *scaling* possiede il grande vantaggio di trasformare i dati, rendendoli analizzabili e interpretabili secondo criteri più precisi che ne facilitano l'esame. Invece, per quanto riguarda la *cluster analysis* si tratta di una tecnica impiegata per creare gruppi omogenei, detti per l'appunto cluster, tra un insieme di dati, che contengano al loro interno unità statistiche che presentino caratteristiche simili tra loro. In molti casi, il processo di *scaling* viene eseguito preliminarmente rispetto alla *cluster analysis*, tanto da essere considerata come una fase propedeutica a quella del clustering dal momento che getta le basi per un'efficace suddivisione in gruppi standardizzando le variabili e riducendo le conseguenze delle differenze di scala. Sebbene questi strumenti statistici possano essere utilizzati per obiettivi diversi, condividono la loro dipendenza dai dati di prossimità, detti anche dati di similarità, per identificare la struttura sottostante di un insieme di oggetti. I dati di prossimità riguardano la misurazione di quanto gli oggetti siano simili o dissimili tra loro e sono tipicamente discussi in termini di somiglianza perché, in parte, ai partecipanti viene chiesto di fornire valutazioni di somiglianza, le quali vengono utilizzate perché valutare la somiglianza degli oggetti è psicologicamente più facile che valutarne la dissomiglianza. Per cui, per avendo presupposti diversi, possono essere utilizzate in modo complementare per aiutare a interpretare un insieme di dati. Date queste premesse, è possibile considerare questi due passaggi come parti integranti del procedimento di analisi dei dati, andando ad esaminare i loro obiettivi comuni e i vantaggi che possono generarsi da un loro utilizzo congiunto. Tra gli obiettivi comuni rientrano sicuramente quello di semplificare e descrivere i dati, poiché tutti i metodi di *scaling* e *clustering* aiutano i ricercatori a ridurre la complessità di un insieme di dati

identificando la struttura sottostante all'interno di un insieme di dati, e quello di verificare un'ipotesi riguardo la struttura di un certo fenomeno osservato, ad esempio possono aiutare nella ricerca di caratteristiche simili tra i differenti gruppi. Un terzo obiettivo delle tecniche di *scaling* e di *clustering* è quello di identificare la rappresentazione psicologica sottostante a un insieme di stimoli, in quanto rappresentano uno strumento molto efficace per comprendere al meglio i processi mentali innescati dai rispondenti quando percepiscono ed interpretano le domande che gli vengono poste.

Passando all'esame dei singoli processi, la *cluster analysis* presuppone che gli oggetti siano differenziati in modo categorico, per cui più caratteristiche due oggetti hanno in comune, più saranno percepiti simili e più alta sarà la probabilità che appartengano alla stessa categoria, ma sono le caratteristiche uniche di un oggetto a definirlo come distinto da altri membri di altre categorie all'interno del dominio. L'analisi dei cluster incorpora sia le caratteristiche condivise che quelle uniche degli oggetti per collocare gli oggetti simili nella categoria appropriata. Esistono diversi tipi di analisi dei cluster che si distinguono tra loro sulla base delle caratteristiche che essi possiedono. Una caratteristica che differenzia le analisi dei cluster è l'unicità o meno dei cluster, infatti, alcune analisi creano cluster unici, mentre altre permettono la sovrapposizione di cluster. Altre caratteristiche dirimenti nella distinzione delle differenti tipologie di *cluster analysis* sono la gerarchia, in quanto le analisi gerarchiche consentono di creare cluster annidati, mentre le analisi non gerarchiche non lo permettono, e il fatto che un'analisi possa essere agglomerativa o divisiva. Un'analisi agglomerativa inizia con tutti gli oggetti che formano un cluster separato e poi crea una struttura combinando cluster più piccoli in cluster più grandi, mentre un'analisi divisiva parte con tutti gli oggetti in un grande cluster e poi crea una struttura dividendo i cluster più grandi in cluster più piccoli. Se ci si volesse soffermare su una tipologia particolare di cluster analysis, sarebbe opportuno farlo focalizzando l'attenzione sui cosiddetti alberi ultrametrici, che rientrano nell'insieme dei metodi gerarchici. Gli alberi ultrametrici sono spesso rappresentati da dendrogrammi, o modelli ad albero, dei dati, a un livello molto elementare, l'analisi dei dati procede mettendo insieme i due unità statistiche più simili in un cluster, poi le successive due unità verranno messe insieme in un cluster. Se i due oggetti simili successivi non fanno parte del cluster originale, viene creato un

nuovo cluster, ma se uno degli oggetti fa parte del cluster originale, allora quest'ultimo verrà aggiunto al cluster originale in modo che il primo cluster sia annidato all'interno del nuovo cluster. Il risultato è che ogni unità viene assegnata ad un cluster non sovrapposto e i cluster sono organizzati gerarchicamente. Questo processo di costruzione e annidamento dei cluster nel tempo, se rappresentato graficamente, produce un albero ultrametrico. L'assunto di base degli alberi ultrametrici è quello della disuguaglianza ultrametrica, secondo il quale per ogni tre oggetti in un albero, due degli oggetti devono formare un livello inferiore, o sottogruppo, in relazione al terzo oggetto (Cortier, 1996). Un aspetto importante legato all'analisi dei cluster è il modo in cui i dati sono collegati tra loro in cluster. A tal proposito, si sottolinea come esistano diverse tecniche per calcolare la distanza tra i gruppi, ognuno dei quali può dar luogo a cluster molto diversi. I metodi più comuni e consolidati sono il metodo del legame semplice, il metodo del legame medio, il metodo del legame completo e il metodo Ward. Volendoli descrivere brevemente, il metodo del legame singolo crea il cluster di livello più basso combinando i due elementi più simili, o per meglio dire meno dissimili. In altri termini, il legame singolo calcola la distanza tra due gruppi come la più piccola distanza tra gli elementi di un cluster e dell'altro considerato. Il legame medio, invece, per calcolare la distanza tra due gruppi impiega la media aritmetica delle distanze tra le unità statistiche contenute nel primo e quelle nel secondo gruppo considerato, mentre il metodo del legame impiega la massima distanza tra i soggetti di un gruppo e quelli dell'altro per rappresentare la distanza che sussiste tra i cluster oggetto d'osservazione. Il metodo di Ward adotta un approccio differente per quanto riguarda il calcolo della distanza tra i gruppi. Infatti, l'obiettivo del metodo è minimizzare la variabilità dei punteggi di dissimilarità tra i cluster, per cui i primi elementi che saranno uniti saranno quelli che determinano la minore variabilità, in particolare la minore somma delle differenze al quadrato tra le unità statistiche. Per quanto riguarda gli altri elementi che saranno uniti nel cluster, questi saranno quelli che determinano il minore aumento della variabilità dei punteggi di dissomiglianza tra i cluster. Il metodo Ward tende a produrre cluster con un numero approssimativamente uguale di oggetti al loro interno. In sintesi, il metodo di Ward sceglie di unire i due gruppi che presentano un incremento minimo della varianza infragruppo. In conclusione, una volta portata a termine la *cluster analysis*, i ricercatori dovranno procedere all'individuazione del numero ottimale di cluster. Non esistono

regole precise per determinare il numero di cluster ottimale, ma esistono numerose tecniche come quella basata sull'entità della varianza, che prevede l'utilizzo degli *scree plot*⁵¹ per identificare il numero ottimale dei gruppi. Nello specifico, sarà necessario individuare il punto del flesso della curva, detto anche gomito, perché indica il punto oltre il quale non si ottengono informazioni unendo ulteriormente i cluster, andando a sottolineare la presenza di dissimmetrie sostanziale tra i due gruppi che sarebbero uniti successivamente e indicando il numero di cluster ottimale per l'indagine che si sta conducendo.

Per quanto concerne, invece, lo *scaling*, si deve sottolineare come in letteratura esistano numerose classificazioni, che identificano diverse tipologie di *scaling*. Tra queste meritano una particolare attenzione lo *scaling* multidimensionale e lo *scaling* per differenze individuali, che sono le due principali tecniche per quanto riguarda la trasformazione dei dati multivariati. L'obiettivo dello *scaling* multidimensionale (MDS) è quello di identificare la rappresentazione dimensionale degli oggetti come rappresentata in una matrice spaziale di tali oggetti. L'MDS si basa sul presupposto della "minimalità", secondo cui ogni punto deve essere almeno tanto vicino a se stesso quanto a qualsiasi altro punto, che se viene inteso in termini di somiglianza, un oggetto deve essere il più simile a se stesso. Esiste anche un secondo presupposto sul quale si fonda lo *scaling* multidimensionale, ossia quello della simmetria, secondo il quale un qualsiasi punto x è vicino ad un altro punto y, quanto il punto y lo è al punto x. Oltre alla minimalità e alla simmetria, vi è ancora un ultimo presupposto, che è quello della disuguaglianza dei triangoli, secondo il quale due punti, x e y, che sono entrambi vicini a un terzo punto, z, devono essere almeno moderatamente vicini tra loro. Sebbene questi tre presupposti siano generalmente dati per scontati, sono importanti da considerare. Quando i presupposti sono soddisfatti, gli oggetti sono probabilmente meglio rappresentati dimensionalmente e l'MDS sarebbe la tecnica migliore da utilizzare per l'analisi dei dati. Quando i presupposti non sono soddisfatti, invece, è probabile che gli oggetti siano probabilmente meglio rappresentati in modo categorico, e l'analisi dei cluster sarebbe lo strumento più appropriato da utilizzare. Lo *scaling* multidimensionale trasforma le relazioni di vicinanza tra dati in distanze misurabili in uno spazio

⁵¹ Particolare rappresentazione grafica impiegati specialmente nell'analisi dei dati multivariati, che consente di stabilire quanti fattori dovranno essere coinvolti nell'ambito dell'analisi.

dimensionale. Una volta convertiti i dati, esistono diverse tecniche che possono essere utilizzate per misurare le distanze nello spazio, tra cui la distanza euclidea, la distanza *city-block* e la distanza massima. In particolare, la scelta tra distanza euclidea e quella *city-block* riguarda l'interazione tra le dimensioni. Ad esempio, con la *city-block*, le dimensioni sono indipendenti e, di conseguenza, vengono trattate come dimensioni separate quando vengono espressi giudizi di somiglianza e sono additive quando si determinano le vicinanze, mentre con la distanza euclidea, le dimensioni dipendono l'una dall'altra. Infine, nel caso insolito in cui opera la distanza massima, gli oggetti sono rappresentati in uno spazio multidimensionale, ma le persone si concentrano solo sulla dimensione lungo la quale gli oggetti sono più dissimili quando esprimono giudizi di somiglianza (Carroll & Wish 1974). Riassumendo, la distanza euclidea è la più appropriata quando le dimensioni di uno spazio sono integrali, quella *city-block* è la migliore quando le dimensioni lungo le quali gli oggetti differiscono sono separabili e la distanza massima è appropriata quando solo una dimensione tra le varie viene utilizzata per formulare giudizi.

A differenza dello *scaling* multidimensionale, la tecnica dello *scaling* per differenze individuali (INDSCAL) non si fonda sul presupposto che tutti i partecipanti giudichino le somiglianze tra i diversi oggetti in modo fondamentalmente identico. Infatti, questa tecnica ipotizza che lo spazio sottostante sia lo stesso per entrambi i campioni, per cui la dimensionalità e la posizione degli oggetti sono simili, ma i due campioni pesano in modo diverso le dimensioni nel formulare i loro giudizi. Lo *scaling* per differenze individuali parte da quattro presupposti fondamentali. In primo luogo, presuppone l'esistenza di un insieme di dimensioni comuni che sono alla base dei giudizi dei diversi gruppi sugli oggetti e, in secondo luogo, presuppone che queste dimensioni siano condivise da tutti gli individui dei diversi gruppi. Il terzo presupposto è rappresentato dal fatto che si assume che i giudizi di somiglianza di tutti i partecipanti dei diversi gruppi siano linearmente correlati alle distanze nello spazio e che tutti i giudizi all'interno di un gruppo di partecipanti siano ponderati allo stesso modo. In quarto luogo, lo *scaling* per differenze individuali assume che le distanze nello spazio siano rappresentate da una funzione euclidea modificata, in cui viene introdotto il peso che i partecipanti applicano ad una dimensione, corrispondente all'importanza di tale dimensione per i giudizi di somiglianza dei partecipanti. Se una dimensione ha un peso

elevato per un gruppo di partecipanti, significa che i partecipanti di quel gruppo pongono una forte enfasi su quella dimensione e che gli oggetti lungo questa dimensione sono percepiti come maggiormente dissimili, rispetto a un altro gruppo il cui peso è minore su quella dimensione. Poiché le distanze sono ponderate, l'INDSCAL viene talvolta definito un modello euclideo ponderato. L'*output* dello *scaling* per differenze individuali contiene due spazi diversi, ossia lo spazio degli stimoli di gruppo, che è la configurazione degli oggetti all'interno di uno spazio, e lo spazio dei pesi, che indica quanto ogni gruppo pesi una dimensione o quanto la dimensione in questione sia saliente per ciascun gruppo. Questa tecnica di classificazione si differenzia da uno *scaling* multidimensionale per tre aspetti principali, di cui il primo è rappresentato dal fatto che le dimensioni fornite in un output INDSCAL sono significative, in quanto rappresenta i fattori che influenzano le differenze individuali nei dati. Un altro modo in cui lo *scaling* per differenze individuali differisce da uno *scaling* multidimensionale è che il numero appropriato di dimensioni da utilizzare nella prima metodologia è determinato esclusivamente dal miglioramento della varianza percentuale contabilizzata. Una terza differenza è che lo *scaling* per differenze individuali può tenere traccia dei cambiamenti nel tempo. Per fare questo, la configurazione iniziale è specificata a priori. Questa specificazione iniziale consente un confronto con studi precedenti, se la specifica adattata dai vecchi dati si adatta bene ai nuovi dati, lo studio precedente viene replicato. Ancora più importante, specificare la configurazione a priori in uno studio longitudinale può dire al ricercatore se il passare del tempo ha influenzato i pesi dimensionali.

4.3.1 La scala Likert.

Una tipologia particolare di scala psicometrica è rappresentata dalla scala Likert. Questo strumento, che deve il nome al suo inventore il sociologo americano Rensis Likert, viene utilizzato per misurare le opinioni, gli atteggiamenti e il modo in cui le persone percepiscono una certa tematica. Questa scala prevede che un gruppo di persone giudichi un insieme di categorie indicando quanto sono d'accordo o in disaccordo, approvano o disapprovano o credono che sia vero o falso. È un metodo che lascia ampia discrezionalità ai ricercatori, dal momento che non esiste una struttura predefinita della scala Likert, anche se l'aspetto più

importante è includere almeno cinque categorie di risposta. Spesso, però, gli estremi della scala vengono aumentati per creare una scala a sette punti, aggiungendo più informazione ai rispettivi vertici rispetto a quanto non avviene con le scale a cinque punti. È stato dimostrato che la scala a sette punti raggiunge i limiti superiori dell'affidabilità della scala, per questo motivo è opportuno utilizzare la scala più ampia possibile. Tuttavia, è necessario porre attenzione a non realizzare una scala Likert con un numero di punti troppo elevato, e quindi con un livello di dettaglio eccessivo, in quanto potrebbe condurre l'intervistato a non rispondere in maniera coerente e potrebbe causare una dispersione nell'informazione. Per l'analisi è sempre possibile comprimere le risposte in categorie condensate, se ritenuto opportuno. A questo proposito, le scale possono essere talvolta limitate ad un numero pari di categorie, in genere quattro, per eliminare l'opzione neutrale in una scala di sondaggio a scelta forzata. Si tratta di un metodo di codifica per assegnare numeri interi a categorie di risposta ordinate, quali mai, a volte, spesso, sempre. La scala Likert è una delle tecniche più note ed utilizzate, adatto per differenti contesti, al punto che viene impiegato nelle cliniche mediche, negli istituti scolastici, nelle società di consulenza, nelle università e in molti altri settori. È un metodo semplice e pratico per la raccolta dei dati, che vengono raccolti chiedendo alle persone di scegliere categorie di risposta appropriate. Nello specifico, gli intervistati sono chiamati a valutare una certa tematica, ad esempio potrebbe essere chiesto ai clienti di una compagnia crocieristica di valutare la qualità dei servizi di ristorazione presenti a bordo della nave, e si chiede loro di rispondere assegnando un valore che convenzionalmente può andare da zero⁵² a cinque, a cui corrisponde un giudizio. Tenendo sempre in considerazione l'esempio dei crocieristi che valutano i servizi di ristorazione su una scala da zero a cinque, il giudizio che fa riferimento a zero potrebbe corrispondere a "Per niente soddisfacente" e quello che viene assegnato a cinque potrebbe essere "Estremamente soddisfacente", mentre per i valori compresi nell'intervallo da zero a cinque si potrebbe dire che questi forniscono una

⁵² Esistono casi in cui la scala può partire da numeri inferiori allo zero. In genere quando si utilizza una scala con valori minori di zero, i ricercatori tendono a farla partire da valori come -2. Così facendo, si vuole agevolare il rispondente nella corretta comprensione delle modalità di valutazione, in quanto il valore 0 viene posto come valore mediano e ciò può aiutare gli intervistati a comprendere la scala di valutazione.

rappresentazione graduale dell'ordine di soddisfazione. Chiaramente è essenziale che coloro che redigono il questionario si occupino di specificare a quali giudizi corrispondano i valori, in modo da orientare i rispondenti nell'effettuare le proprie valutazioni in maniera corretta, evitando di generare in loro confusione circa le modalità di compilazione. Una volta che la scala sarà stata compilata, i ricercatori possono procedere all'analisi dei dati. È importante sottolineare che, come regola generale, la media e la deviazione standard sono parametri non validi per le statistiche descrittive quando i dati sono su scale ordinali, così come qualsiasi analisi parametrica basata sulla distribuzione normale. Le procedure non parametriche basate sul rango, la mediana o il range sono appropriate per l'analisi di questi dati, così come i metodi privi di distribuzione, quali tabulazioni, frequenze, contingenze e così via. Rimanendo sull'analisi delle scale Likert, un altro esempio si ha quando gli insiemi di item Likert possono essere combinati per formare indici. Tuttavia, si tratta di un approccio al quale occorre prestare molta attenzione, in quanto la maggior parte dei ricercatori insiste sul fatto che tali combinazioni di scale superino i test di interrelazione e validità⁵³. Inoltre, la combinazione di scale per formare un indice a livello di intervallo presuppone che questa combinazione formi una caratteristica o una variabile sottostante. Per riassumere quanto detto finora sulla scala Likert, l'analisi iniziale dei dati scalari Likert non deve coinvolgere le statistiche parametriche, ma deve basarsi sulla natura ordinale dei dati. Sebbene le variabili della scala Likert rappresentino solitamente una misura continua sottostante, l'analisi dei singoli item dovrebbe utilizzare procedure parametriche solo come analisi pilota. La combinazione delle scale Likert in indici aggiunge valori e variabilità ai dati e se le ipotesi di normalità sono soddisfatte, si può procedere all'analisi con procedure parametriche. Infine, la conversione di uno strumento a cinque o sette categorie in una variabile continua è possibile con una linea o una barra calibrata.

4.3.2 Data cleaning.

Prima di iniziare ad analizzare i dati, è necessario “pulirli”. La pulizia dei dati è il processo di identificazione e spesso di correzione dei dati "sporchi", ciò significa che

⁵³ Test dell'alfa di Cronbach o il test Kappa.

è necessario rimuovere tutte le variabili non necessarie, gli errori di inserimento, le risposte illogiche, ad esempio quelle al di fuori di un intervallo possibile, le informazioni identificative dell'intervistato, gli errori di salto e così via. Inoltre, è necessario creare nomi di variabili, etichette di variabili ed etichette di valori adeguati. Sebbene molti tipi di analisi riportino errori nei dati, è necessario eseguire una pulizia sistematica dei dati per eliminare tutte le voci errate, in questo modo si evita di trarre conclusioni sbagliate basate, ad esempio, sul peso relativo di valori estremi elevati di *outlier*. Il processo di data *cleaning* ha inizio con lo *screening* dei dati, ovvero saranno passati al vaglio i dati grezzi, per caso e per variabile, alla ricerca di ridondanze, di un'etichettatura corretta, di informazioni identificative e di incongruenze e, in particolare, uno dei primi *step* consiste nell'unione dei dataset. Dopo aver unito tutti i set di dati desiderati e aver raccolto tutte le informazioni in un unico set di dati, si può iniziare a indagare sui casi in esso contenuti. È necessario verificare se qualcuno ha iniziato il questionario più di una volta ed è quindi un duplicato nel dataset. Se ci sono duplicati, è necessario confrontare i casi e tenere il migliore, ad esempio si potrebbe optare di considerare il migliore il tentativo che ha cercato di andare più avanti possibile nel questionario. Inoltre, è necessario rimuovere i casi di prova e tutti i casi in cui i rispondenti non hanno completato alcuna domanda, questo può avvenire perché spesso le persone che guardano la schermata introduttiva ma poi abbandonano il sondaggio vengono memorizzate nel database. Quindi, queste persone dovrebbero essere rimosse e trattate come non rispondenti, mentre risulta più difficile sapere come trattare le persone che hanno completato una o più domande, ma che poi si sono ritirate. Questa scelta può essere fatta per le indagini con un numero relativamente alto di casi e per le indagini in cui i tassi di abbandono sono bassi, anche se va detto che in generale la tendenza per i ricercatori sarebbe quella di mantenere il maggior numero possibile di rispondenti. Si può anche scegliere di mantenere gli intervistati che hanno completato la maggior parte delle domande dell'indagine o che hanno fornito informazioni cruciali per l'indagine. Si procede poi a ripulire le variabili, eliminando quelle ridondanti, ad esempio i testi introduttivi e le intestazioni che possono essere memorizzate all'interno del dataset ma che non forniscono nessun tipo di informazione e quindi possono essere eliminati dal file. Successivamente, si procede ad assegnare i nomi alle variabili, le etichette delle variabili e le etichette dei valori, che devono essere controllati e migliorati, se

necessario. È importante che ogni variabile abbia un'etichetta adeguata, in modo da fornire il testo giusto nelle analisi, ciò consente anche di risparmiare tempo se si desidera produrre tabelle o figure nel report, in quanto il significato delle variabili è già presente nell'output. Si passa alla fase di ricodifica, perché per molte variabili è necessaria una trasformazione. Infatti, in alcune circostanze può capitare di trovarsi di fronte ad una scala con etichette di risposta da 0 a 10, mentre i software di indagine assegnano sempre alla prima opzione di risposta il valore 1, con il risultato che le etichette dei valori vanno da 1 a 11, per cui è necessario ricodificare 1-11 in 0-10. Dopodiché, si passa all'osservazione dei paradatai, che specialmente nei sondaggi online sono presenti in grandissime quantità. Questi possono fornire informazioni molto preziose, come la durata del questionario o indicazioni sulla serietà con cui il rispondente ha risposto all'indagine. Ad esempio, una durata di un minuto per un sondaggio di trenta domande indica che l'intervistato è stato veloce e non ha fornito risposte significative e si potrebbe decidere di eliminare il caso. Quindi, paradatai come l'ora del giorno in cui l'indagine è stata avviata, il browser o il dispositivo utilizzato dal rispondente e così via, possono fornire ulteriori informazioni sul processo di risposta. È quindi possibile spiegare i casi di abbandono e anticipare i potenziali errori nelle indagini future. Un'altra fase di estrema rilevanza è quella costituita dall'eliminazione degli *outliers*, ossia di quei valori che si differenziano in maniera significativa dalla distribuzione generale, i quali possono essere generati da errori di misurazione, da errori nella compilazione o da altri fenomeni anomali. La loro identificazione è cruciale perché esiste il rischio che questi possano impattare negativamente sui risultati prodotti dall'indagine. Gli *outliers* possono avere un effetto significativo sul risultato delle analisi. È possibile ridurre al minimo la loro influenza modificando il punteggio grezzo con un valore meno estremo, come la media, la modalità o la mediana, ed eliminando il caso estremo dalle analisi. Esistono, chiaramente, molti metodi per identificare gli *outliers*, come i diagrammi a scatola e le tecniche basate sulla distribuzione dei dati. Tuttavia, esistono molte altre tecniche che si possono basare sugli aspetti più disparati. In tal senso, viene proposto un esempio che consente di comprendere le metodologie sottostanti a questa fase. In particolare, si supponga di avere un questionario che ha l'obiettivo di osservare come i clienti di un'azienda operante nel settore dei trasporti valutino i servizi erogati dall'impresa stessa. La struttura del sondaggio in questione si

articola in due parti: una prima parte nella quale si chiede ai rispondenti di scegliere un giudizio tra «irrelevante», «sorprendente», «decisivo» e «necessario» per singole categorie di servizi e una seconda parte dove è presente una scala Likert con scala da 0 a 5, nella quale sono contenuti gli stessi *item* proposti nella prima per verificare che gli intervistati abbiano risposto in maniera coerente. Una volta somministrato il questionario e raccolti i dati, per eliminare dal campione i rispondenti che hanno fornito risposte non coerenti tra le domande della prima e della seconda parte, vengono confrontati i punteggi attribuiti alle variabili categoriche e alla scala Likert. Successivamente viene assegnato un punteggio da 1 a 4, in ordine crescente di rilevanza, alle risposte: irrilevante, sorprendente, decisivo e necessario e, in un secondo momento, vengono calcolate le medie delle domande che avevano ad oggetto lo stesso attributo e sono osservati i minimi e i massimi delle differenti domande al fine di attuare una normalizzazione dei dati. L'esempio in questione fa ricorso alla normalizzazione in quanto quest'ultima consente di trasformare i valori delle variabili in un intervallo compreso tra 0 e 100. La terza fase comprende in primis il calcolo degli scarti (Δ), per valutare quanto una variabile x_i si discosti da un'altra variabile y_i . A questo punto lo studio si concentra sull'osservazione della distribuzione degli scarti mediante i box-plot, per individuare il valore soglia, che verrà impostata al 50%, oltre il quale eliminare le unità statistiche che hanno fornito risposte non coerenti. Si tenga a mente che la rappresentatività può essere messa in discussione all'aumentare dei rispondenti eliminati, per cui sarà necessario agire sempre con cautela quando si trattano i casi estremi. Bisognerà, quindi, analizzare i dati con e senza *outlier* e vedere se i risultati differiscono. In caso affermativo, si dovranno riportare le analisi senza *outlier* e indicare attentamente perché e come sono stati rimossi i casi. Tornando all'esempio, vengono create differenti variabili *dummy* per ognuno degli scarti, le quali potevano assumere il valore 1 se il valore dello scarto era maggiore di 50 e valore 0 altrimenti, e un'altra variabile *dummy* che aveva come valore la somma delle variabili *dummy* precedentemente definite. In questo modo è stata ottenuta una variabile che forniva il numero di risposte che il soggetto ha valutato in maniera incoerente. Sono stati, quindi, eliminati i soggetti che hanno risposto a più del 50% delle domande presenti all'interno del questionario con uno scarto superiore al 50%.

Una volta svolta tutte queste operazioni, il set di dati si può dire pulito ed è pronto per essere utilizzato. Ma prima di distribuire il set di dati, è necessario codificarlo in modo che nessuno possa identificare i rispondenti.

Conclusioni.

L'analisi dei dati rappresenta un processo molto delicato per i ricercatori che conducono l'indagine e per coloro che devono assumere decisioni in ambito aziendale. Infatti, si tratta di una fase che presenta numerose insidie, che possono derivare da differenti fattori legati alla costruzione delle domande contenute nel questionario, all'ordine delle domande, alla negligenza degli intervistati e così via. Esistono una serie di strategie che possono essere adottate in diverse fasi del processo di analisi dei dati, che vanno dalla fase precedente la redazione del questionario sino ad arrivare alla suddivisione in *cluster* degli intervistati. Infatti, prima di poter procedere alla stesura delle domande, i ricercatori dovranno comprendere lo scopo del questionario, chiarire i termini utilizzati e conoscere i rispondenti. Dopo aver adottato queste accortezze, i ricercatori si dovranno occupare della scrittura vera e propria dei quesiti, che dovranno essere precisi e concreti, perché un maggior grado di dettaglio mette il rispondente nella condizione di non fraintendere la domanda e l'intervistatore in quella di ottenere risposte più affidabili. Non solo, i quesiti dovranno essere scritti con un linguaggio semplice e convenzionale, per cui non dovranno essere presenti *slang* o espressioni gergali, così da poter aumentare la comprensione delle domande, e dovranno essere redatte in modo da evitare di far sentire a disagio il rispondente. Nello specifico, quest'ultima caratteristica permette al ricercatore di considerare anche gli aspetti cognitivi e i processi mentali, come la desiderabilità sociale, che guidano l'intervistato nel processo di risposta. Infatti, per i team di ricercatori conoscere gli aspetti psicologici che influenzano i rispondenti è molto importante, perché una conoscenza di queste dinamiche consente di realizzare domande in grado di mitigare le implicazioni negative di questi processi mentali sulle risposte, senza considerare il fatto che formulare domande che riducano l'impatto di questi aspetti permette di ottenere risposte più affidabili e prive di errori. Per limitare l'influenza di questi bias cognitivi, si può fare ricorso ad una serie di tecniche, come l'*item count*, che consiste nel nascondere l'item che provoca questi bias in un elenco di comportamenti innocui, e la casualizzazione dei quesiti, che prevede un cambiamento nell'ordine delle domande in modo da ridurre gli effetti di *primacy* e di *recency*. In questa prima fase più legata alla redazione di un questionario, anche i metodi di somministrazione del sondaggio giocano un ruolo fondamentale nell'ambito delle strategie per eliminare gli errori, dal momento che

alcune tecniche, come l'intervista personale, permettono all'intervistatore di avere un maggiore controllo sulla bontà delle risposte fornite dall'intervistato. Nelle fasi successive più connesse all'analisi dei dati vera e propria, è fondamentale osservare attentamente i dati a disposizione e verificare quanto questi possano essere pronti per essere analizzati. Si possono adottare, anche in questo caso, metodologie differenti per andare a pulire i dati dagli errori, come il ricorso alla scala Likert come strumento di conferma per eventuali domande poste in precedenza oppure il ricorso alla normalizzazione dei dati quando si devono confrontare dati misurati con scale di valori diverse. Queste fasi di *data cleaning* sono di estrema rilevanza, in quanto permettono ai ricercatori di poter procedere all'applicazione dei metodi di analisi dati, come l'ANOVA, il test di Wilcoxon e così via. In conclusione, nell'economia dell'analisi dati raccolti da un questionario diventa fondamentale predisporre un questionario corretto, ovvero un sondaggio contenente domande formulate in modo da poter essere chiare e comprensibili a tutti gli intervistati, che contenga al suo interno le accortezze necessarie per aumentare la validità dello strumento e l'affidabilità delle risposte, riducendo così la possibilità di analizzare dati errati e di prendere, conseguentemente, decisioni sulla base di queste informazioni viziate da errori.

Bibliografia

Designing and Doing Survey Research, Lesley Andres, 2012.

Survey Methodology, Robert M. Groves, Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, Roger Tourangeau, 2004.

Materiale del corso “Analisi di mercato per il trasporto marittimo di persone”, “Redazione e somministrazione di un questionario”.

Asking Questions: A Practical Guide to Questionnaire Design, Sudman e Bradburn, 1982.

How to ask survey questions, Arlene Fink, 2002.

Sage research methods, constructing effective questionnaires, Robert A. Peterson, 2013.

Design and Evaluation of Survey Questions, Floyd J. Jr. Fowler, C. Cosenza, 2009

Desiderabilità Sociale e Acquiescenza. Alcune Trappole delle Inchieste e dei Sondaggi., Roccato M., (2003).

A companion to survey research, Micheal Ornstein, 2013

Designing surveys: A guide to decisions and Procedures, Blair Johnny J, Czaja Ronald F., Blair Edward, 2013.

Statistica per analisi di mercato, Francesca Bassi e Salvatore Ingrassia, 2022.

Fundamentals of statistical inference: what is the meaning of random error? Norbert Hirschauer, Sven Gruner e Oliver Mubbhoff, 2022.

Sandip Sinharay, “An overview of statistics in education”, 2010, Princeton, NJ, USA.

Sampling Theory and Practice, Changbao Wu e Mary E. Thompson

Anita S Acharya, Anupam Prakash, Pikee Saxena e Aruna Nigam, 2013, Indian Journal of medical specialities, “Sampling: why and how of it?”

Allison Shorten e Calvin Moorley, “Selecting the sample”, 2014.

Andrea E. Berndt, “Sampling methods”, Journal of human lactation, Los Angeles, 2020.

El-Masri, Maher M., “Probability sampling”, Canada, 2017.

How to sample in surveys, Arlene Fink, 2003.

El-Masri, Maher M, “Non-probability sampling”, Canada, 2017.

The SAGE Handbook of Survey Methodology, Vesja Vehovar, Vera Toepoel e Stephanie Steinmetz, Londra, 2016.

The SAGE Encyclopedia of Educational Research, Measurement and Evaluation, Jill Coleman e Karen Multon, 2018.

The SAGE Encyclopedia of Social Science Research, Karen A. Bosch, Alan Bryman and Tim F Liao, Londra, 2014.

Paul P. Biemer, “Total survey error: design, implementation and evaluation”, 2010.

Robert M. Groves, Lars Lyberg, “Total survey error: past, present and future”, 2010.

The SAGE Handbook of Survey Methodology, Christian Wolf, Dominique Joye, Tom W. Smith and Yang-chih Fu, 2017.

Tom W. Smith, “Refining the total survey error perspective”, Chicago, 2011.

How Data Quality Affects our Understanding of the Earnings Distribution, Reza Che Daniels, Singapore, 2022.

Survey Methodology, Robert M. Groves, Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, Roger Tourangeau, 2004.

Conducting online surveys, Valerie M. Sue, Lois A. Ritter, Los Angeles, 2007.

Philip Sedgwick, “What is sampling error?”, Londra, 2012.

Douglas G. Altman, Martin J Bland, “Uncertainty and sampling error”, Oxford, 2014.

Encyclopedia of Survey Research Methods, Paul J. Lavrakas, Londra, 2008.

Non sampling error in social surveys, Davide E. McNabb, Los Angeles, 2014.

J.N.K Rao “On Measuring the Quality of Survey Estimates”, Ottawa, 2005.

Virginia M. Lesser, William D. Kalsbeek, “Nonsampling Errors in Environmental Surveys”, 1999.

Statistica per analisi di mercato, Francesca Bassi e Salvatore Ingrassia, 2022.

The SAGE Encyclopedia of Social Research Methods, Michael S. Lewis-Beck, Alan Bryman and Tim Futing Liao, 2004.

How to Assess and Interpret Survey Psychometrics, Mark S. Litwin, 2003.

Analysing Quantitative Survey Data for Business and Management Students, Jeremy Dawson, 2017.

Andrew T. Jebb, Vincent Ng and Louis Ta, “A Review of Key Likert Scale Development Advances: 1995–2019”, 2021.

Samuel Messick, “Meaning and Values in Test Validation: The Science and Ethics of Assessment”, 1989.

Paul E. Newton, Stuart D. Shaw, “Standards for talking and thinking about validity.”, 2013.

Borsboom, D., Mellenbergh, G. J., and van Heerden, “The concept of validity.”, 2004.

Bailey, K. D., “Typologies and taxonomies: An introduction to classification techniques.”, 1994.

Borg, I., Groenen, P., Modern multidimensional scaling: Theory and applications, 1997.

Corter, J. E., “Tree models of similarity and association.”, 1996.

Aldenderfer, M. S., Blashfield, R. K.,” Cluster analysis. Beverly Hills”, 1984.

“The SAGE Sourcebook of Advanced Data Analysis Methods for Communication Research”, Andrew F. Hayes, Michael D. Slater e Leslie B. Snyder, 2008.

“Multidimensional perceptual models and measurement methods”, Carroll, J. D., Wish, M. (1974).

“Multidimensional scaling”, Kruskal, J. P., Wish, M., 1978.

“Measurement, Mathematics and New Quantification Theory”, Nishisato Shizuhiko, 2023, Singapore.

Elaine Allen and Christopher A. Seaman, “Likert Scales and Data Analyses”, 2007.

“Doing Surveys Online”, Vera Toepoel, 2017.

“The SAGE encyclopedia of Qualitative Research Methods”, Lisa M. Given, 2008.

Zulfiqar Ali, S Bala Bhaskar, “Basic statistical tools in research and data analysis”, 2016, Indian Journal of Anaesthesia.