

***Indagine multiscopo sulle famiglie
Viaggi e vacanze
Anno 2013***

Manuale per l'utente

I file standard vengono rilasciati per finalità di studio e ricerca. Per ottenere tali file è necessario registrarsi al Contact centre. Una volta effettuata la registrazione, la richiesta deve essere formulata selezionando nel Contact centre l'area "Collezioni campionarie di dati elementari (file standard) e compilando un modulo on-line.

Per informazioni sull'indagine rivolgersi a:
Istat - Servizio condizioni economiche delle famiglie
U.O. Viaggi e vacanze
viale Oceano Pacifico, 171 – 00144
Roma
tel: 06.4673.4655
fax: 06.4673.3908
e-mail: maditorr@istat.it

INDICE

1. PREMESSA.....	3
2. L'INDAGINE TRIMESTRALE “VIAGGI E VACANZE”	3
3. I DATI CAMPIONARI E IL LORO UTILIZZO	4
APPENDICE A - IL QUESTIONARIO	7
APPENDICE B - ALCUNE DEFINIZIONI.....	8
APPENDICE C - STRATEGIA DI CAMPIONAMENTO E VALUTAZIONE DEGLI ERRORI CAMPIONARI	9

1. PREMESSA

Il decreto legislativo n. 322 del 6/9/1989 regola la diffusione delle informazioni statistiche prodotte nell'ambito del Sistema Statistico Nazionale al fine di garantire la riservatezza dei rispondenti. In particolare, per la diffusione di dati elementari, l'articolo 10, comma 2, dispone quanto segue: "Sono distribuite altresì, ove disponibili, su richiesta motivata e previa autorizzazione del Presidente dell'Istat, collezioni campionarie di dati elementari, resi anonimi e privi di ogni riferimento che ne permetta il collegamento con singole persone fisiche e giuridiche".

Nell'osservanza di tale disposizione e del d. lgs del 30/06/2003 n. 196 (*Codice in materia di protezione dei dati personali*) l'Istat ha adottato misure e tecniche che rendono impossibile, o altamente improbabile, il collegamento dei dati rilasciati con l'unità statistica a cui si riferiscono. Per tale motivo vengono apportate alcune modifiche sui file originali delle indagini, nell'intento di garantire la massima protezione ai dati, contenendo al minimo la perdita di informazioni. Le metodologie applicate si concretizzano nell'accorpamento e/o riclassificazione di modalità di variabili e nell'oscuramento di variabili. In quest'ultimo caso nei campi del tracciato record è riportata la dicitura "RISERVATO ISTAT".

Va considerato inoltre che la stessa dicitura è stata utilizzata anche per quelle informazioni che, pur essendo state oggetto di indagine, non sono risultate essere attendibili dal punto di vista campionario e quindi statisticamente non analizzabili.

2. L'INDAGINE TRIMESTRALE "VIAGGI E VACANZE"

A partire dal 1997, l'Istat conduce l'indagine trimestrale telefonica CATI¹ "Viaggi e vacanze" su un campione nazionale annuo di circa 14.000 famiglie (circa 3.500 per trimestre per un totale annuo di circa 40.000 individui). L'indagine, in conformità al Regolamento europeo 692/2011 (che ha sostituito dal 1 gennaio 2012 la Direttiva 95/57/CE), ha l'obiettivo di quantificare e analizzare i flussi turistici dei residenti in Italia, sia all'interno del Paese che all'estero, oltre che di fornire informazioni circa le modalità di effettuazione dei viaggi e le caratteristiche socio-demografiche dei turisti².

Le informazioni vengono rilevate per tutti i componenti della famiglia. L'intervista viene eseguita intervistando, nella famiglia, una persona di almeno 18 anni che fornisce notizie sia per sé che per ciascun altro componente.

I periodi di osservazione sono i trimestri gennaio-marzo, aprile-giugno, luglio-settembre e ottobre-dicembre: per ciascun periodo vengono rilevate, per tutti i componenti delle famiglie del campione, informazioni sui viaggi conclusi nel trimestre, indipendentemente dalla loro data di inizio.

Rispetto alle precedenti indagini dell'Istat sulle vacanze degli italiani, esperienza iniziata nel 1959, proseguita con cadenza pluriennale sino al 1993³ e successivamente con cadenza annuale nell'indagine "Aspetti della vita quotidiana", l'indagine trimestrale "Viaggi e vacanze" rileva, per la prima volta e con carattere di continuità, anche informazioni sulle vacanze di breve durata (1-3 notti) e sul turismo per motivi di lavoro.

L'indagine costituisce, infatti, la prima esperienza nella raccolta di informazioni trimestrali sui viaggi con almeno un pernottamento realizzati dai residenti per ragioni di lavoro o di vacanza e tale peculiarità permette di disporre sistematicamente, in aggiunta ai dati sul movimento dei clienti nelle strutture ricettive⁴, anche di informazioni sul numero e sulle caratteristiche dei viaggi e dei turisti.

Tali informazioni sono ormai ritenute necessarie a comprendere più a fondo la complessità del fenomeno turismo; l'indagine arricchisce così il quadro informativo finora esistente a livello nazionale.

¹ Computer Assisted Telephone Interview.

² La metodologia e l'organizzazione dell'indagine sono disponibili all'indirizzo www.istat.it/dati/catalogo/20060615_00/ e all'indirizzo http://www3.istat.it/dati/catalogo/20030717_01/volmetodviaggiavacanze.pdf.

³ Le indagini sulle vacanze degli italiani condotte dall'Istat fino al 1982, e in particolare con riferimento agli anni 1959, 1965, 1968, 1972, 1975, 1978 e 1982 hanno avuto per oggetto le vacanze di almeno 3 notti consecutive di durata. Per le successive indagini, la definizione di vacanza è stata estesa a periodi di almeno 4 notti consecutive. Dal 1982 al 1993 tali indagini hanno avuto per oggetto, con riferimento però al solo mese di ottobre, anche le brevi vacanze (di durata non superiore a 3 notti consecutive, ma comunque con un pernottamento fuori dal luogo di residenza abituale) ed i viaggi per altri motivi (di almeno una notte di durata), cioè effettuati per motivi diversi da quelli della vacanza (affari, lavoro temporaneo fuori sede, convegni, salute, studio, pellegrinaggi, ecc.).

⁴ I dati si riferiscono alle statistiche correnti Istat dell'attività degli esercizi ricettivi alberghieri e delle altre strutture ricettive (consistenza degli esercizi e movimento dei clienti).

I risultati dell'indagine sono pubblicati annualmente nel volume "I viaggi in Italia e all'estero" della Collana Informazioni Istat. A partire dal 2007, la pubblicazione dei volumi è sostituita dal rilascio di tavole di dati⁵.

La rilevazione utilizza le definizioni standard indicate nella metodologia adottata per l'attuazione del Regolamento europeo 692/2011 sulle Statistiche del Turismo. In particolare, i flussi turistici vengono individuati in quegli spostamenti con pernottamento effettuati fuori dal luogo in cui si vive. In tal senso, per una corretta interpretazione dei dati, assume notevole importanza il significato delle definizioni adottate:

- **viaggio**: spostamento realizzato, per turismo di vacanza o per ragioni di lavoro (purché lo scopo non sia quello di esercitare un'attività alle dipendenze di un soggetto residente nel luogo di destinazione), fuori dal luogo dove si vive e che comporta almeno un pernottamento nel luogo visitato. Sono esclusi i viaggi e gli spostamenti effettuati nelle località frequentate abitualmente tutte le settimane con soste di uno o più pernottamenti, nonché i viaggi di durata superiore ad un anno. In questi casi, infatti, il viaggio non costituisce flusso turistico poiché la località visitata viene associata al luogo dove si vive;
- **turista**: un individuo viene considerato "turista" se ha effettuato uno o più viaggi nel trimestre di riferimento dell'indagine.

I viaggi sono distinti per motivo, secondo la seguente tipologia:

- **viaggio per motivi di lavoro o professionali**: Il viaggio svolto per motivi prevalenti di lavoro quali missioni, partecipazione a congressi, riunioni d'affari o esercizio di attività di rappresentanza, docenza o altre attività professionali. Sono esclusi gli impieghi presso il luogo di destinazione (lavoro stagionale, supplenze, altri lavori temporanei);
- **viaggio di vacanza**: viaggio svolto per motivi prevalenti di relax, piacere, svago o riposo, per visite a parenti o amici, per motivi religiosi/pellegrinaggio, per cure termali o trattamenti di salute.

In relazione alla durata, le vacanze sono distinte in:

- **vacanza "breve"**: quando la durata del soggiorno è inferiore a 4 notti;
- **vacanza "lunga"**: quando la durata del soggiorno è di 4 o più notti.

Le informazioni **destinazione, mezzo di trasporto, tipo di alloggio, motivo della vacanza** **motivo del viaggio di lavoro** sono rilevate sulla base del concetto della "prevalenza". In particolare, la destinazione del viaggio ed il tipo di alloggio vengono associate rispettivamente alla località ed al tipo di alloggio in cui è stato trascorso il maggior numero di notti, mentre il mezzo di trasporto viene indicato come il mezzo con cui è stata coperta la maggiore distanza. Dal 2013, il nuovo Regolamento Europeo 692/2011 per le Statistiche del Turismo richiede di rilevare l'informazione "Tipo di luogo" come quesito con possibilità di risposta multipla e legata alle caratteristiche del luogo visitato. Gli items di risposta previsti (città, mare, campagna, crociera, montagna, altro), erano in precedenza rilevati dall'indagine insieme ad altri sulle possibili attività svolte in vacanza e/o altre specifiche caratteristiche del luogo, in modo da poter definire univocamente, con un criterio di prevalenza (una sola risposta), la tipologia della vacanza effettuata. Pertanto, a partire dal 2013, il quesito ha assunto una diversa valenza e le risposte fornite non possono essere confrontate con le stesse modalità presenti negli anni precedenti.

3. I DATI CAMPIONARI E IL LORO UTILIZZO

Vengono messi a disposizione degli utenti i seguenti file di dati elementari:

- **un file degli individui** (tipo A o tipo B)
lunghezza record: 107 bytes
numero record: 31271
- **un file dei viaggi** (tipo A o tipo B)
lunghezza record: 132 bytes

⁵ Il volume e le tavole sono consultabili sul sito Istat <http://www.istat.it>.

Per garantire la riservatezza delle informazioni a tutela della privacy e per limiti connessi alla numerosità campionaria, non è possibile fornire contemporaneamente i dati sulle suddivisioni territoriali di residenza: regione e tipologia socio-demografica del comune. Per questo motivo, del file individui e del file viaggi sono predisposte due versioni (tipo A e tipo B):

- la prima (**tipo A**) con i codici della regione (REG) e della ripartizione geografica (RIP);
- la seconda (**tipo B**) con i codici della tipologia socio-demografica del comune (DOM) e della ripartizione geografica (RIP).

Le due versioni non possono essere ricongiunte in quanto ogni singola famiglia ha nei due file un identificativo familiare (IDFAM) diverso.

I file sono riferiti all'intero anno e sono ottenuti mettendo in sequenza i quattro file trimestrali; in ogni record è specificato il trimestre di riferimento (TRIM).

Il **file individui** è composto da un record per ciascun individuo componente la famiglia. Ogni record "individuo" contiene una prima parte (coll. 1-44) di informazioni generali sulla famiglia e sull'individuo (codici identificativi e caratteristiche territoriali), una seconda parte (coll. 45-64) contenente il numero di viaggi effettuati dalla famiglia e una terza parte (coll. 65-107) di informazioni sull'individuo (caratteristiche socio-demografiche e numero di viaggi effettuati dall'individuo). Ciò significa che nel file individui sono rappresentati tanto gli individui che hanno viaggiato (turisti), quanto quelli che non hanno viaggiato: quindi è possibile, con riferimento ad ogni trimestre, individuare la quota di popolazione che ha viaggiato e la quota di popolazione che non ha viaggiato.

Il **file viaggi** è composto da un record per ciascun viaggio realizzato da un individuo. Si ricorda che dal 2011, recepite le raccomandazioni internazionali IRTS (International Recommendations for Tourism Statistics) 2008, sono stati esclusi i record corrispondenti a viaggi di lavoro, il cui motivo principale riguarda l'esercizio di un'attività alle dipendenze di un soggetto residente nel luogo di destinazione (come lavori stagionali, supplenze o altri lavori a tempo determinato). Si avverte, pertanto, di tenerne conto nel caso di eventuali confronti con dati di anni precedenti, riguardanti i viaggi totali e/o di lavoro (e relative durate in notti). Ogni record "viaggio" contiene una prima parte di informazioni generali sulla famiglia e sull'individuo (che corrisponde alle coll. 1-31 del file individui) e una seconda parte (coll. 32-132) di informazioni dettagliate sul viaggio (durata, destinazione, mezzo di trasporto, alloggio, ecc.).

Nel caso in cui due o più componenti della famiglia abbiano effettuato un viaggio insieme, sono presenti tanti record "viaggio" quanti sono i partecipanti al viaggio. In ogni record "viaggio" è presente il numero dei partecipanti della famiglia allo stesso viaggio (NPART) ed una variabile (INSIEME) che permette di individuare i record "viaggio" effettuati insieme (si tratta di un numero progressivo entro la famiglia che si ripete per tutti i viaggi fatti insieme da componenti della famiglia stessa). Per distinguere i viaggi di lavoro dai viaggi di vacanza è presente una variabile (TIPOVIA) che individua la tipologia del viaggio. Nel caso di più componenti familiari partecipanti allo stesso viaggio, coincidono tutte le caratteristiche del viaggio (la data di inizio, la durata, il mezzo, la destinazione, ecc.), ad eccezione della tipologia del viaggio (TIPOVIA): è infatti ammesso che alcuni dei partecipanti abbiano viaggiato per motivi di lavoro e altri per motivi personali (vacanza, ecc.).

Per una corretta utilizzazione dei dati, si rammenta che l'indagine rileva i viaggi che si concludono in ciascun trimestre di riferimento, indipendentemente dalla data di inizio.

Per un corretto utilizzo dei file si precisa che è possibile effettuare elaborazioni sulle seguenti unità di analisi:

- a) **famiglie**: volendo analizzare le famiglie occorre selezionare solo il primo componente di ciascuna, utilizzando il numero d'ordine all'interno della famiglia (IDIND=01). Ogni famiglia è individuata da: trimestre di riferimento (TRIM), numero progressivo della famiglia all'interno del trimestre (IDFAM);
- b) **individui**: ogni individuo è identificato da: trimestre di riferimento (TRIM), numero progressivo della famiglia (IDFAM) all'interno del trimestre, numero d'ordine dell'individuo all'interno della famiglia (IDIND);
- c) **viaggi**: ogni viaggio è individuato da: trimestre di riferimento (TRIM), numero progressivo della famiglia all'interno del trimestre (IDFAM), numero progressivo del viaggio all'interno della famiglia (IDVIA);
- d) **notti**: le notti trascorse in viaggio sono espresse dalla durata del viaggio (DURATA) presente nel file viaggi.

Le informazioni riportate nei file sono di carattere campionario. In particolare, ciascun sotto-file trimestrale si riferisce ad un campione di circa 3.500 famiglie che viene selezionato in modo indipendente dai campioni delle altre tre rilevazioni.

Dato il carattere campionario dell'indagine, per ottenere stime relative all'intera popolazione di riferimento (le famiglie residenti in Italia e gli individui che le compongono, al netto dei membri permanenti delle convivenze), è necessario procedere alla ponderazione dei dati elementari, attraverso i coefficienti di riporto all'universo (contenuti, in ciascun record, nella variabile COE). I valori contenuti in questa variabile, calcolati separatamente per ciascun trimestre, rappresentano il peso che ogni unità campionaria ha nella popolazione e nel trimestre di riferimento. Il COE è espresso con due valori decimali separati dal punto (es. "04782.72"). Per l'elaborazione dei dati si consiglia di utilizzare il COE arrotondato all'unità. I coefficienti di riporto all'universo sono stati determinati in modo da poter essere utilizzati indifferente per costruire stime trimestrali relative alle famiglie, agli individui, ai viaggi e alle notti.

Ad esempio, se in un record individuale del primo trimestre la variabile COE vale "04782.72", significa che l'individuo rappresenta 4.782,72 individui nella popolazione residente in Italia nel primo trimestre dell'anno. Ponderando quindi per questo valore i dati presenti nel suo record, le caratteristiche relative a tale individuo saranno stimate per 4.782,72 individui. Di conseguenza, ponderando i dati di tutti gli individui campionati in un trimestre, si otterranno stime riferite all'intera popolazione residente in Italia in quel trimestre. Analogamente, ponderando i dati contenuti nel file viaggi sarà possibile ottenere stime sui viaggi effettuati dai residenti in Italia con riferimento a ciascun trimestre di indagine.

Per quanto appena detto, è importante sottolineare che, mentre i viaggi e le notti rilevati in ciascun trimestre possono essere sommati per ottenere la stima annuale, non è corretto sommare i file trimestrali degli individui, per non quadruplicare la popolazione. Per evitare ciò, qualora si volessero comunque utilizzare dati sugli individui rilevati in tutti e quattro i trimestri, sarà necessario far riferimento ad un valore medio di popolazione, utilizzando il coefficiente COE diviso per 4. Analogamente, i dati trimestrali sul numero dei turisti (cioè il numero di individui che hanno fatto almeno un viaggio in un dato trimestre) non sono sommabili, in quanto una stessa persona può essere stata turista in trimestri diversi. Un'eventuale somma dei turisti trimestrali comporterebbe, pertanto, una sovrastima del numero annuale dei turisti.

L'indagine ha la finalità di fornire stime di parametri di diversa natura con riferimenti temporali e territoriali differenti. In particolare, il disegno campionario consente di ottenere stime trimestrali con riferimento soltanto all'intero territorio nazionale, mentre per le stime annuali sono previsti i seguenti dettagli territoriali:

1. l'intero territorio nazionale;
2. le tre grandi ripartizioni geografiche (Nord, Centro, Mezzogiorno);
3. le regioni geografiche (solamente per alcune stime più importanti);
4. le sei tipologie socio-demografiche dei comuni.

Trattandosi di un'indagine di natura campionaria, l'utilizzatore dei dati dovrà valutare l'attendibilità delle stime ottenibili, sfruttando tutte le informazioni e gli strumenti di calcolo messi a disposizione, per procedere ad una corretta interpretazione dei dati. A tale scopo, si riportano, in Appendice C, le informazioni relative alla strategia di campionamento e alla valutazione degli errori di stima, con esempi di calcolo.

APPENDICE A - IL QUESTIONARIO

Il questionario, somministrato con tecnica CATI, è strutturato in Sezioni in modo da garantire continuità di contenuto fra le domande e da facilitare la concentrazione del rispondente, al quale viene chiesto di rispondere su un certo argomento fino a che questo non sia stato del tutto esaurito e solo successivamente di passare ad un altro. All'interno del questionario sono quindi individuabili una prima fase di presentazione dell'intervista, una seconda fase in cui vengono raccolti i dati individuali ed una terza fase in cui vengono raccolti i dati sui viaggi.

La fase di presentazione rappresenta una componente cruciale dell'intervista. L'intervistatore si presenta specificando che l'indagine è effettuata per conto dell'Istat, riassume il contenuto dell'indagine e ricorda che le notizie raccolte sono coperte da segreto d'ufficio. Verifica, inoltre, che la famiglia abbia ricevuto la lettera firmata dal Presidente dell'Istat che avvertiva dell'intervista telefonica e spiegava dettagliatamente i motivi dell'indagine. Successivamente si passa a valutare l'eleggibilità della famiglia contattata. Essendo il target dell'indagine costituito dalle famiglie, si verifica in primo luogo che l'intestatario del contratto telefonico sia effettivamente una famiglia e non un'impresa, un'azienda, un ufficio o un gruppo di individui. Per evitare, inoltre, ridondanze all'interno del campione si escludono anche le dimore non abituali della famiglia, come ad esempio le seconde case.

Una volta accertata l'eleggibilità della famiglia, viene intervistata una persona, di almeno 18 anni (esclusi domestici, amici o visitatori casuali), che fornirà informazioni per sé e per ciascuno degli altri componenti della famiglia. Quindi, per ogni famiglia del campione, viene realizzata l'intervista "diretta" per un solo componente (persona di riferimento), mentre per ciascun altro componente l'intervista è di tipo "proxy".

Il questionario si compone delle seguenti sezioni:

- "Sezione generale", dove vengono rilevate le informazioni strutturali di ciascun componente della famiglia;
- "Scheda riepilogativa viaggi", dove viene riportato il numero dei viaggi conclusi da ciascun componente nel trimestre di riferimento;
- "Scheda viaggio di lavoro", dove vengono rilevate le informazioni per ciascun viaggio effettuato dai componenti della famiglia per motivi professionali;
- "Scheda viaggio di vacanza", dove vengono rilevate le informazioni per ciascun viaggio effettuato dai componenti della famiglia per motivi di vacanza/personali.

Di seguito è riportato l'elenco dei quesiti posti all'intervistato in ciascuna sezione.

APPENDICE B - ALCUNE DEFINIZIONI

Le **ripartizioni geografiche** sono una suddivisione amministrativa del territorio italiano e sono così definite:

- **Nord**: Piemonte, Valle d'Aosta, Lombardia, Liguria, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna
- **Centro**: Toscana, Umbria, Marche, Lazio
- **Mezzogiorno**: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna.

La **tipologia socio-demografica del comune (dominio)** è una suddivisione in classi dei comuni italiani, in base all'ampiezza demografica e alle caratteristiche socio-economiche:

- **comuni centro dell'area metropolitana**: Torino, Milano, Venezia, Genova, Bologna, Firenze, Roma, Napoli, Bari, Palermo, Catania, Cagliari;
- **comuni che gravitano intorno al centro dell'area metropolitana** come definiti sulla base dei risultati del Censimento della Popolazione;
- **altri comuni** suddivisi per dimensione demografica (oltre i 50.000 abitanti, da 10.001 a 50.000 abitanti, da 2.001 a 10.000 abitanti, fino a 2.000 abitanti).

Le caratteristiche socio-demografiche dell'individuo (stato civile, titolo di studio, condizione professionale, posizione nella professione, attività economica) fanno riferimento alla situazione al momento dell'intervista e sono quelle standard dell'Istat. In particolare:

- **l'età** è espressa in anni compiuti;
- **la condizione professionale** è quella dichiarata come unica o prevalente;
- **occupato** è chi svolge un'occupazione in proprio o alle dipendenze da cui trae un profitto o una retribuzione (utile, onorario, salario, stipendio, ecc.) o collabora con un familiare che svolge attività lavorativa in conto proprio, senza avere un regolare contratto di lavoro (coadiuvante);
- **in cerca di nuova occupazione** è chi avendo perduto una precedente occupazione è alla ricerca attiva di una nuova occupazione ed è in grado di accettarla se gli viene offerta;
- **in cerca di prima occupazione** è chi avendo concluso o sospeso o abbandonato un ciclo di studi e chi non avendo mai esercitato un'attività lavorativa è alla ricerca attiva di un'occupazione ed è in grado di accettarla se gli viene offerta;
- **casalinga** è chi si dedica prevalentemente alle cure della propria famiglia o della propria casa;
- **studente** è chi si dedica prevalentemente allo studio;
- **ritirato dal lavoro** è chi ha cessato un'attività lavorativa per raggiunti limiti di età, invalidità o altra causa; la figura di ritirato dal lavoro non coincide necessariamente con quella di pensionato in quanto non sempre il ritirato dal lavoro gode di una pensione;
- **in altra condizione** è chi si trova in condizione diversa da quelle sopra elencate, e dalla condizione di inabile al lavoro o militare in servizio di leva/servizio civile (ad es. benestante, pensionato per motivi diversi dall'attività lavorativa, detenuto agli arresti domiciliari, ecc.).

APPENDICE C - STRATEGIA DI CAMPIONAMENTO E VALUTAZIONE DEGLI ERRORI CAMPIONARI

1. Introduzione

La popolazione di interesse dell'indagine – ossia l'insieme delle unità statistiche oggetto di investigazione – è costituita dalle famiglie residenti in Italia e dagli individui che le compongono, al netto dei membri permanenti delle convivenze. Per famiglia si intende la famiglia di fatto, ossia un insieme di persone coabitanti e legate da vincoli di matrimonio, parentela, affinità, adozione, tutela o affettivi.

L'indagine viene svolta mediante intervista telefonica e utilizza come lista di selezione l'archivio ufficiale nazionale degli abbonati alle utenze private di telefonia fissa; le unità di campionamento sono, pertanto, i numeri telefonici appartenenti a detto archivio.

Nel corso di un anno sono previste quattro rilevazioni, nei mesi di aprile, luglio, ottobre e gennaio. Ciascuna rilevazione ha come periodo di riferimento il trimestre immediatamente precedente. Ad esempio, la rilevazione di aprile si riferisce al primo trimestre dell'anno.

Ciascuna delle quattro rilevazioni trimestrali utilizza un campione di circa 3.500 famiglie che viene selezionato in modo indipendente dai campioni delle altre tre rilevazioni.

L'indagine ha la finalità di fornire stime di parametri di diversa natura (totali, medie, rapporti, frequenze assolute e relative) con diversi riferimenti temporali e territoriali. Le stime relative ai singoli trimestri si riferiscono all'intero territorio nazionale, mentre i domini territoriali di riferimento per le stime relative all'intero anno sono:

- l'intero territorio nazionale;
- le tre grandi ripartizioni geografiche (Nord, Centro e Mezzogiorno);
- le regioni geografiche (solamente per alcune stime più importanti);
- la tipologia comunale, ottenuta suddividendo i comuni italiani in sei classi formate in base a caratteristiche socio-economiche e demografiche (sulla base della popolazione residente per l'anno di riferimento):

A, *area metropolitana* suddivisa in :

A₁, comuni centro dell'area metropolitana: Torino, Milano, Venezia, Genova, Bologna, Firenze, Roma, Napoli, Bari, Palermo, Catania e Cagliari;

A₂, comuni che gravitano intorno al centro dell'area metropolitana;

B, *area non metropolitana* suddivisa in:

B₁, comuni aventi fino a 2.000 abitanti;

B₂, comuni con 2.001-10.000 abitanti;

B₃, comuni con 10.001-50.000 abitanti;

B₄, comuni con oltre 50.000 abitanti.

La base di campionamento adottata, ossia la lista di selezione delle unità campionarie, è l'archivio informatizzato ufficiale degli abbonati alle utenze private di telefonia fissa. Tale scelta è motivata dal fatto che le informazioni dell'archivio in oggetto sono contenute in un file che viene costantemente aggiornato sulle variazioni degli intestatari e degli indirizzi telefonici; esso è, inoltre, di agevole utilizzo per la selezione delle unità campionarie in quanto si presta facilmente alla scelta di diversi criteri di ordinamento.

Le informazioni relative a ciascun indirizzo, utilizzabili per la stratificazione delle unità della popolazione di riferimento, sono essenzialmente di tipo territoriale; esse sono: la provincia, il comune, la sezione di censimento, la via, il numero civico e infine l'ampiezza del comune di appartenenza definita sia in termini demografici sia in termini di numero di indirizzi.

I principali problemi che derivano dall'utilizzo della lista telefonica come base di campionamento per le indagini sulle famiglie, sono legati alla non perfetta coincidenza tra la popolazione oggetto d'indagine e l'insieme delle unità contenute nella lista; indicando con L ed U rispettivamente la lista e la popolazione d'interesse, le possibili situazioni in cui gli insiemi L e U differiscono sono in generale le seguenti:

- I. sovracopertura, quando alcuni elementi dell'insieme L non appartengono all'insieme U;
- II. sottocopertura, nel caso in cui alcuni elementi di U non sono contenuti in L;
- III. duplicazione di alcune unità, se alcuni elementi di U sono presenti più volte in L;
- IV. grappoli di unità, quando alcuni elementi dell'insieme L contengono grappoli di elementi dell'insieme U.

La lista telefonica utilizzata per l'indagine presenta in misura diversa tutte le situazioni elencate; i fenomeni che appaiono più rilevanti sono, tuttavia, quelli della sovracopertura e della sottocopertura. Rientrano, infatti, nel caso della sovracopertura tutti i numeri telefonici che corrispondono a seconde case oppure ad attività professionali o a pubblici esercizi, ecc.; mentre rientrano nel caso della sottocopertura tutte le famiglie non abbonate al telefono o intestatarie di numeri riservati. Rientrano, poi, nel caso III le famiglie intestatarie di più di un numero riferito all'abitazione principale e nel caso IV i numeri telefonici a cui corrispondono più famiglie.

L'effetto di ciascuna delle situazioni appena elencate è quello di modificare in modo non controllato le probabilità di inclusione delle unità della popolazione rispetto alle probabilità di inclusione teoriche assegnate alle stesse unità dal disegno di campionamento adottato. Ciò può essere causa di distorsioni nelle stime prodotte dall'indagine.

In particolare la sottocopertura determina una distorsione delle stime prodotte tanto maggiore quanto più, relativamente alle variabili d'interesse dell'indagine, le unità di U che appartengono a L sono differenti dalle unità di U non presenti in L. È possibile in parte attenuare tale effetto distorsivo ricorrendo a opportune procedure di stima, note in letteratura come stimatori di ponderazione vincolata⁶, che tengono conto di totali noti (desunti da fonti esterne all'indagine) della popolazione relativi a variabili ausiliarie correlate a quelle d'interesse.

Un effetto dovuto al fenomeno della sovracopertura, è il fatto che la numerosità realizzata del campione sia minore di quella definita nella fase di progettazione del disegno e ciò comporta un aumento di variabilità delle stime. Si tratta allora di decidere se accettare una dimensione campionaria inferiore a quella prevista o, alternativamente, ricorrere a un sovra-dimensionamento del campione o a una lista di unità sostitutive. Nelle indagini di tipo telefonico, in cui la rilevazione è affidata ad una ditta esterna si ricorre in genere alla sostituzione in quanto è necessario garantire che tale ditta svolga un numero prefissato di interviste utili. Al fine di evidenziare il fenomeno della sostituzione, nei prospetti 1 e 2 vengono presentati i tassi di sostituzione, rispettivamente per regione e per tipologia comunale, osservati nel trimestre di indagine ottobre-dicembre.

3. Disegno di campionamento

3.1 Descrizione generale

Il disegno di campionamento è un disegno stratificato a grappoli, in cui ciascun grappolo è un indirizzo telefonico presente nell'archivio di selezione. Tutti gli individui appartenenti ai grappoli campione, cioè appartenenti alle famiglie a cui corrispondono i numeri selezionati, vengono inclusi nel campione.

Gli indirizzi telefonici sono stati stratificati, all'interno di ogni regione geografica, secondo la tipologia socio-demografica dei comuni (descritta nel paragrafo 1) di appartenenza degli indirizzi. Poiché tale informazione non è presente nell'archivio di selezione, è stato necessario integrare l'archivio attribuendo ad ogni record il relativo codice di tipologia.

La determinazione del numero totale di unità campionarie e la sua allocazione tra gli strati, è in genere, per un'indagine ad obiettivi plurimi come quella in esame, un'operazione complessa. È poco realistico, infatti, pensare di poter definire un campione che assicuri prefissati livelli di precisione per tutte le stime d'interesse, considerando anche il fatto che le stime vengono prodotte con diversi riferimenti temporali e territoriali. L'allocazione ottimale delle unità del campione con riferimento ad un dato tipo di dominio può risultare contrastante con l'allocazione ottimale con riferimento ad un altro tipo di dominio territoriale. Infatti, per quanto riguarda le stime riferite all'intero territorio nazionale l'allocazione ottimale risulta vicina a quella proporzionale tra le diverse regioni; per quanto riguarda, invece, le stime riferite alle regioni, l'allocazione ottimale risulta prossima a quella che assegna a tutte le regioni un campione di uguale numerosità. È necessario quindi ricorrere a un procedimento complesso articolato in più fasi.

⁶ Noti nella letteratura in lingua anglosassone sul tema come Deville J.C. e Särndal C.E. "Calibration Estimators in Survey Sampling", *Journal of the American Statistical Association*, 87 (1992): 1013-1020.

Dapprima, mediando tra esigenze operative e di costo ed esigenze relative all'attendibilità delle principali stime di interesse, si è definito un numero complessivo di indirizzi campione n pari a circa 3.500. Successivamente, sulla base di valutazioni dell'errore di campionamento atteso delle principali stime a livello regionale e nazionale, è stata determinata l'allocazione del campione tra le regioni; si è ottenuta in tal modo un'allocazione di compromesso tra l'allocazione uniforme e quella proporzionale al peso demografico di ciascuna regione. Infine, le numerosità campionarie regionali sono state ripartite tra le diverse tipologie di comune in modo proporzionale alla popolazione residente.

Nei prospetti 1 e 2 sono riportati, rispettivamente per regione e per tipologia di comune, le numerosità campionarie teoriche e i tassi di sostituzione riferiti alla rilevazione del quarto trimestre d'indagine; nell'ambito di ciascuna regione il tasso di sostituzione è calcolato come rapporto percentuale tra il numero di indirizzi sostituiti ed il numero teorico di indirizzi campione.

Prospetto 1 - Distribuzione regionale del campione e tassi di sostituzione - Quarto trimestre 2013

REGIONI	Indirizzi campione	Tasso di sostituzione totale (a)	Tasso di sostituzione al netto dei mancati contatti (b)
Piemonte	266	21,1	10,2
Valle d'Aosta	92	31,5	16,3
Lombardia	330	29,1	16,7
Trentino-Alto Adige	207	22,2	10,1
<i>Bolzano-Bozen</i>	100	11,0	3,0
<i>Trento</i>	107	32,7	16,8
Veneto	202	28,7	14,9
Friuli-Venezia Giulia	139	20,1	11,5
Liguria	152	34,9	19,1
Emilia-Romagna	198	27,8	15,7
Toscana	213	28,2	16,4
Umbria	104	20,2	7,7
Marche	141	29,1	16,3
Lazio	219	36,5	18,3
Abruzzo	143	23,8	12,6
Molise	96	19,8	9,4
Campania	197	33,0	22,3
Puglia	187	23,5	13,4
Basilicata	99	20,2	9,1
Calabria	158	27,2	13,9
Sicilia	212	31,1	15,1
Sardegna	151	37,1	21,2
Italia	3506	27,7	14,9

(a) Il tasso di sostituzione totale è calcolato come rapporto percentuale tra il numero di famiglie sostituite ed il numero teorico di famiglie campione. Esso esclude le sostituzioni rese necessarie dagli 'errori' di lista, molto elevati in alcune zone soprattutto per la presenza di numerici telefonici riferiti ad abitazioni non principali (seconde case).

(b) Il tasso di sostituzione al netto dei mancati contatti è calcolato come rapporto percentuale tra il numero di famiglie sostituite, depurate dalle sostituzioni per mancato contatto, ed il numero teorico di famiglie campione. Per mancato contatto con la famiglia ci si riferisce ai casi di 7 tentativi di contatto telefonico con esito 'libero', effettuati in giorni e orari diversi.

Prospetto 2 - Distribuzione del campione per tipologia di comune e tassi di sostituzione - Quarto trimestre 2013

TIPOLOGIA DEL COMUNE	Indirizzi campione	Tasso di sostituzione totale	Tasso di sostituzione al netto dei mancati contatti
Comuni centro dell'area metropolitana	413	37,0	17,4
Comuni che gravitano intorno al centro dell'area metropolitana	323	35,0	20,7
Comuni aventi fino a 2.000 abitanti	286	22,7	9,4
Comuni con 2.001-10.000 abitanti	920	23,4	12,9
Comuni con 10.001-50.000 abitanti	926	25,5	14,3
Comuni con oltre 50.000 abitanti	638	29,5	16,3
Italia	3506	27,7	14,9

3.2. Selezione del campione

L'estrazione degli indirizzi campione da ciascuno strato, ottenuto come incrocio della regione e della tipologia comunale, è stata effettuata con probabilità uguali e senza reimmissione mediante tecnica di selezione sistematica. La selezione del campione di indirizzi avviene nel seguente modo:

- per ciascuno strato si ordinano gli indirizzi dell'Archivio Telefonico delle Famiglie per ampiezza demografica del comune, strada, numero civico e piano;

- si definisce il passo di estrazione p come rapporto tra il numero di indirizzi in archivio e il numero di indirizzi da estrarre; si seleziona un numero casuale compreso tra 1 e p e si procede, poi, alla selezione degli indirizzi campione sulla base del passo di estrazione a partire dal punto di partenza casuale;
- ogni indirizzo selezionato dall'Archivio entra a far parte dell'*elenco base* e per ognuno di essi si selezionano i tre indirizzi successivi, che entrano a far parte dell'*elenco sostitutivo*; in tal modo le unità sostitutive sono le più vicine possibile dal punto di vista territoriale alle rispettive unità del campione base e quindi, presumibilmente, sono anche *simili* dal punto di vista delle caratteristiche socio-economiche.

4. Procedimento per il calcolo delle stime

Le stime prodotte dall'indagine sono stime del numero di viaggi e del numero di notti trascorse fuori casa per tipologia di viaggio, nonché del numero e delle caratteristiche di individui che effettuano tali viaggi.

Tali stime sono ottenute mediante uno stimatore di ponderazione vincolata (cfr. par. 2), che è il metodo di stima standard per la maggior parte delle indagini Istat sulle imprese e sulle famiglie.

Il principio su cui è basato ogni metodo di stima campionaria è che le unità appartenenti al campione rappresentino anche le unità della popolazione che non sono incluse nel campione.

Questo principio viene realizzato attribuendo a ogni unità campionaria un peso che indica il numero di unità della popolazione rappresentate dall'unità medesima. Se, ad esempio, a un'unità campionaria viene attribuito un peso pari a 100, vuol dire che questa unità rappresenta se stessa ed altre 99 unità della popolazione che non sono state incluse nel campione.

Al fine di rendere più chiara la successiva esposizione, introduciamo la seguente simbologia: d , indice di livello territoriale di riferimento delle stime; t , indice di trimestre ($t=1, \dots, 4$); j , indice di famiglia; p , indice del componente della famiglia; h , indice di strato; y , generica variabile oggetto di indagine; Y_{thjp} valore di y osservato sul componente p della famiglia j dello strato h rilevata nel trimestre t ; P_{hj} , numero di componenti della famiglia j dello strato h ; M_h , numero di famiglie residenti nello strato h ; m_h , campione di famiglie nello strato h ; H_d , numero di strati nel dominio d .

Ipotizziamo di voler stimare, con riferimento a un generico dominio d (ad esempio una regione geografica) e relativamente a un generico trimestre t , il totale della variabile y , espresso dalla seguente relazione:

$${}_dY_t = \sum_{h=1}^{H_d} \sum_{j=1}^{M_h} \sum_{p=1}^{P_{hj}} Y_{thjp} \quad (1)$$

Una stima del totale (1) è data dalla seguente espressione:

$${}_d\hat{Y}_t = \sum_{h=1}^{H_d} \hat{Y}_h = \sum_{h=1}^{H_d} \sum_{j=1}^{m_h} \sum_{p=1}^{P_{hj}} Y_{thjp} W_{thj}, \quad (2)$$

in cui W_{thj} è il peso finale da attribuire alla famiglia j dello strato h e a tutti i componenti a essa appartenenti.

Dalla precedente relazione si desume, quindi, che per ottenere la stima del totale (1) occorre moltiplicare il peso finale associato a ciascuna unità campionaria per il valore della variabile y assunto da tale unità ed effettuare, a livello del dominio di interesse, la somma dei prodotti così ottenuti.

Il peso da attribuire alle unità campionarie è ottenuto per mezzo di una procedura complessa che ha le seguenti finalità:

- correggere l'effetto distorsivo dovuto agli errori di lista e al fenomeno della mancata risposta totale;

- tenere conto della conoscenza di alcuni totali noti sulla popolazione oggetto di studio, nel senso che le stime campionarie di tali totali devono coincidere con i rispettivi valori noti. Nell'indagine in oggetto i totali noti⁷ sono:
 - popolazione residente per le 6 tipologie comunali;
 - popolazione residente per sesso nelle 20 regioni italiane;
 - popolazione residente per 8 classi di età⁸;
 - famiglie unipersonali di età inferiore a 65 anni, per sesso;
 - famiglie unipersonali di 65 anni e più, per sesso;
 - famiglie con più di un componente per classi di ampiezza⁹

Indicando, quindi, con ${}_kX$ il totale noto della k -ma variabile ($k=1\dots K$) ausiliaria e con ${}_kX_{thjp}$ il valore assunto dalla k -ma variabile ausiliaria per l'individuo rispondente $thjp$, la condizione sopra descritta è espressa dalla seguente uguaglianza:

$${}_kX_t = \hat{X}_t = \sum_{h=1}^H \sum_{j=1}^{m_h} \sum_{p=1}^{P_{hj}} {}_kX_{thjp} W_{thj} \quad k=1\dots K.$$

La procedura per la costruzione dei pesi finali da attribuire alle unità campionarie, è articolata nelle seguenti fasi :

- 1) viene dapprima calcolato il peso base (o peso diretto), ottenuto come reciproco della probabilità di inclusione di ogni unità campionaria;
- 2) si calcola quindi il fattore correttivo che consente di soddisfare la condizione di uguaglianza tra i totali noti della popolazione e le corrispondenti stime campionarie;
- 3) il peso finale è dato dal prodotto del peso base per i fattori correttivi sopra indicati.

Il fattore correttivo del punto 2) è ottenuto mediante la risoluzione di un problema di minimo vincolato, in cui la funzione da minimizzare è la distanza tra i pesi base ed i pesi finali e i vincoli sono definiti dalla condizione che le stime campionarie dei totali di popolazione sopra definiti coincidano con i valori noti degli stessi.

Con riferimento all'intero anno, i parametri d'interesse possono essere sia dei totali annui, ${}_dY$, che delle medie annue di dati trimestrali ${}_d\bar{Y}$. I totali annui - ottenuti come somma dei totali (1), relativi ai quattro trimestri - si riferiscono al totale annuo dei viaggi ed al totale annuo delle notti trascorse fuori casa per viaggi. Le medie annue di dati trimestrali - ottenute come media aritmetica dei totali (1) - si riferiscono, invece, al numero medio di persone che viaggiano nel trimestre. Con riferimento al generico dominio d , le stime dei suddetti parametri, ${}_dY$ e ${}_d\bar{Y}$, si ottengono rispettivamente mediante le seguenti espressioni:

$${}_d\hat{Y} = \sum_{t=1}^4 {}_d\hat{Y}_t, \quad (3)$$

$${}_d\hat{\bar{Y}} = \frac{1}{4} \sum_{t=1}^4 {}_d\hat{Y}_t \quad (4)$$

Alcuni parametri d'interesse, infine, sono dati dal rapporto tra totali; in tal caso la stima può essere ottenuta come rapporto tra i totali stimati:

$${}_d\hat{R} = \frac{{}_d\hat{Y}_1}{{}_d\hat{Y}_2}.$$

⁷ I primi tre totali sono desunti da fonti demografiche (anagrafiche), mentre i totali riferiti alle famiglie derivano da stime dell'indagine annuale multiscope "Aspetti della vita quotidiana".

⁸ Le classi di età considerate sono: 0-5 anni, 6-14 anni, 15-24 anni, 25-34 anni, 35-44 anni, 45-54 anni, 55-64 anni, 65 anni e più.

⁹ Le classi di ampiezza considerate sono: 2 componenti e 3 componenti e più.

5. Valutazione del livello di precisione delle stime

Le principali statistiche di interesse per valutare la variabilità campionaria delle stime prodotte dall'indagine sono l'errore di campionamento assoluto e l'errore di campionamento relativo.

Indicando con $\hat{\text{Var}}(\hat{Y}_t)$ la stima della varianza della stima \hat{Y}_t , riferita al dominio d e al trimestre di indagine t , la stima dell'errore di campionamento assoluto di \hat{Y}_t si può ottenere mediante la seguente espressione:

$$\hat{\sigma}(\hat{Y}_t) = \sqrt{\hat{\text{Var}}(\hat{Y}_t)} . \quad (5)$$

La stima dell'errore di campionamento relativo di \hat{Y}_t , è invece definita dall'espressione:

$$\hat{\varepsilon}(\hat{Y}_t) = \frac{\sqrt{\hat{\text{Var}}(\hat{Y}_t)}}{\hat{Y}_t} . \quad (6)$$

La stima della varianza, $\hat{\text{Var}}(\hat{Y}_t)$, viene ottenuta come somma della stima della varianza dei singoli strati appartenenti al dominio d ; in simboli:

$$\hat{\text{Var}}(\hat{Y}_t) = \sum_{h=1}^{H_d} \hat{\text{Var}}(\hat{Y}_{th}) = \sum_{h=1}^{H_d} M_h^2 \frac{M_h - m_h}{m_h M_h} \sum_{j=1}^{m_h} \frac{(e_{thj} - \bar{e}_{th})^2}{m_h - 1} \quad (7)$$

dove

$$e_{thj} = \sum_{p=1}^{P_{hj}} Y_{thjp} - \hat{\beta} \sum_{p=1}^{P_{hj}} X_{thjp} \quad \text{e} \quad \bar{e}_{th} = \frac{1}{m_h} \sum_{j=1}^{m_h} e_{thj} ;$$

in cui \underline{X}_{thjp} è il vettore delle K variabili ausiliarie utilizzate per la costruzione dello stimatore di ponderazione vincolata e $\hat{\beta}$ è il vettore dei coefficienti di regressione stimati del modello lineare che lega la variabile Y alle variabili ausiliarie \underline{X} .

Passiamo ora alla descrizione della metodologia per il calcolo degli errori di campionamento delle stime dei parametri riferiti all'intero anno. Per quanto riguarda la stima di un totale annuo (3), l'errore di campionamento assoluto e relativo, tenendo presente l'indipendenza dei quattro campioni, sono espressi rispettivamente dalle seguenti formule

$$\hat{\sigma}(\hat{Y}) = \sqrt{\sum_{t=1}^4 \hat{\text{Var}}(\hat{Y}_t)} , \quad \hat{\varepsilon}(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}} \quad (8)$$

Per quanto riguarda, invece, la stima di una media annua (4), l'errore di campionamento assoluto e relativo sono dati rispettivamente dalle seguenti espressioni:

$$\hat{\sigma}(\hat{\bar{Y}}) = \frac{1}{4} \sqrt{\sum_{t=1}^4 \hat{\text{Var}}(\hat{Y}_t)} , \quad \hat{\varepsilon}(\hat{\bar{Y}}) = \frac{\hat{\sigma}(\hat{\bar{Y}})}{\hat{\bar{Y}}} . \quad (9)$$

Infine, la stima degli errori relativi di stime di rapporto tra totali, nell'ipotesi semplificatrice che il rapporto sia indipendente dal valore del denominatore, può essere ottenuta come

$$\hat{\varepsilon}(\hat{R}) = \sqrt{\hat{\varepsilon}^2(\hat{Y}_1) - \hat{\varepsilon}^2(\hat{Y}_2)} . \quad (10)$$

Gli errori campionari consentono di valutare il grado di precisione delle stime; inoltre, l'errore assoluto permette di costruire l'intervallo di confidenza che con una certa fiducia P contiene il parametro di interesse. Con riferimento ad una generica stima \hat{Y} tale intervallo assume la seguente forma:

$$\{\hat{Y} - k\hat{\sigma}(\hat{Y}), \hat{Y} + k\hat{\sigma}(\hat{Y})\}$$

(11)

Nella (11) il valore di k dipende dal valore fissato per P ; ad esempio, per $P=0,95$ si ha $k=2$.

6. Presentazione sintetica degli errori campionari

Poiché a ciascuna stima ${}_d\hat{Y}_t$ corrisponde un errore campionario relativo $\hat{\varepsilon}({}_d\hat{Y}_t)$, per consentire un uso corretto delle informazioni prodotte dall'indagine sarebbe necessario pubblicare per ogni stima anche il corrispondente errore di campionamento relativo. Tuttavia sia per limiti di tempo e di costi di elaborazione, sia perché le tavole di pubblicazione risulterebbero appesantite e di non facile consultazione per l'utente finale, non è possibile pubblicare anche tutti gli errori di campionamento delle stime fornite. Inoltre, non sarebbero comunque disponibili gli errori delle stime non pubblicate, che l'utente può ricavare in modo autonomo.

Al fine di permettere comunque una valutazione della variabilità campionaria di tutte le stime d'interesse, si ricorre a una presentazione sintetica degli errori relativi basata su modelli regressivi; ossia fondata sulla determinazione di una funzione matematica che mette in relazione ciascuna stima con il proprio errore di campionamento. L'approccio utilizzato per la costruzione dei modelli è differente a seconda che la variabile oggetto di stima sia qualitativa ovvero quantitativa. Infatti, mentre per le stime di frequenze assolute (o relative) riferite alle modalità di variabili qualitative è possibile utilizzare modelli che hanno un fondamento teorico, secondo cui gli errori relativi delle stime di frequenze assolute sono funzione decrescente dei valori delle stime stesse, per le stime di totali di variabili quantitative il problema è piuttosto complesso, dal momento che non è stata ancora elaborata un'adeguata base teorica per l'interpolazione degli errori campionari delle stime in questione. L'approccio adottato per trattare il caso di variabili quantitative è pertanto di tipo empirico ed è fondato sull'evidenza sperimentale che l'errore assoluto di un totale è una funzione crescente del totale stesso.

Vengono di seguito descritti i modelli adottati per i due differenti gruppi di variabili.

6.1. Presentazione sintetica degli errori campionari per stime di frequenze riferite agli individui

Il modello utilizzato per le stime di frequenze assolute, con riferimento al generico dominio d e al trimestre t , è del tipo seguente:

$$\log \hat{\varepsilon}^2({}_d\hat{Y}_t) = a + b \log({}_d\hat{Y}_t) \quad (12)$$

in cui i parametri a e b vengono stimati utilizzando il metodo dei minimi quadrati. Dall'analisi dei coefficienti di regressione ottenuti è emerso che i coefficienti relativi al primo, secondo e quarto trimestre sono sostanzialmente simili tra loro, mentre differenti risultati si ottengono per il terzo trimestre; pertanto, per semplicità, vengono presentati solamente i risultati relativi al primo e al terzo trimestre. È stato inoltre stimato un modello analogo al (12) con riferimento alla media dei 4 trimestri.

I prospetti 3 (per i trimestri) e 4 (per la media dei 4 trimestri) riportano i valori dei coefficienti a e b e dell'indice di determinazione R^2 delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di frequenze riferite agli individui, per area territoriale.

Inoltre, allo scopo di facilitare il calcolo degli errori campionari, nei prospetti 5 (per i trimestri) e 6 (per la media dei 4 trimestri) sono riportati, per area territoriale, i valori interpolati degli errori campionari relativi percentuali di alcuni valori tipici assunti dalle stime di frequenze assolute.

La prima colonna dei prospetti riporta K valori crescenti delle stime ${}_d\hat{Y}_t^k$ ($k=1, \dots, K$), la seconda riporta i rispettivi errori relativi interpolati $\hat{\varepsilon}({}_d\hat{Y}_t^k)$. Le informazioni contenute nei prospetti permettono di calcolare l'errore relativo di una generica stima di frequenza assoluta mediante due procedimenti che risultano di facile applicazione, anche se conducono a risultati meno precisi di quelli ottenibili mediante l'applicazione diretta dell'espressione (12).

Il metodo consiste nell'approssimare l'errore relativo $\hat{\varepsilon}({}_d\hat{Y}_t)$ della stima di interesse ${}_d\hat{Y}_t$ con l'errore relativo corrispondente al livello stima, presente nel prospetto (5 o 6) che più si avvicina al valore della stima ${}_d\hat{Y}_t$.

Con il secondo metodo, l'errore campionario della stima ${}_d\hat{Y}_t$, si ricava mediante la seguente espressione:

$$\hat{\varepsilon}({}_d\hat{Y}_t) = \hat{\varepsilon}({}_d\hat{Y}_t^{k-1}) + \frac{\hat{\varepsilon}({}_d\hat{Y}_t^k) - \hat{\varepsilon}({}_d\hat{Y}_t^{k-1})}{{}_d\hat{Y}_t^k - {}_d\hat{Y}_t^{k-1}} ({}_d\hat{Y}_t - {}_d\hat{Y}_t^{k-1}) \quad (13)$$

dove: ${}_d\hat{Y}_t^{k-1}$ e ${}_d\hat{Y}_t^k$ sono i valori delle stime entro i quali è compresa la stima d'interesse ${}_d\hat{Y}_t$ e $\hat{\varepsilon}({}_d\hat{Y}_t^{k-1})$ e $\hat{\varepsilon}({}_d\hat{Y}_t^k)$ sono i corrispondenti errori relativi presenti nel prospetto.

6.2. Presentazione sintetica degli errori campionari per stime di totali riferiti ai viaggi e alle notti

Il modello utilizzato per le stime del totale del numero di viaggi e di notti, con riferimento al generico dominio d e al trimestre t , è del tipo seguente:

$$\hat{\sigma}({}_d\hat{Y}_t) = a + b \cdot {}_d\hat{Y}_t + c \cdot {}_d\hat{Y}_t^2 \quad (14)$$

dove i parametri a , b e c vengono stimati utilizzando il metodo dei minimi quadrati, adattando il modello (14) a una nuvola di punti costituita da un consistente numero di coppie $(\hat{\sigma}({}_d\hat{Y}_t), {}_d\hat{Y}_t)$. In considerazione del fatto che il modello (14) è di tipo empirico, l'insieme delle stime ${}_d\hat{Y}_t$ utilizzate per interpolare il modello è stato determinato in modo da includere la maggior parte delle stime pubblicate. Al fine di ottenere un migliore adattamento alla nuvola di punti, è stato interpolato un modello - con riferimento ad ogni trimestre e con riferimento all'intero anno - per ciascuno dei seguenti sottogruppi di stime:

- numero di notti per viaggi di lavoro;
- numero di notti per viaggi di vacanza;
- numero totale di notti in viaggio;
- numero di viaggi di lavoro;
- numero di viaggi di vacanza
- numero totale di viaggi.

Analogamente al caso precedente, dall'analisi dei coefficienti di regressione ottenuti è emerso che i coefficienti relativi al primo, al secondo e al quarto trimestre sono sostanzialmente uguali tra loro, mentre differiscono significativamente da quelli relativi al terzo trimestre; ciò è spiegabile considerando il fatto che i fenomeni indagati presentano andamenti difforni nel terzo trimestre rispetto agli altri tre.

Al fine di non appesantire la trattazione vengono qui, pertanto, riportati i risultati relativi al primo trimestre, al terzo trimestre e all'intero anno.

I prospetti 7, 9 e 11 (per le stime relative alle notti e ai viaggi) riportano i valori dei coefficienti a , b , c e dell'indice di determinazione R^2 dei modelli utilizzati per l'interpolazione degli errori campionari delle stime di totali per area territoriale e per ciascuno dei sei tipi di stime sopra elencati, con riferimento al primo trimestre, al terzo trimestre e all'intero anno.

Nel paragrafo successivo verrà illustrato come calcolare l'errore di campionamento assoluto e relativo di una stima a partire dalle informazioni contenute nei prospetti.

Utilizzando i valori stimati dei parametri del modello (14) e dividendo ambo i membri del modello per il valore della stima, ${}_d\hat{Y}_t$, si perviene alla seguente equazione di secondo grado:

$$a + [b - \varepsilon({}_d\hat{Y}_t)] \cdot {}_d\hat{Y}_t + c({}_d\hat{Y}_t)^2 = 0$$

la cui radice positiva è espressa dalla seguente formula:

$${}_d\hat{Y}_t = \frac{-[b - \varepsilon({}_d\hat{Y}_t)] - \sqrt{[b - \varepsilon({}_d\hat{Y}_t)]^2 - 4ac}}{2c} \quad (15)$$

Nei prospetti 8, 10 e 12 (per le stime relative alle notti e ai viaggi) vengono riportati i valori delle stime ${}_d\hat{Y}_t^*$ ottenuti sulla base della relazione (15) in corrispondenza di alcuni valori tipici prefissati dell'errore relativo percentuale, indicati come $\varepsilon^* = \varepsilon({}_d\hat{Y}_t)$; le stime con valori superiori a ${}_d\hat{Y}_t^*$ presentano valori dell'errore relativo inferiori a ε^* , mentre le stime che assumono valori inferiori a ${}_d\hat{Y}_t^*$ presentano valori dell'errore relativo superiori a ε^* . In tali prospetti, articolati per gruppo di stime,

trimestre e area territoriale, i valori prefissati di ε^* sono: 1, 5, 10, 15, 20, 25 e 30%. L'utilizzo di tali prospetti verrà meglio chiarito nel paragrafo seguente attraverso un esempio numerico.

Prospetto 3 - Valori dei coefficienti a, b e di R^2 delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di frequenze assolute riferite agli individui per aree territoriali - Trimestri – Anno 2013

AREE TERRITORIALI	Trimestri 1,2,e 4			Trimestre 3		
	a	b	R^2	a	b	R^2
Nord	10,65599	-1,11654	90,55380	10,33403	-1,05931	87,08463
Centro	11,00078	-1,15865	89,65229	10,06551	-1,05864	86,13075
Mezzogiorno	10,52173	-1,12317	88,98511	9,79958	-1,04370	86,31648
Italia	11,00256	-1,13479	92,14140	10,30754	-1,05217	92,38320

Prospetto 4 - Valori dei coefficienti a, b e di R^2 delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di frequenze assolute riferite agli individui per aree territoriali - Media dei 4 trimestri – Anno 2013

AREE TERRITORIALI	a	b	R^2
Nord	9,02025	-1,05455	91,8
Centro	9,87409	-1,14850	90,7
Mezzogiorno	9,22315	-1,09012	92,3
Italia	10,24850	-1,14229	94,2

Prospetto 5 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite agli individui per aree territoriali - Trimestri – Anno 2013

STIME	Trimestri 1,2 e 4				Trimestre 3			
	Aree territoriali				Aree territoriali			
	Nord	Centro	Mezzogiorno	Italia	Nord	Centro	Mezzogiorno	Italia
50.000	49,0	46,4	44,2	52,8	56,9	49,9	47,4	58,4
60.000	44,3	41,8	39,9	47,7	51,7	45,3	43,1	53,0
70.000	40,6	38,2	36,6	43,7	47,6	41,8	39,8	48,9
80.000	37,7	35,3	34,0	40,5	44,4	38,9	37,1	45,6
90.000	35,3	33,0	31,8	37,9	41,7	36,6	34,9	42,8
100.000	33,3	31,1	30,0	35,7	39,4	34,6	33,0	40,5
200.000	22,6	20,8	20,3	24,1	27,3	24,0	23,0	28,1
300.000	18,0	16,4	16,2	19,1	22,0	19,3	18,6	22,7
400.000	15,4	13,9	13,8	16,2	18,9	16,6	16,0	19,5
500.000	13,6	12,2	12,1	14,3	16,8	14,8	14,3	17,4
750.000	10,8	9,7	9,7	11,4	13,6	11,9	11,5	14,0
1.000.000	9,2	8,2	8,2	9,7	11,6	10,2	9,9	12,1
2.000.000	6,3	5,5	5,6	6,5	8,1	7,1	6,9	8,4
3.000.000	5,0	4,3	4,4	5,2	6,5	5,7	5,6	6,8
4.000.000	4,2	3,7	3,8	4,4	5,6	4,9	4,8	5,8
5.000.000	3,8	3,2	3,3	3,9	5,0	4,4	4,3	5,2
7.500.000	3,0	2,5	2,7	3,1	4,0	3,5	3,5	4,2
10.000.000	2,5	2,2	2,3	2,6	3,4	3,0	3,0	3,6
15.000.000	2,0	1,7	1,8	2,1	2,8	2,4	2,4	2,9
20.000.000	1,7	1,4	1,5	1,8	2,4	2,1	2,1	2,5
25.000.000	1,5	1,3	1,3	1,6	2,1	1,9	1,9	2,2

Prospetto 6 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite agli individui per aree territoriali - Media dei 4 trimestri – Anno 2013

STIME	Aree territoriali			
	Nord	Centro	Mezzogiorno	Italia
20.000	30,3	27,9	27,6	34,8
30.000	27,5	25,1	25,0	31,4
40.000	25,4	23,0	23,0	28,7
50.000	23,6	21,3	21,4	26,6
60.000	22,2	19,9	20,1	24,9
70.000	21,0	18,7	18,9	23,4
80.000	14,6	12,6	13,0	15,8
90.000	11,8	10,0	10,4	12,5
100.000	10,1	8,5	8,9	10,6
200.000	9,0	7,4	7,9	9,3
300.000	7,3	5,9	6,3	7,4
400.000	6,2	5,0	5,4	6,3
500.000	4,3	3,4	3,7	4,2
750.000	3,5	2,7	3,0	3,4
1.000.000	3,0	2,3	2,5	2,8
2.000.000	2,7	2,0	2,2	2,5
3.000.000	2,2	1,6	1,8	2,0
4.000.000	1,9	1,3	1,5	1,7
5.000.000	1,5	1,1	1,2	1,3
7.500.000	1,3	0,9	1,1	1,1
10.000.000	1,1	0,8	0,9	1,0

Prospetto 7 - Valori dei coefficienti a, b, c e di R² delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di totali per gruppi di stime, trimestre e anno, aree territoriali - Stime del numero di notti – Anno 2013

Notti – lavoro				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	7.412,418	0,559735637	-0,000000133524	95,17
Centro	-2.671,589	0,649924668	-0,000000201797	93,79
Mezzogiorno	6.515,030	0,652436699	-0,000000200990	93,71
Italia	32.568,166	0,376918784	-0,000000042240	95,74
Luglio-settembre				
Nord	18.112,074	0,535461988	-0,000000256474	93,34
Centro	-7.761,716	0,924586713	-0,000000554969	96,20
Mezzogiorno	7.601,992	0,726317438	-0,000000311374	80,56
Italia	47.136,247	0,398216216	-0,000000071087	88,70
Anno				
Nord	25.242,77	0,4221528951	-0,000000030737	82,27
Centro	35.488,10	0,4334456124	-0,000000061181	90,99
Mezzogiorno	55.107,51	0,3651623847	-0,000000040476	90,86
Italia	91.384,55	0,2410673251	-0,000000008112	81,26
Notti – vacanze				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	90.664,36	0,177459917	-0,000000004846	92,55
Centro	51.872,35	0,296201216	-0,000000011560	92,80
Mezzogiorno	12.151,62	0,455019034	-0,000000015776	89,83
Italia	155.201,05	0,172592346	-0,000000001946	87,24
Luglio-settembre				
Nord	722.059,41	0,1159850356	-0,000000000271	89,75
Centro	431.214,69	0,1546219014	-0,000000001242	82,70
Mezzogiorno	314.443,08	0,1646330943	-0,000000001388	89,25
Italia	946.071,00	0,0829027483	-0,000000000122	90,02
Anno				
Nord	846.525,44	0,1027802117	-0,000000000160	91,44
Centro	453.363,54	0,1417111173	-0,000000000705	89,38
Mezzogiorno	389.164,74	0,1405041463	-0,000000000783	91,45
Italia	1.124.098,31	0,0727333622	-0,000000000070	92,46
Notti - totali				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	102.660,505	0,172904720	-0,000000004360	93,53
Centro	59.880,360	0,287003510	-0,000000010471	93,88
Mezzogiorno	30.092,266	0,409628796	-0,000000013107	89,03
Italia	179.138,059	0,161	-0,000000001699	88,58
Luglio-settembre				
Nord	712.437,738	0,114132946	-0,000000000249	91,24
Centro	435.802,633	0,149469560	-0,000000001083	84,97
Mezzogiorno	318.209,820	0,159695549	-0,000000001206	91,08
Italia	921.881,321	0,081476529	-0,000000000111	91,47
Anno				
Nord	817.123,050	0,101572153	-0,000000000148	92,28
Centro	449.306,946	0,139529202	-0,000000000659	90,44
Mezzogiorno	380.727,636	0,137910925	-0,000000000741	91,27
Italia	1.075.196,176	0,071881368	-0,000000000065	93,00

Prospetto 8 - Valori dei totali corrispondenti ad alcuni valori tipici degli errori relativi percentuali per gruppi di stime, trimestre e anno, aree territoriali - Stime del numero di notti – Anno 2013

Notti – lavoro							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	4.130.585	3.832.059	3.459.155	3.086.626	2.714.625	2.343.398	1.973.374
Centro	3.166.952	2.968.453	2.720.273	2.472.010	2.223.638	1.975.115	1.726.375
Mezzogiorno	3.206.474	3.008.125	2.760.324	2.512.712	2.265.352	2.018.334	1.771.800
Italia	8.774.384	7.837.915	6.671.405	5.512.004	4.365.048	3.242.490	2.175.418
Luglio-settembre							
Nord	2.082.702	1.929.434	1.738.502	1.548.533	1.359.907	1.173.219	989.447
Centro	1.639.464	1.566.994	1.476.351	1.385.636	1.294.833	1.203.922	1.112.876
Mezzogiorno	2.311.071	2.183.227	2.023.531	1.863.985	1.704.630	1.545.526	1.386.756
Italia	5.579.949	5.030.247	4.347.584	3.672.271	3.008.732	2.365.320	1.758.662
Anno							
Nord	13.470.062	12.175.176	10.558.789	8.946.096	7.339.478	5.743.849	4.171.049
Centro	7.004.027	6.358.634	5.554.590	4.754.904	3.962.066	3.180.777	2.420.780
Mezzogiorno	8.927.241	7.957.575	6.752.788	5.560.699	4.390.635	3.262.543	2.222.513
Italia	28.874.007	24.021.976	18.014.808	12.152.889	6.735.013	2.850.633	1.313.258
Notti – vacanze							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	35.089.158	26.994.882	17.079.541	8.003.925	2.585.333	1.159.964	719.408
Centro	24.938.084	21.506.551	17.232.988	12.992.665	8.830.166	4.910.499	1.960.373
Mezzogiorno	28.235.800	25.702.999	22.537.847	19.374.067	16.212.459	13.054.590	9.903.998
Italia	84.495.184	64.238.037	39.330.773	16.455.980	4.330.909	1.912.984	1.196.287
Luglio-settembre							
Nord	397.233.073	253.632.745	88.848.011	18.497.761	8.368.214	5.330.365	3.901.468
Centro	119.350.542	88.173.348	50.811.444	20.586.359	7.826.263	4.282.309	2.894.579
Mezzogiorno	113.415.797	85.254.378	51.011.691	21.220.424	6.979.358	3.485.879	2.270.061
Italia	611.004.531	296.235.386	42.476.160	13.756.337	8.012.554	5.638.614	4.347.214
Anno							
Nord	587.126.267	344.214.291	81.805.817	16.950.830	8.585.658	5.714.483	4.277.407
Centro	190.233.570	134.875.048	68.555.392	20.153.819	7.158.221	4.078.342	2.828.525
Mezzogiorno	169.672.743	119.786.088	60.033.709	17.042.712	6.058.223	3.468.173	2.411.433
Italia	916.377.562	369.317.171	37.606.061	14.361.976	8.790.247	6.325.531	4.938.674
Notti – totali							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	37.981.553	28.999.685	18.026.572	8.144.138	2.654.747	1.244.073	786.522
Centro	26.668.658	22.884.065	18.173.764	13.507.400	8.948.053	4.740.283	1.849.994
Mezzogiorno	30.565.792	27.522.199	23.720.693	19.924.262	16.136.437	12.364.961	8.630.434
Italia	90.219.723	67.084.509	38.806.828	14.128.736	3.947.713	1.947.602	1.272.010
Luglio-settembre							
Nord	425.137.498	268.352.165	88.962.930	17.691.469	8.106.513	5.194.217	3.813.576
Centro	131.836.908	96.039.128	53.238.348	19.816.878	7.438.649	4.149.542	2.837.201
Mezzogiorno	126.261.356	93.805.101	54.371.415	20.758.006	6.594.417	3.371.954	2.225.440
Italia	656.755.869	310.411.700	40.123.672	13.172.513	7.722.216	5.450.780	4.209.684
Anno							
Nord	625.470.168	362.481.970	79.662.720	16.080.145	8.200.297	5.475.196	4.105.373
Centro	199.846.320	140.620.590	69.720.932	19.344.583	6.909.553	3.972.984	2.768.436
Mezzogiorno	175.554.833	122.827.482	59.762.266	15.933.313	5.738.915	3.323.630	2.324.185
Italia	965.937.443	379.020.237	35.341.056	13.609.018	8.356.645	6.023.122	4.706.986

Prospetto 9 - Valori dei coefficienti a, b, c e di R² delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di totali per gruppi di stime, trimestre e anno, aree territoriali - Stime del numero di viaggi – Anno 2013

Viaggi – lavoro				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	7.314,270	0,397380000	-0,000000154000	92,41
Centro	4.107,930	0,498040000	-0,000000215000	95,03
Mezzogiorno	6.727,020	0,440070000	-0,000000418000	93,30
Italia	12.981,710	0,262720000	-0,000000053000	93,64
Luglio-settembre				
Nord	8.456,437	0,549111533	-0,000000397191	88,67
Centro	4.787,925	0,610755962	-0,000001819983	94,87
Mezzogiorno	5.100,353	0,471497997	-0,000000844736	91,59
Italia	12.196,555	0,345974117	-0,000000148225	87,36
Anno				
Nord	17.641,40	0,2765364621	-0,000000055779	92,13
Centro	11.535,08	0,4428008001	-0,000000139364	94,76
Mezzogiorno	12.472,56	0,2932743974	-0,000000100056	93,07
Italia	27.711,48	0,1988731409	-0,000000018586	91,87
Viaggi – vacanza				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	14.887,422	0,265869007	-0,000000028893	91,54
Centro	20.071,594	0,308987568	-0,000000070267	94,72
Mezzogiorno	14.204,408	0,263510112	-0,000000095362	90,54
Italia	30.810,351	0,173732952	-0,000000010486	93,39
Luglio-settembre				
Nord	62.736,665	0,132268977	-0,000000004943	91,41
Centro	40.444,658	0,182282739	-0,000000021742	86,80
Mezzogiorno	28.483,826	0,194298272	-0,000000023354	91,14
Italia	85.136,563	0,094085391	-0,000000002137	91,03
Anno				
Nord	81.013,80	0,1185346595	-0,000000002098	90,90
Centro	54.415,75	0,1530098170	-0,000000007195	92,61
Mezzogiorno	40.577,33	0,1440651544	-0,000000007499	90,44
Italia	112.431,97	0,0811200147	-0,000000000831	91,24
Viaggi – totali				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	17.846,663	0,292150174	-0,000000030795	86,25
Centro	21.077,617	0,319033173	-0,000000060255	92,20
Mezzogiorno	15.421,366	0,253347894	-0,000000074188	89,49
Italia	35.351,659	0,181074109	-0,000000009693	87,98
Luglio-settembre				
Nord	58.320,456	0,133788077	-0,000000004911	91,13
Centro	35.799,166	0,184032660	-0,000000021883	86,59
Mezzogiorno	24.013,900	0,195390942	-0,000000023091	91,18
Italia	71.624,966	0,096868111	-0,000000002187	90,43
Anno				
Nord	58.236,89	0,1183897382	-0,000000002290	83,43
Centro	45.009,27	0,1590259605	-0,000000007457	85,19
Mezzogiorno	31.641,74	0,1443890856	-0,000000007949	86,55
Italia	75.696,03	0,0829825396	-0,000000000918	84,82

Prospetto 10 - Valori dei totali corrispondenti ad alcuni valori tipici degli errori relativi percentuali per gruppi di stime, trimestre e anno, aree territoriali - Stime del numero di viaggi – Anno 2013

Viaggi – lavoro							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	2.534.196	2.276.577	1.955.329	1.635.406	1.317.732	1.004.305	700.171
Centro	2.278.340	2.093.036	1.861.612	1.630.509	1.399.881	1.170.005	941.412
Mezzogiorno	1.044.286	950.120	832.887	716.411	601.103	487.711	377.704
Italia	4.819.128	4.073.711	3.147.996	2.236.320	1.363.089	629.252	255.451
Luglio-settembre							
Nord	1.372.820	1.273.324	1.149.246	1.025.595	902.541	780.351	659.468
Centro	337.875	316.425	289.718	263.162	236.802	210.705	184.969
Mezzogiorno	557.159	510.791	453.105	395.843	339.200	283.507	229.346
Italia	2.302.391	2.037.183	1.707.652	1.381.694	1.062.276	756.290	481.172
Anno							
Nord	4.843.769	4.137.788	3.261.911	2.400.314	1.573.189	848.495	390.100
Centro	3.131.965	2.847.587	2.492.949	2.139.659	1.788.483	1.440.875	1.099.910
Mezzogiorno	2.874.536	2.481.624	1.994.180	1.514.269	1.050.850	630.282	321.054
Italia	10.306.807	8.191.989	5.586.663	3.109.127	1.191.123	463.812	261.463
Viaggi – vacanze							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	8.913.403	7.539.541	5.829.100	4.134.826	2.486.905	1.043.159	338.936
Centro	4.321.125	3.761.699	3.067.318	2.382.514	1.717.377	1.099.318	602.226
Mezzogiorno	2.713.293	2.303.603	1.797.491	1.304.491	842.738	463.226	239.440
Italia	15.800.364	12.043.764	7.427.155	3.185.633	870.475	383.734	239.256
Luglio-settembre							
Nord	25.236.931	17.372.881	8.095.448	2.194.998	870.905	521.464	369.997
Centro	8.152.088	6.375.912	4.224.784	2.295.252	1.016.007	512.820	324.166
Mezzogiorno	8.043.201	6.370.244	4.320.127	2.404.147	989.037	432.821	255.096
Italia	40.336.930	22.408.475	5.077.995	1.443.036	791.193	542.020	411.697
Anno							
Nord	52.468.230	33.808.705	12.041.271	2.240.109	970.215	610.293	444.161
Centro	20.249.531	14.826.778	8.280.810	2.967.174	1.003.753	539.456	363.724
Mezzogiorno	18.176.467	12.961.888	6.685.845	1.963.917	665.980	373.183	257.043
Italia	87.136.564	40.767.707	4.898.808	1.601.351	939.589	663.584	512.672
Viaggi – totali							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	9.225.065	7.936.348	6.331.215	4.738.340	3.174.921	1.708.038	644.413
Centro	5.196.040	4.541.896	3.728.887	2.924.873	2.139.012	1.396.213	770.105
Mezzogiorno	3.342.348	2.814.832	2.163.117	1.529.006	940.185	479.048	239.412
Italia	17.852.677	13.786.437	8.779.199	4.096.038	1.168.533	480.433	290.385
Luglio-settembre							
Nord	25.669.667	17.731.563	8.309.473	2.170.422	829.751	491.632	347.316
Centro	8.153.468	6.381.296	4.227.079	2.274.460	965.216	469.544	292.550
Mezzogiorno	8.156.167	6.457.444	4.369.090	2.399.193	924.856	379.003	218.974
Italia	40.525.376	22.861.209	5.051.223	1.280.557	684.560	464.650	351.275
Anno							
Nord	47.864.124	30.693.724	10.461.574	1.646.056	699.854	439.140	319.383
Centro	20.283.551	15.023.325	8.616.544	3.135.561	938.274	476.165	314.057
Mezzogiorno	17.139.587	12.201.240	6.224.115	1.673.223	528.988	293.139	201.270
Italia	80.485.844	38.075.260	3.706.621	1.112.534	643.627	452.098	348.288

Prospetto 11 - Valori dei coefficienti a, b, c e di R^2 delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di totali per gruppi di stime e tipologie comunali - Stime del numero di viaggi - Anno 2013

TIPOLOGIE COMUNALI	Viaggi – lavoro			R^2
	a	b	c	
Comuni centro dell'area metropolitana	15.169,265	0,391417929	-0,000000112726	93,88
Comuni periferia dell'area metropolitana	12.853,115	0,409269599	-0,000000288243	91,02
Comuni con oltre 50.000 abitanti	8.826,475	0,352582912	-0,000000295555	88,60
Comuni da 10.001 a 50.000 abitanti	8.764,279	0,446514211	-0,000000149786	92,60
Comuni da 2.001 a 10.000 abitanti	9.205,322	0,514061757	-0,000000297898	86,08
Comuni aventi fino a 2.000 abitanti	-254,031	0,999459506	-0,0000006677270	90,66
TIPOLOGIE COMUNALI	Viaggi – vacanze			R^2
	a	b	c	
Comuni centro dell'area metropolitana	29.046,367	0,200566360	-0,000000011045	82,79
Comuni periferia dell'area metropolitana	23.805,080	0,314030597	-0,000000022126	85,27
Comuni con oltre 50.000 abitanti	22.623,396	0,216324553	-0,000000013619	84,22
Comuni da 10.001 a 50.000 abitanti	25.774,703	0,182354047	-0,000000009966	81,96
Comuni da 2.001 a 10.000 abitanti	24.175,689	0,203299993	-0,000000012250	82,88
Comuni aventi fino a 2.000 abitanti	8.828,101	0,420154720	-0,000000160706	86,45
TIPOLOGIE COMUNALI	Viaggi – totali			R^2
	a	b	c	
Comuni centro dell'area metropolitana	175.388,663	0,079022043	-0,000000001491	89,01
Comuni periferia dell'area metropolitana	90.866,430	0,115702454	-0,000000001308	92,82
Comuni con oltre 50.000 abitanti	48.371,215	0,151820462	-0,000000007202	93,73
Comuni da 10.001 a 50.000 abitanti	137.602,779	0,081224827	-0,000000001826	91,26
Comuni da 2.001 a 10.000 abitanti	99.555,784	0,115474720	-0,000000004105	95,41
Comuni aventi fino a 2.000 abitanti	17.994,051	0,219637161	-0,000000051317	90,72

Prospetto 12 - Valori dei totali corrispondenti ad alcuni valori tipici degli errori relativi percentuali per gruppi di stime e tipologie comunali - Stime del numero di viaggi – Anno 2013

TIPOLOGIE COMUNALI	Viaggi – lavoro						
	1%	5%	10%	15%	20%	25%	30%
Comuni centro dell'area metropolitana	3.422.895	3.072.536	2.636.231	2.202.724	1.773.938	1.353.918	952.284
Comuni periferia dell'area metropolitana	1.416.661	1.281.217	1.113.012	946.591	782.970	624.013	473.302
Comuni con oltre 50.000 abitanti	1.184.332	1.052.161	888.226	726.536	568.765	418.453	283.320
Comuni da 10.001 a 50.000 abitanti	2.934.200	2.669.132	2.338.422	2.008.719	1.680.596	1.355.147	1.034.708
Comuni da 2.001 a 10.000 abitanti	1.710.133	1.577.380	1.411.834	1.246.886	1.082.799	920.006	759.273
Comuni aventi fino a 2.000 abitanti	147.926	141.925	134.422	126.917	119.410	111.900	104.388
TIPOLOGIE COMUNALI	Viaggi – vacanze						
	1%	5%	10%	15%	20%	25%	30%
Comuni centro dell'area metropolitana	17.404.495	13.822.157	9.385.229	5.094.365	1.647.502	525.809	283.209
Comuni periferia dell'area metropolitana	13.819.030	12.022.795	9.783.444	7.556.029	5.354.731	3.227.344	1.401.708
Comuni con oltre 50.000 abitanti	15.258.870	12.347.419	8.731.729	5.190.147	2.020.747	549.635	259.418
Comuni da 10.001 a 50.000 abitanti	17.442.946	13.472.883	8.565.666	3.908.287	950.464	361.745	215.165
Comuni da 2.001 a 10.000 abitanti	15.904.186	12.670.447	8.660.797	4.765.315	1.545.985	461.753	242.554
Comuni aventi fino a 2.000 abitanti	2.573.554	2.326.914	2.019.382	1.713.118	1.408.914	1.108.359	815.066
TIPOLOGIE COMUNALI	Viaggi – totali						
	1%	5%	10%	15%	20%	25%	30%
Comuni centro dell'area metropolitana	48.703.060	24.302.745	5.892.532	2.354.560	1.424.737	1.016.781	789.487
Comuni periferia dell'area metropolitana	81.647.889	51.568.743	16.271.351	2.425.039	1.060.472	672.204	491.328
Comuni con oltre 50.000 abitanti	20.026.164	14.597.166	8.031.161	2.720.986	886.497	476.056	321.415
Comuni da 10.001 a 50.000 abitanti	40.849.332	20.733.932	4.947.900	1.904.462	1.138.584	808.235	625.701
Comuni da 2.001 a 10.000 abitanti	26.606.372	17.348.371	7.158.027	2.270.588	1.117.207	724.055	533.199
Comuni aventi fino a 2.000 abitanti	4.169.223	3.408.527	2.473.107	1.579.052	813.626	366.103	198.699

7. Esempi di calcolo degli errori campionari

7.1. Esempi di calcolo degli errori campionari per stime di totali riferiti ai viaggi e alle notti

Esempio 1

Come visto nel paragrafo precedente, il prospetto 7 presenta il valore dei coefficienti a, b e c delle funzioni utilizzate per l'interpolazione degli errori con riferimento alle stime del numero di notti, mentre il prospetto 9 presenta gli stessi coefficienti con riferimento alle stime del numero di viaggi.

Sulla base dei dati riportati in tali prospetti, è possibile calcolare l'errore di campionamento assoluto e relativo di una stima a livello trimestrale.

Se ad esempio si volesse calcolare l'errore di campionamento assoluto per una stima appartenente al gruppo del numero di notti per lavoro riferita al Centro per il terzo trimestre, ${}_C\hat{Y}_3$, si possono leggere i valori dei coefficienti a, b e c dal prospetto 7 e si può utilizzare la seguente espressione:

$$\hat{\sigma}({}_C\hat{Y}_3) = -7.761,716 + 0,924586713{}_C\hat{Y}_3 - 0,000000554969({}_C\hat{Y}_3)^2$$

Esempio 2

La valutazione degli errori di campionamento delle stime mediante un metodo approssimato si può ottenere utilizzando i dati dei prospetti 8 e 10. Il prospetto 8 fa riferimento alle stime del numero di notti; il prospetto 10 fa riferimento alle stime del numero di viaggi. Tali prospetti sono articolati per trimestre e area territoriale e presentano gruppi di stime per valori prefissati degli errori di campionamento relativi pari all'1, 5, 10, 15, 20, 25 e 30%.

Considerando ad esempio il prospetto 8, si può osservare che il valore riportato all'incrocio tra la riga riferita alle notti per vacanza del terzo trimestre per l'Italia e la colonna relativa all'errore del 10% indica che tutte le stime superiori a 42.476.160 presentano un errore relativo inferiore al 10%; analogamente si avrà che le stime inferiori a tale valore presenteranno un errore relativo superiore al 10%.

Esempio 3

Per fare un ulteriore esempio e meglio specificare l'utilizzo del suddetto prospetto 8, si consideri la stima del numero di notti per lavoro, riferita al totale Italia nel trimestre luglio-settembre, pari a 3.057.285.

Dal prospetto 8, considerando le notti-lavoro ed esaminando l'ultima riga corrispondente al totale Italia per il trimestre luglio-settembre, si può osservare che il valore di tale stima, pari a 3.057.285 notti, risulta essere compreso tra i valori 3.672.271 e 3.008.732

Da ciò si può facilmente desumere che il valore dell'errore relativo percentuale corrispondente è compreso tra il 15% e il 20%.

7.2. Esempi di calcolo degli errori campionari per stime di frequenze riferite agli individui

Esempio 1

Il totale delle persone che hanno viaggiato nel trimestre luglio-settembre è pari 23.472.267.

Si cerca il livello di stima che più si avvicina a 23.472.267 nella prima colonna del prospetto 5, che è pari a 25.000.000. Dalla colonna del prospetto, riferita all'Italia per il terzo trimestre, si ricava il corrispondente errore relativo percentuale della stima considerata, che è pari a 2,2%.

L'errore assoluto sarà:

$$\sigma(23.472.267) = 2,2/100 \times 23.472.267 = 516.390$$

L'intervallo di confidenza (per $P=0,95$ e $k=2$), avrà come estremi:

$$23.472.267 - (2 \times 516.390) = 22.439.487$$

$$23.472.267 + (2 \times 516.390) = 24.505.047$$

Esempio 2

Considerando la stima precedente, si possono ottenere valori più precisi dell'errore di campionamento mediante interpolazione lineare dei due livelli di stima consecutivi tra i quali è compreso il valore della stima considerato.

Tali livelli sono 20.000.000 e 25.000.000 ai quali corrispondono i valori 2,5 e 2,2 percentuali.

L'errore corrispondente a 23.472.267 è pari a:

$$\sigma(23.472.267) = 2,5 - ((2,5 - 2,2) / (25.000.000 - 20.000.000)) \times (23.472.267 - 20.000.000) = 2,29\%$$

Il corrispondente errore assoluto è $2,29/100 \times 23.472.267 = 537.905$ e l'intervallo di confidenza avrà come estremi:

$$23.472.267 - (2 \times 537.905) = 22.396.457$$

$$23.472.267 + (2 \times 537.905) = 24.548.077$$

Esempio 3

Il calcolo dell'errore può essere effettuato direttamente tramite la funzione interpolante (12):

$$\log \varepsilon^2({}_d\hat{Y}_t) = a + b \log({}_d\hat{Y}_t)$$

Per la stima di 23.472.267 si ha:

$$\log \varepsilon^2 = 10,30754 - 1,05217 \times \log(23.472.267) = -7,54918$$

Da questa il valore di ε è facilmente desumibile, calcolando la radice quadrata dell'antilogaritmo dell'espressione di sopra.

Infatti, da $\log(\varepsilon^2) = -7,54918$ si ha:

$$\varepsilon^2 = \exp(-7,54918)$$

e dunque:

$$\varepsilon = 0,023$$

L'errore relativo percentuale sarà dunque pari a 2,3%.