

***Indagine multiscopo sulle famiglie
Viaggi e vacanze
Anno 2010***

Manuale per l'utente

Per informazioni sull'acquisto rivolgersi a:
Istat – Servizio Sistema informativo
diffusione e servizi all'utenza
via Cesare Balbo, 16 – 00184 Roma
tel: 06.4673.4126
fax: 06.46678200
<https://contact.istat.it/>

Per informazioni sull'indagine rivolgersi a:
Istat - Servizio struttura e dinamica sociale
U.O. Integrazione sociale dei migranti,
mobilità, domanda turistica
via Adolfo Ravà, 150 – 00142 Roma
tel: 06.4673.4655, 4616
fax: 06.4673.4596
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 alla Direttiva 95/57/CE attuata dal Consiglio dell'Unione Europea il 23/11/1995, 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/.

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 della Direttiva 95/57/CE 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, 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**: viaggio svolto per motivi prevalenti di lavoro temporaneo fuori sede o più in generale per motivi professionali (missione, partecipazione a congressi, convegni ecc.). In questi viaggi sono inclusi anche quelli effettuati per formazione/aggiornamento professionale o per sostenere concorsi;
- **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, tipo della vacanza di piacere/svago, 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.

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: 34551
- **un file dei viaggi** (tipo A o tipo B)
lunghezza record: 132 bytes
numero record: 11924

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.

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

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. 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 indifferentemente 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.

2. Caratteristiche della lista di selezione

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,

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.

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.

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 2010

REGIONI	Indirizzi campione	Tasso di sostituzione totale (a)	Tasso di sostituzione al netto dei mancati contatti (b)
Piemonte	266	37,2	20,7
Valle d'Aosta	92	41,3	27,2
Lombardia	330	28,2	19,4
Trentino-Alto Adige	207	27,5	16,4
<i>Bolzano-Bozen</i>	100	33,0	21,0
<i>Trento</i>	107	22,4	12,1
Veneto	202	31,2	25,2
Friuli-Venezia Giulia	139	24,5	14,4
Liguria	152	33,6	24,3
Emilia-Romagna	198	26,8	18,2
Toscana	213	32,9	21,6
Umbria	104	20,2	14,4
Marche	141	21,3	15,6
Lazio	219	28,8	17,4
Abruzzo	143	37,1	26,6
Molise	96	20,8	17,7
Campania	197	34,0	25,9
Puglia	187	25,7	19,3
Basilicata	99	23,2	18,2
Calabria	158	33,5	22,8
Sicilia	212	36,3	25,0
Sardegna	151	31,8	15,9
Italia	3506	30,3	20,4

(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 2010

TIPOLOGIA DEL COMUNE	Indirizzi campione	Tasso di sostituzione totale	Tasso di sostituzione al netto dei mancati contatti
Comuni centro dell'area metropolitana	413	32,9	21,1
Comuni che gravitano intorno al centro dell'area metropolitana	323	28,5	19,2
Comuni aventi fino a 2.000 abitanti	286	30,8	19,6
Comuni con 2.001-10.000 abitanti	920	29,3	20,3
Comuni con 10.001-50.000 abitanti	926	30,9	20,7
Comuni con oltre 50.000 abitanti	638	29,6	20,7
Italia	3506	30,3	20,4

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

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

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

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

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

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

dove

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

in cui 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 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

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

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

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

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

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

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:

$$\{\hat{Y} - k\hat{\sigma}(\hat{Y}), \hat{Y} + k\hat{\sigma}(\hat{Y})\} \quad (11)$$

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:

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

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, \hat{Y}_t , si perviene alla seguente equazione di secondo grado:

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

la cui radice positiva è espressa dalla seguente formula:

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

Nei prospetti 8, 10 e 12 (per le stime relative alle notti e ai viaggi) vengono riportati i valori delle stime \hat{Y}_t^* ottenuti sulla base della relazione (15) in corrispondenza di alcuni valori tipici prefissati dell'errore relativo percentuale, indicati come $\varepsilon^* = \varepsilon(\hat{Y}_t^*)$; le stime con valori superiori a \hat{Y}_t^* presentano valori dell'errore relativo inferiori a ε^* , mentre le stime che assumono valori inferiori a \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 2010

AREE TERRITORIALI	Trimestri 1,2,e 4			Trimestre 3		
	a	b	R^2	a	b	R^2
Nord	10,09278	-1,07958	86,4	10,59087	-1,11448	89,5
Centro	10,39193	-1,11650	90,0	10,49735	-1,12093	89,5
Mezzogiorno	10,54282	-1,14001	91,1	10,70850	-1,14168	91,0
Italia	11,81043	-1,19813	92,7	11,49473	-1,17267	92,0

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 2010

AREE TERRITORIALI	a	b	R^2
Nord	9,95533	-1,15998	92,1
Centro	9,64054	-1,15289	92,0
Mezzogiorno	10,23510	-1,19971	92,9
Italia	10,61562	-1,20051	93,8

Prospetto 5 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite agli individui per aree territoriali - Trimestri – Anno 2010

STIME	Trimestri 1,2 e 4				Trimestre 3			
	Aree territoriali				Aree territoriali			
	Nord	Centro	Mezzogiorno	Italia	Nord	Centro	Mezzogiorno	Italia
50.000	45,2	43,0	40,8	56,2	48,0	44,2	43,9	48,0
60.000	41,0	38,8	36,8	50,4	43,4	39,9	39,6	43,4
70.000	37,7	35,6	33,7	45,9	39,8	36,6	36,3	39,8
80.000	35,1	33,1	31,2	42,4	36,9	34,0	33,6	36,9
90.000	32,9	31,0	29,2	39,5	34,6	31,8	31,4	34,6
100.000	31,1	29,2	27,5	37,1	32,6	30,0	29,6	32,6
200.000	21,4	19,8	18,5	24,5	22,2	20,3	19,9	22,2
300.000	17,2	15,8	14,7	19,2	17,7	16,2	15,8	17,7
400.000	14,7	13,5	12,5	16,2	15,1	13,8	13,4	15,1
500.000	13,0	11,9	11,0	14,1	13,3	12,2	11,8	13,3
750.000	10,5	9,5	8,7	11,1	10,6	9,7	9,4	10,6
1.000.000	9,0	8,1	7,4	9,3	9,0	8,3	7,9	9,0
2.000.000	6,2	5,5	5,0	6,2	6,1	5,6	5,4	6,1
3.000.000	5,0	4,4	4,0	4,8	4,9	4,5	4,2	4,9
4.000.000	4,2	3,7	3,4	4,1	4,2	3,8	3,6	4,2
5.000.000	3,8	3,3	3,0	3,6	3,7	3,3	3,2	3,7
7.500.000	3,0	2,6	2,3	2,8	2,9	2,7	2,5	2,9
10.000.000	2,6	2,2	2,0	2,4	2,5	2,3	2,1	2,5
15.000.000	2,1	1,8	1,6	1,8	2,0	1,8	1,7	2,0
20.000.000	1,8	1,5	1,3	1,6	1,7	1,5	1,4	1,7
25.000.000	1,6	1,3	1,2	1,4	1,5	1,4	1,3	1,5

Prospetto 6 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite agli individui per aree territoriali - Media dei 4 trimestri – Anno 2010

STIME	Aree territoriali			
	Nord	Centro	Mezzogiorno	Italia
20.000	46,5	41,1	43,9	52,9
30.000	36,7	32,6	34,4	41,5
40.000	31,1	27,6	29,0	34,9
50.000	27,3	24,2	25,3	30,5
60.000	24,6	21,8	22,7	27,4
70.000	22,5	20,0	20,7	24,9
80.000	20,8	18,5	19,1	23,0
90.000	19,4	17,3	17,8	21,4
100.000	18,3	16,3	16,7	20,1
200.000	12,2	10,9	11,0	13,3
300.000	9,7	8,6	8,7	10,4
400.000	8,2	7,3	7,3	8,8
500.000	7,2	6,4	6,4	7,7
750.000	5,7	5,1	5,0	6,0
1.000.000	4,8	4,3	4,2	5,1
2.000.000	3,2	2,9	2,8	3,3
3.000.000	2,5	2,3	2,2	2,6
4.000.000	2,2	1,9	1,8	2,2
5.000.000	1,9	1,7	1,6	1,9
7.500.000	1,5	1,3	1,3	1,5
10.000.000	1,3	1,1	1,1	1,3
15.000.000	1,0	0,9	0,8	1,0
20.000.000	0,8	0,8	0,7	0,8

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 2010

Notti – lavoro				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	36.646,784	0,481096	-0,000000061875	86,68
Centro	4.497,016	0,628989	-0,000000088307	81,91
Mezzogiorno	12.594,622	0,567607	-0,000000077594	95,84
Italia	92.051,459	0,320742	-0,000000015836	88,08
Luglio-settembre				
Nord	64.279,587	0,443681	-0,000000050472	87,54
Centro	13.735,525	0,571410	-0,000000107727	98,41
Mezzogiorno	25.932,045	0,507730	-0,000000103887	95,30
Italia	104.458,159	0,280005	-0,000000014880	92,47
Anno				
Nord	156.183,216	0,237307	-0,000000006988	82,55
Centro	82.400,794	0,299285	-0,000000015138	86,84
Mezzogiorno	117.425,209	0,352201	-0,000000014814	76,23
Italia	292.868,317	0,164350	-0,000000002204	81,23
Notti – vacanze				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	174.769,729	0,162222	-0,000000002096	88,27
Centro	116.682,066	0,198366	-0,000000006631	90,10
Mezzogiorno	97.264,036	0,188281	-0,000000004925	89,02
Italia	267.523,110	0,107185	-0,000000000745	90,12
Luglio-settembre				
Nord	721.459,393	0,085872	-0,000000000219	87,62
Centro	446.857,315	0,121141	-0,000000000633	88,33
Mezzogiorno	279.338,843	0,128653	-0,000000000964	89,88
Italia	909.611,465	0,062175	-0,000000000088	88,88
Anno				
Nord	797.602,484	0,074537	-0,000000000120	91,31
Centro	479.843,174	0,104723	-0,000000000364	90,79
Mezzogiorno	341.857,744	0,103691	-0,000000000444	92,88
Italia	1.059.717,922	0,052603	-0,000000000046	92,27
Notti - totali				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	200.802,314	0,151443	-0,000000001721	89,60
Centro	138.739,643	0,206771	-0,000000006342	88,59
Mezzogiorno	110.880,669	0,191626	-0,000000004468	91,89
Italia	315.544,086	0,101715	-0,000000000614	91,51
Luglio-settembre				
Nord	729.504,591	0,083071	-0,000000000195	88,86
Centro	454.789,514	0,120039	-0,000000000601	89,98
Mezzogiorno	295.048,524	0,123213	-0,000000000834	91,12
Italia	917.297,592	0,060438	-0,000000000079	90,17
Anno				
Nord	822.380,434	0,072692	-0,000000000112	91,99
Centro	514.144,175	0,101042	-0,000000000327	91,43
Mezzogiorno	432.305,441	0,099993	-0,000000000389	92,23
Italia	1.136.758,716	0,050847	-0,000000000042	92,80

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 2010

Notti – lavoro							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	7.690.658	7.051.179	6.253.811	5.459.511	4.669.779	3.887.233	3.116.816
Centro	7.016.741	6.564.278	5.998.806	5.433.485	4.868.369	4.303.538	3.739.121
Mezzogiorno	7.208.719	6.694.944	6.053.136	5.411.934	4.771.580	4.132.463	3.495.245
Italia	19.914.955	17.430.593	14.344.875	11.296.759	8.323.144	5.520.296	3.153.286
Luglio-settembre							
Nord	8.738.191	7.959.929	6.991.453	6.029.862	5.078.774	4.144.651	3.239.828
Centro	5.235.777	4.866.316	4.404.923	3.944.167	3.484.296	3.025.705	2.569.058
Mezzogiorno	4.842.629	4.461.992	3.987.358	3.514.488	3.044.168	2.577.713	2.117.468
Italia	18.524.999	15.899.331	12.652.332	9.477.861	6.463.059	3.843.201	2.061.548
Anno							
Nord	33.202.834	27.614.708	20.728.177	14.081.692	8.098.814	3.905.934	2.031.334
Centro	19.390.046	16.791.213	13.565.446	10.385.456	7.303.746	4.472.631	2.309.570
Mezzogiorno	23.437.303	20.780.478	17.477.483	14.206.812	10.994.729	7.901.831	5.083.027
Italia	71.875.770	54.325.410	33.197.543	15.233.142	5.993.901	3.162.067	2.088.155
Notti – vacanze							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	73.765.323	55.062.773	32.274.003	12.502.213	3.817.699	1.904.449	1.244.914
Centro	29.013.867	23.135.590	15.938.565	9.205.589	4.073.453	1.829.813	1.072.953
Mezzogiorno	36.733.938	28.761.734	18.964.897	9.789.360	3.410.619	1.415.916	839.536
Italia	133.176.860	81.199.530	24.379.299	5.685.928	2.818.574	1.855.263	1.380.102
Luglio-settembre							
Nord	355.356.768	181.740.385	33.574.263	10.847.975	6.246.530	4.370.196	3.357.744
Centro	179.558.309	118.382.373	48.090.624	12.213.305	5.429.943	3.410.677	2.476.678
Mezzogiorno	125.453.972	85.038.498	37.473.544	9.235.568	3.727.559	2.261.375	1.615.573
Italia	609.084.394	191.956.550	22.833.594	10.251.696	6.572.162	4.831.925	3.819.311
Anno							
Nord	547.958.467	232.253.492	27.695.846	10.396.920	6.318.938	4.531.604	3.530.959
Centro	265.213.538	158.655.504	43.371.545	9.822.315	4.942.946	3.276.059	2.446.088
Mezzogiorno	214.493.249	126.925.553	32.204.524	6.922.428	3.493.308	2.320.202	1.734.619
Italia	944.239.505	181.967.582	21.890.055	10.824.657	7.173.370	5.361.709	4.280.038
Notti – totali							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	83.567.777	60.850.699	33.380.576	11.227.938	3.660.393	1.969.657	1.331.150
Centro	31.715.685	25.574.303	18.047.288	10.949.299	5.241.327	2.379.046	1.361.970
Mezzogiorno	41.250.679	32.461.294	21.652.570	11.478.144	4.131.823	1.682.740	983.267
Italia	152.724.335	89.932.809	24.108.126	6.066.939	3.148.565	2.109.529	1.583.600
Luglio-settembre							
Nord	384.206.995	189.239.201	31.589.818	10.573.727	6.175.234	4.348.051	3.352.762
Centro	187.018.797	122.629.078	48.813.414	12.194.466	5.463.159	3.444.523	2.506.165
Mezzogiorno	138.328.054	91.660.435	37.319.070	8.672.951	3.694.209	2.292.546	1.656.010
Italia	654.435.468	192.065.601	22.199.647	10.150.920	6.548.356	4.829.295	3.824.228
Anno							
Nord	572.962.610	234.131.565	27.104.202	10.478.763	6.423.502	4.624.649	3.611.492
Centro	283.656.656	165.420.053	41.254.342	9.852.617	5.109.214	3.425.807	2.573.286
Mezzogiorno	236.285.556	136.779.081	33.344.087	8.131.084	4.252.474	2.860.697	2.152.447
Italia	1.003.135.267	175.248.116	22.688.869	11.409.801	7.605.223	5.701.150	4.559.010

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 2010

Viaggi – lavoro				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	9.559,605	0,267211	-0,000000108270	94,65
Centro	8.661,985	0,375142	-0,000000275329	90,44
Mezzogiorno	5.310,523	0,488970	-0,000000250052	87,68
Italia	15.661,415	0,227310	-0,000000043159	86,87
Luglio-settembre				
Nord	14.739,900	0,285405	-0,000000133266	91,01
Centro	9.813,555	0,423532	-0,000000219488	97,89
Mezzogiorno	5.573,717	0,384909	-0,000000383827	96,39
Italia	16.944,492	0,220309	-0,000000045003	95,00
Anno				
Nord	22.412,154	0,172089	-0,000000018484	93,37
Centro	19.230,713	0,231198	-0,000000032785	95,56
Mezzogiorno	18.390,088	0,268800	-0,000000044477	83,62
Italia	38.185,014	0,128827	-0,000000006107	91,76
Viaggi – vacanza				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	35.313,021	0,168952	-0,000000010687	92,68
Centro	20.376,740	0,243962	-0,000000028995	96,76
Mezzogiorno	19.010,828	0,221323	-0,000000028832	90,91
Italia	51.651,344	0,116563	-0,000000003652	93,33
Luglio-settembre				
Nord	56.569,219	0,108925	-0,000000003180	89,73
Centro	32.719,527	0,158805	-0,000000010976	90,82
Mezzogiorno	29.725,280	0,156536	-0,000000012231	91,65
Italia	73.199,872	0,078526	-0,000000001281	89,96
Anno				
Nord	79.798,166	0,074194	-0,000000001110	89,33
Centro	52.371,414	0,102032	-0,000000003536	89,22
Mezzogiorno	44.294,124	0,106983	-0,000000004328	90,41
Italia	112.893,799	0,052191	-0,000000000433	89,35
Viaggi – totali				
AREE TERRITORIALI	a	b	c	R ²
Gennaio-marzo				
Nord	25.374,099	0,110392	-0,000000005805	86,19
Centro	17.435,557	0,168502	-0,000000018847	86,78
Mezzogiorno	16.676,219	0,145603	-0,000000016217	94,97
Italia	34.213,492	0,079820	-0,000000002195	89,26
Luglio-settembre				
Nord	49.795,610	0,108957	-0,000000003207	88,89
Centro	33.560,568	0,156280	-0,000000010310	90,53
Mezzogiorno	25.170,489	0,158062	-0,000000012086	91,89
Italia	62.163,188	0,079285	-0,000000001288	89,53
Anno				
Nord	50.971,375	0,089033	-0,000000001173	87,23
Centro	37.831,095	0,123176	-0,000000003510	87,46
Mezzogiorno	34.855,332	0,129638	-0,000000003923	86,42
Italia	69.779,226	0,063767	-0,000000000446	87,24

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 2010

Viaggi – lavoro							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	2.412.249	2.049.284	1.599.587	1.158.776	740.074	387.070	182.077
Centro	1.349.512	1.206.984	1.029.866	854.533	682.231	515.540	360.246
Mezzogiorno	1.926.504	1.767.529	1.569.091	1.371.087	1.173.734	977.410	782.852
Italia	5.106.127	4.194.768	3.068.038	1.974.997	996.802	394.380	193.275
Luglio-settembre							
Nord	2.118.777	1.826.965	1.466.649	1.115.225	782.252	490.956	282.292
Centro	1.907.514	1.727.711	1.503.762	1.281.126	1.060.581	843.619	633.406
Mezzogiorno	991.415	888.891	761.360	634.892	510.214	388.832	274.181
Italia	4.752.487	3.881.428	2.807.488	1.774.511	879.425	366.784	191.843
Anno							
Nord	8.905.251	6.783.812	4.189.471	1.850.324	580.117	270.327	170.992
Centro	6.832.783	5.631.031	4.143.341	2.694.382	1.377.433	531.050	249.778
Mezzogiorno	5.888.934	5.002.045	3.901.197	2.817.775	1.779.251	888.206	381.715
Italia	19.773.065	13.374.697	5.798.458	1.309.134	513.853	310.276	221.331
Viaggi – vacanze							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	15.092.649	11.420.108	6.928.967	2.909.217	874.272	413.192	263.791
Centro	8.155.182	6.792.924	5.102.767	3.444.638	1.888.354	740.637	312.951
Mezzogiorno	7.418.230	6.050.986	4.359.120	2.716.431	1.262.009	454.885	223.350
Italia	29.659.917	18.974.267	6.659.905	1.346.697	603.129	383.070	280.015
Luglio-settembre							
Nord	31.669.042	19.443.985	5.848.109	1.255.219	608.208	397.425	294.612
Centro	13.773.728	10.205.104	5.865.811	2.173.653	673.431	344.503	227.703
Mezzogiorno	12.180.713	8.981.293	5.099.188	1.848.932	586.964	305.804	203.661
Italia	54.531.476	24.587.788	2.905.182	1.006.004	598.815	425.533	329.883
Anno							
Nord	59.038.292	24.701.558	2.763.606	1.036.910	630.782	452.604	352.780
Centro	26.583.291	15.660.067	4.146.402	1.015.734	524.640	350.992	263.306
Mezzogiorno	22.854.049	13.901.142	4.105.797	940.662	466.087	306.863	228.314
Italia	100.145.040	18.884.953	2.312.938	1.148.392	762.081	570.010	455.206
Viaggi – totali							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Gennaio-marzo							
Nord	17.543.979	10.808.336	3.169.446	589.671	278.156	180.399	133.280
Centro	8.518.676	6.431.524	3.873.521	1.570.689	438.493	204.287	130.163
Mezzogiorno	8.483.243	6.064.954	3.139.655	887.519	282.733	155.960	106.810
Italia	32.286.502	14.647.303	1.462.665	480.293	283.220	200.524	155.149
Luglio-settembre							
Nord	31.352.939	19.193.525	5.577.200	1.115.950	536.796	350.263	259.521
Centro	14.413.937	10.615.040	6.001.157	2.134.268	663.733	344.999	229.727
Mezzogiorno	12.418.342	9.168.199	5.204.227	1.814.692	521.737	264.575	174.735
Italia	54.656.644	24.683.443	2.585.184	865.420	512.158	363.139	281.183
Anno							
Nord	68.024.680	34.538.815	3.406.573	823.014	457.128	315.929	241.284
Centro	32.572.317	21.351.070	7.956.845	1.216.644	481.831	295.873	213.047
Mezzogiorno	30.787.508	20.730.487	8.589.980	1.357.034	482.400	286.906	203.641
Italia	121.882.552	35.310.888	1.882.232	805.833	511.348	374.351	295.218

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 2010

TIPOLOGIE COMUNALI	Viaggi – lavoro			R^2
	a	b	c	
Comuni centro dell'area metropolitana	21.594,394	0,282199	-0,000000044558	93,72
Comuni periferia dell'area metropolitana	12.186,936	0,334044	-0,000000139569	94,11
Comuni con oltre 50.000 abitanti	17.159,233	0,289417	-0,000000071554	81,40
Comuni da 10.001 a 50.000 abitanti	16.220,528	0,259408	-0,000000062181	88,53
Comuni da 2.001 a 10.000 abitanti	14.091,250	0,248208	-0,000000070242	93,06
Comuni aventi fino a 2.000 abitanti	5.784,892	0,568139	-0,000000616864	92,44
TIPOLOGIE COMUNALI	Viaggi – vacanze			R^2
	a	b	c	
Comuni centro dell'area metropolitana	59.050,768	0,112990	-0,000000004244	91,57
Comuni periferia dell'area metropolitana	42.141,296	0,157522	-0,000000008920	90,37
Comuni con oltre 50.000 abitanti	42.559,574	0,111014	-0,000000005041	91,47
Comuni da 10.001 a 50.000 abitanti	50.742,771	0,099622	-0,000000004017	85,99
Comuni da 2.001 a 10.000 abitanti	56.016,477	0,115616	-0,000000004728	83,37
Comuni aventi fino a 2.000 abitanti	12.424,959	0,308690	-0,000000056854	84,75
TIPOLOGIE COMUNALI	Viaggi – totali			R^2
	a	b	c	
Comuni centro dell'area metropolitana	45.991,606	0,130046	-0,000000003940	87,23
Comuni periferia dell'area metropolitana	35.449,005	0,177903	-0,000000008710	83,11
Comuni con oltre 50.000 abitanti	31.251,653	0,135029	-0,000000005071	88,58
Comuni da 10.001 a 50.000 abitanti	33.831,846	0,123713	-0,000000004032	85,60
Comuni da 2.001 a 10.000 abitanti	33.559,063	0,143030	-0,000000005010	85,49
Comuni aventi fino a 2.000 abitanti	9.798,894	0,372714	-0,000000056666	85,21

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 2010

TIPOLOGIE COMUNALI	Viaggi – lavoro					
	1%	5%	10%	15%	20%	30%
Comuni centro dell'area metropolitana	6.187.145	5.302.514	4.204.266	3.122.098	2.077.972	1.145.646
Comuni periferia dell'area metropolitana	2.358.760	2.077.182	1.727.448	1.381.845	1.044.044	722.946
Comuni con oltre 50.000 abitanti	3.965.440	3.416.148	2.734.866	2.064.564	1.418.677	837.282
Comuni da 10.001 a 50.000 abitanti	4.075.006	3.443.465	2.661.617	1.897.018	1.177.031	591.969
Comuni da 2.001 a 10.000 abitanti	3.449.392	2.891.163	2.201.094	1.529.308	907.392	435.319
Comuni aventi fino a 2.000 abitanti	915.049	850.977	771.064	691.410	612.112	533.320
TIPOLOGIE COMUNALI	Viaggi – vacanze					
	1%	5%	10%	15%	20%	30%
Comuni centro dell'area metropolitana	24.825.230	15.725.348	5.561.835	1.377.809	657.570	425.389
Comuni periferia dell'area metropolitana	16.818.447	12.433.414	7.112.535	2.635.629	842.877	437.247
Comuni con oltre 50.000 abitanti	20.452.153	12.765.450	4.196.754	970.003	465.972	302.887
Comuni da 10.001 a 50.000 abitanti	22.862.350	13.302.167	3.507.394	937.208	495.686	334.448
Comuni da 2.001 a 10.000 abitanti	22.857.968	14.685.728	5.469.387	1.370.781	640.821	410.899
Comuni aventi fino a 2.000 abitanti	5.294.874	4.597.581	3.729.208	2.867.380	2.019.916	1.212.519
TIPOLOGIE COMUNALI	Viaggi – totali					
	1%	5%	10%	15%	20%	30%
Comuni centro dell'area metropolitana	30.846.088	20.874.839	8.932.431	1.720.422	634.760	378.699
Comuni periferia dell'area metropolitana	19.486.374	14.957.094	9.378.303	4.177.850	1.114.588	465.508
Comuni con oltre 50.000 abitanti	24.901.656	17.126.524	7.706.951	1.412.080	464.194	268.640
Comuni da 10.001 a 50.000 abitanti	28.494.626	18.728.408	7.067.806	1.101.057	433.547	265.644
Comuni da 2.001 a 10.000 abitanti	26.802.760	18.922.828	9.308.443	1.984.385	561.356	309.246
Comuni aventi fino a 2.000 abitanti	6.427.837	5.725.244	4.848.340	3.973.822	3.103.655	2.242.677

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:

$$\hat{\sigma}({}_C\hat{Y}_3) = 13.735,525 + 0,57141{}_C\hat{Y}_3 - 0,000000107727({}_C\hat{Y}_3)^2.$$

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 22.833.594 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 9.651.000.

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 9.651.000 risulta essere compreso tra i valori 12.652.332 e 9.477.861

Da ciò si può facilmente desumere che il valore dell'errore relativo percentuale corrispondente è compreso tra il 10 ed il 15%.

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 gennaio-marzo è pari a 13.093.000.

Si cerca il livello di stima che più si avvicina a 13.093.000 nella prima colonna del prospetto 5, che è pari a 15.000.000. Dalla colonna del prospetto, riferita all'Italia per il primo trimestre, si ricava il corrispondente errore relativo percentuale della stima considerata, che è pari a 1,8%.

L'errore assoluto sarà:

$$\sigma(13.093.000) = 1,8/100 \times 13.093.000 = 235.674$$

L'intervallo di confidenza avrà come estremi:

$$13.093.000 - (2 \times 235.674) = 12.621.652$$

$$13.093.000 + (2 \times 235.674) = 13.564.348$$

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 10.000.000 e 15.000.000 ai quali corrispondono i valori 2,4 e 1,8 percentuali.

L'errore corrispondente a 13.093.000 è pari a:

$$\sigma(13.093.000) = 2,4 - ((2,4 - 1,8) / (15.000.000 - 10.000.000)) \times (13.093.000 - 10.000.000) = 2,36\%$$

Il corrispondente errore assoluto è $2,36/100 \times 13.093.000 = 308.994$ e l'intervallo di confidenza avrà come estremi:

$$13.093.000 - (2 \times 308.994) = 12.475.010$$

$$13.093.000 + (2 \times 308.994) = 13.710.990$$

Esempio 3

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

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

Per la stima di 13.093.000 si ha:

$$\log \varepsilon^2 = 10,09278 - 1,07958 \times \log(13.093.000) = 2,40936$$

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

Infatti, da $\log(\varepsilon^2) = 2,40936$ si ha:

$$\varepsilon^2 = \exp(2,40936)$$

e dunque:

$$\varepsilon = 3,335$$

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