

Nota Metodologica

Cenni sull'indagine Viaggi e vacanze

L'indagine Viaggi e vacanze sulla domanda turistica è condotta dall'Istat dal 1997 con cadenza trimestrale e fa parte del Sistema delle Indagini Multiscopo sulle famiglie.

L'indagine è di tipo campionario (cfr. Strategia di campionamento e valutazione degli errori campionari). 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 (cittadini italiani e stranieri residenti in famiglia, di qualunque età). Le informazioni sono raccolte tramite intervista telefonica assistita da computer (CATI).

L'indagine si presenta come uno strumento idoneo ad analizzare le diverse tipologie di domanda turistica, i comportamenti turistici e le modalità di effettuazione dei viaggi (con pernottamento) attraverso la raccolta di informazioni sulla destinazione, la forma di organizzazione scelta, il tipo di alloggio utilizzato, il mezzo di trasporto, la durata del soggiorno, il periodo dell'anno in cui si è viaggiato, ecc. Inoltre, è la fonte privilegiata per conoscere la domanda turistica 'sommersa', vale a dire quella che si realizza negli alloggi privati a pagamento, nelle seconde case di proprietà oppure tramite l'ospitalità di parenti o amici.

Gli spostamenti turistici sono classificati, secondo gli standard internazionali, distinguendo i viaggi per motivi di lavoro da quelli per motivi di vacanza e le vacanze 'brevi' da quelle 'lunghe'. Tra le vacanze rientrano i viaggi per svago, piacere, relax, per visitare parenti o amici, per trattamenti di salute o per motivi religiosi.

Oltre ai viaggi, costituiscono oggetto di analisi anche gli individui che hanno viaggiato e le notti trascorse in viaggio.

Avvertenze

Per una corretta interpretazione dei dati riportati nelle tavole si suggerisce la lettura della sezione "Definizioni e classificazioni" con particolare riferimento alla definizione di viaggio adottata e che riguarda, più in particolare, gli spostamenti turistici con pernottamento effettuati fuori dall' "ambiente abituale", vale a dire fuori dal luogo in cui si vive, nonché dai luoghi frequentati settimanalmente con soste di almeno un pernottamento. Inoltre, poiché la stagionalità del turismo e le caratteristiche che contraddistinguono le diverse tipologie di viaggio fanno sì che alcuni aspetti del viaggiare siano, in termini quantitativi, particolarmente rari in particolari periodi dell'anno, è opportuno porre particolare cura ed attenzione nell'analisi e valutazione delle stime. A tal fine è bene sempre far riferimento agli errori campionari (sezione "Strategia di campionamento e valutazione degli errori campionari").

Si rammenta che l'indagine rileva i flussi turistici che si concludono in ciascun trimestre di rilevazione anche se iniziati precedentemente. Per quanto concerne i prospetti e le tavole si precisa, pertanto, che i dati trimestrali fanno riferimento ai viaggi conclusi nel trimestre, indipendentemente dalla data di inizio.

I dati sui viaggi, nonché sulle notti, si riferiscono a stime trimestrali e a stime annuali; queste ultime sono ottenute come somma dei valori trimestrali.

I dati sulle persone che viaggiano, diversamente dai viaggi e dalle notti, si riferiscono soltanto a stime trimestrali o, in alcuni casi, a valori medi trimestrali ottenuti sulla base delle stime dei quattro trimestri.

Si avverte che i dati trimestrali sul numero di persone che viaggiano non sono cumulabili, in quanto una persona può viaggiare in trimestri diversi: l'eventuale somma dei dati trimestrali comporterebbe pertanto una sovrastima del numero di persone che hanno effettuato viaggi nell'anno. Per analogo motivo, si segnala che anche i totali sulle persone che hanno viaggiato per tipo di viaggio o per destinazione non corrispondono alla somma dei parziali, per la presenza di persone che risultano viaggiare contemporaneamente per diverse

tipologie di viaggio o per diverse destinazioni. Ad esempio, un individuo che nel corso del trimestre luglio-settembre ha effettuato sia vacanze brevi che vacanze lunghe viene conteggiato tanto tra i turisti per vacanza breve quanto tra i turisti per vacanza lunga, ma una sola volta come turista per vacanza. Analogamente, un individuo che ha realizzato sia vacanze brevi in Italia che vacanze brevi all'estero risulta presente sia tra i turisti per vacanza breve in Italia sia tra quelli per vacanza breve all'estero, pur essendo conteggiato una sola volta nel totale dei turisti per vacanza breve.

Si fa notare che i prospetti riguardano sia i confronti temporali che i dati relativi all'anno di riferimento dell'indagine, mentre le tavole di approfondimento soltanto i dati relativi all'anno di riferimento.

Si fa presente che i dati assoluti sono espressi in migliaia; inoltre, i totali possono non quadrare per effetto di arrotondamento. Le percentuali e i quozienti, essendo calcolati sui dati assoluti prima dell'arrotondamento, possono non coincidere con i risultati che si ottengono elaborando i dati espressi in migliaia.

Si segnala che il segno (-) indica che il fenomeno esiste e viene rilevato, ma i casi non si sono verificati nel campione; (..) indica che la numerosità non raggiunge la metà della cifra dell'ordine minimo considerato oppure che l'esiguità del fenomeno rende i valori calcolati non significativi.

Si ricorda che i risultati riguardano la popolazione residente al netto dei membri permanenti delle convivenze (ospedali, carceri, comunità religiose, eccetera), poiché trattasi di un'indagine campionaria effettuata presso le famiglie, le quali rappresentano le unità di rilevazione.

Per confronti con i dati delle precedenti indagini "Viaggi e vacanze", si deve tener conto di alcune modifiche introdotte nel corso dei primi anni di indagine al fine di migliorare la qualità delle stime. Ci si riferisce, in particolare, alle modifiche apportate alla scheda riepilogativa dei viaggi nel questionario del 1998 che hanno contribuito in maniera sostanziale a migliorare le stime sui flussi turistici, soprattutto per quanto riguarda le vacanze brevi ed i viaggi di lavoro.¹

Per i confronti storici con i dati anteriori al 1997, ricavati dalle precedenti indagini Istat sulle vacanze dei residenti in Italia e dalle indagini "Aspetti della vita quotidiana", si deve tener conto, oltre che degli usuali accorgimenti nel paragonare fonti diverse, di problematiche che emergono nel confrontare dati ricavati da indagini che hanno periodi di riferimento diversi, dal momento che in passato il periodo di riferimento era l'anno solare. Si può ipotizzare che più ampio è il periodo a cui ci si riferisce e maggiore è la difficoltà per l'intervistato a ricordare gli eventi accaduti in quel lasso di tempo. Si osservano, infatti, differenze tra i dati dell'indagine trimestrale e quelli relativi alle precedenti indagini sulle vacanze, basate su periodi di riferimento annuali. Analizzando le distribuzioni mensili dei viaggi risulta che, nel caso delle stime ottenute con periodi di riferimento annuali, vi è una maggiore concentrazione delle vacanze nei mesi ad esse abitualmente dedicati (luglio e agosto) rispetto alle informazioni rilevate trimestralmente. Quando il periodo di riferimento dell'indagine è annuale si può verificare, infatti, che l'intervistato ricordi principalmente i viaggi effettuati nel periodo dell'anno in cui è solito trascorrere le proprie vacanze, dimenticando più facilmente quelli effettuati negli altri periodi ed in particolare quelli più remoti rispetto al momento della rilevazione.

¹ Per approfondimenti si suggerisce la lettura dell'appendice B del volume: Istat. *I viaggi in Italia e all'estero nel 1998* Roma: Istat, 1999. (Informazioni n. 25); oppure del volume: Istat. *Metodologia e organizzazione dell'indagine multiscopo sulla domanda turistica 'Viaggi e vacanze'*. Roma: Istat, 2003. (Metodi e Norme n. 17).

Principali definizioni e classificazioni

L'indagine "Viaggi e vacanze" utilizza le definizioni standard indicate nella metodologia adottata per l'attuazione della Direttiva 95/57/CE sulle Statistiche del Turismo.²

I flussi turistici vengono individuati in quegli spostamenti con pernottamento effettuati fuori dall' "ambiente abituale" ovvero fuori dal luogo in cui si vive, nonché dai luoghi frequentati settimanalmente con soste di almeno un pernottamento.

In tal senso, per una corretta interpretazione dei risultati, assume notevole importanza il significato delle definizioni adottate. In particolare:

- **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; vengono esclusi i viaggi e gli spostamenti effettuati nelle località frequentate 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**

persona che ha effettuato uno o più viaggi nel trimestre di riferimento dell'indagine.

I viaggi sono, inoltre, distinti secondo le seguenti tipologie:

- **viaggio per motivi di lavoro o professionali**

viaggio svolto per motivi prevalenti di lavoro temporaneo fuori sede o più in generale per motivi professionali (missioni, partecipazione a congressi, convegni, eccetera): in questi viaggi vengono 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;

nella presentazione dei risultati, il soggiorno di vacanza è stato distinto, in relazione alla durata, in:

- *vacanza breve* : quando la durata del soggiorno è inferiore a 4 pernottamenti;

- *vacanza lunga* : quando la durata del soggiorno è di 4 o più notti.

Sono state adottate le seguenti definizioni:

- **destinazione, mezzo di trasporto, tipo di alloggio, motivo della vacanza, tipo della vacanza di piacere/svago, motivo del viaggio di lavoro**

sono informazioni che vengono rilevate sulla base del concetto di "prevalenza": in particolare, la destinazione del viaggio ed il tipo di alloggio vengono associate rispettivamente alla località ed al tipo di alloggio in cui si è trascorso il maggior numero di notti, mentre il mezzo di trasporto viene individuato nel mezzo con cui è stata coperta la maggiore distanza;

- **organizzazione**

- per *prenotazione diretta* si intende la prenotazione dell'alloggio e/o del trasporto direttamente presso la struttura ricettiva o la compagnia di viaggio;

- per *prenotazione presso agenzia* si intende la prenotazione dell'alloggio e/o del trasporto tramite agenzia di viaggi o tour operator;

- La prenotazione tramite internet riguarda l'utilizzo di internet per prenotare direttamente o presso agenzia l'alloggio e/o il trasporto;

² Eurostat. Community methodology on Tourism statistics. Bruxelles: Eurostat, 1998.

- **estero**
 - *Europa* include i paesi dell'Unione europea e gli altri paesi europei;
 - *Unione europea* comprende: Austria, Belgio, Danimarca, Finlandia, Francia, Germania, Grecia, Irlanda, Lussemburgo, Olanda, Portogallo, Regno Unito, Spagna, Svezia, Cipro, Estonia, Latvia (Lettonia), Lituania, Malta, Polonia, Repubblica Ceca, Slovenia, Slovacchia, Ungheria, Bulgaria, Romania;
 - *altri paesi europei* comprende: Norvegia, Russia, Svizzera, Liechtenstein, Turchia e altri paesi europei altrove non classificati;
 - *paesi extra-europei* include tutti i paesi non menzionati tra quelli dell'Europa;
- **ripartizione geografica**
 le ripartizioni geografiche sono suddivisioni del territorio italiano derivate dalle regioni amministrative e così definite:
 - *Nord*: Piemonte, Valle d'Aosta/Vallée d'Aoste, 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;
- **tipo di comune**
 i comuni italiani sono suddivisi nelle seguenti tipologie:
 - *comuni centro dell'area metropolitana*: Torino, Milano, Venezia, Genova, Bologna, Firenze, Roma, Napoli, Bari, Palermo, Catania, Cagliari;
 - *comuni periferia dell'area metropolitana*: comuni che gravitano intorno al centro dell'area metropolitana, come definiti sulla base dei risultati del Censimento Popolazione;
 - *altri comuni* suddivisi in base alla dimensione demografica (oltre 50.000 abitanti, da 10.001 a 50.000 abitanti, da 2.001 a 10.000 abitanti, fino a 2.000 abitanti).

I dati relativi alle caratteristiche delle persone fanno riferimento al momento dell'intervista. In particolare:

- ***l'età*** è espressa in anni compiuti;
- ***la condizione*** è quella dichiarata come unica o prevalente dalle persone di 15 anni e più:
 - *occupato*: chi possiede un'occupazione in proprio o alle dipendenze da cui trae un profitto o una retribuzione (utile, onorario, stipendio, salario) o chi collabora con un familiare che svolge un'attività lavorativa in conto proprio senza avere un regolare contratto di lavoro (coadiuvante);
 - *persona in cerca di prima occupazione*: chi non ha mai esercitato un'attività lavorativa ed è alla ricerca attiva di un'occupazione che è in grado di accettare se gli viene offerta;
 - *persona in cerca di nuova occupazione*: chi ha perduto una precedente occupazione alle dipendenze ed è alla ricerca attiva di un'occupazione che è in grado di accettare se gli viene offerta;
 - *casalinga*: chi si dedica prevalentemente alle faccende domestiche;
 - *studente*: chi si dedica prevalentemente allo studio;
 - *ritirato dal lavoro*: chi ha cessato un'attività lavorativa per raggiunti limiti di età, invalidità o altra causa; la figura del ritirato dal lavoro non coincide necessariamente con quella del pensionato in quanto, non sempre il ritirato dal lavoro gode di una pensione e non sempre chi riceve una pensione la riceve per motivi di lavoro;
 - *in altra condizione*: chi si trova in condizione diversa da quelle sopra elencate (militare di leva, inabile al lavoro, benestante, detenuto eccetera).

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 vincolata³, 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

³ 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.

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.

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 2009

REGIONI	Indirizzi campione	Tasso di sostituzione totale (a)	Tasso di sostituzione al netto dei mancati contatti (b)
Piemonte	268	33,2	25,7
Valle d'Aosta	93	43,0	26,9
Lombardia	329	22,8	16,7
Trentino-Alto Adige	204	31,9	20,1
<i>Bolzano-Bozen</i>	99	37,4	24,2
<i>Trento</i>	105	26,7	16,2
Veneto	201	23,4	18,4
Friuli-Venezia Giulia	139	32,4	23,0
Liguria	152	35,5	26,3
Emilia-Romagna	197	26,9	18,8
Toscana	213	28,2	20,2
Umbria	103	25,2	15,5
Marche	141	30,5	19,9
Lazio	220	26,4	16,4
Abruzzo	144	25,0	17,4
Molise	95	40,0	28,4
Campania	196	34,7	27,6
Puglia	187	33,2	26,2
Basilicata	99	25,3	21,2
Calabria	158	38,0	29,1
Sicilia	213	38,0	28,6
Sardegna	149	37,6	24,2
Italia	3501	30,9	22,2

(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 (secondo 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 2009

TIPOLOGIA DEL COMUNE	Indirizzi campione	Tasso di sostituzione totale	Tasso di sostituzione al netto dei mancati contatti
Comuni centro dell'area metropolitana	419	31,0	21,5
Comuni che gravitano intorno al centro dell'area metropolitana	321	24,9	19,0
Comuni aventi fino a 2.000 abitanti	634	31,4	21,9
Comuni con 2.001-10.000 abitanti	910	32,1	24,5
Comuni con 10.001-50.000 abitanti	922	29,5	21,4
Comuni con oltre 50.000 abitanti	295	36,6	23,1
Italia	3501	30,9	22,2

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:

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

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 noti⁴ sono:
 - popolazione residente per le 6 tipologie comunali;
 - popolazione residente per sesso nelle 20 regioni italiane;
 - popolazione residente per 8 classi di età⁵;
 - 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 ampiezza⁶

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:

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

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:

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

⁵ 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.

⁶ Le classi di ampiezza considerate sono: 2 componenti e 3 componenti e più.

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

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

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:

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

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 $\hat{\text{Var}}({}_d\hat{Y}_t)$ 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:

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

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

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

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

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

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

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

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

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

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

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

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 2009

AREE TERRITORIALI	Trimestri 1,2,e 4			Trimestre 3		
	a	b	R^2	a	b	R^2
Nord	11,70631	-1,20591	92,7	10,95603	-1,14169	91,9
Centro	10,58817	-1,14800	89,4	9,86751	-1,08285	89,7
Mezzogiorno	11,15267	-1,18681	90,4	10,18371	-1,10538	89,3
Italia	11,77505	-1,20427	93,0	11,13746	-1,15215	92,8

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 2009

AREE TERRITORIALI	a	b	R^2
Nord	10,25183	-1,18783	93,6
Centro	9,09837	-1,12181	91,6
Mezzogiorno	10,24829	-1,20505	93,3
Italia	10,43531	-1,19478	93,9

Prospetto 5 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite agli individui per aree territoriali - Trimestri – Anno 2009

STIME	Trimestri 1,2 e 4				Trimestre 3			
	Aree territoriali				Aree territoriali			
	Nord	Centro	Mezzogiorno	Italia	Nord	Centro	Mezzogiorno	Italia
50.000	44,3	41,5	40,2	49,5	41,9	40,0	38,3	42,7
60.000	40,2	37,5	36,4	44,6	38,1	36,3	35,0	38,8
70.000	37,0	34,4	33,5	40,9	35,2	33,5	32,5	35,8
80.000	34,5	31,9	31,2	37,9	32,9	31,2	30,4	33,4
90.000	32,4	29,9	29,2	35,4	30,9	29,3	28,7	31,5
100.000	30,6	28,2	27,6	33,4	29,3	27,7	27,3	29,8
200.000	21,1	19,1	19,0	22,5	20,5	19,2	19,4	20,8
300.000	17,0	15,2	15,2	17,8	16,7	15,5	15,9	16,9
400.000	14,5	13,0	13,0	15,1	14,4	13,3	13,8	14,6
500.000	12,9	11,5	11,5	13,3	12,8	11,9	12,3	13,0
750.000	10,4	9,1	9,3	10,6	10,4	9,6	10,1	10,5
1.000.000	8,9	7,8	7,9	9,0	9,0	8,2	8,8	9,1
2.000.000	6,1	5,3	5,4	6,1	6,3	5,7	6,2	6,3
3.000.000	4,9	4,2	4,4	4,8	5,1	4,6	5,1	5,1
4.000.000	4,2	3,6	3,7	4,1	4,4	4,0	4,4	4,4
5.000.000	3,8	3,2	3,3	3,6	3,9	3,5	4,0	3,9
7.500.000	3,0	2,5	2,7	2,8	3,2	2,8	3,3	3,2
10.000.000	2,6	2,2	2,3	2,4	-	-	-	2,8
15.000.000	-	-	-	1,9	-	-	-	2,2
20.000.000	-	-	-	-	-	-	-	1,9
25.000.000	-	-	-	-	-	-	-	1,7

Prospetto 6 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite agli individui per aree territoriali - Media dei 4 trimestri – Anno 2009

STIME	Aree territoriali			
	Nord	Centro	Mezzogiorno	Italia
20.000	42,5	35,5	37,7	45,8
30.000	33,8	28,5	30,2	36,3
40.000	28,7	24,3	25,8	30,7
50.000	25,3	21,6	22,8	27,0
60.000	22,9	19,5	20,7	24,3
70.000	21,0	17,9	19,0	22,3
80.000	19,4	16,7	17,7	20,6
90.000	18,2	15,6	16,6	19,3
100.000	17,1	14,8	15,6	18,1
200.000	11,6	10,1	10,7	12,2
300.000	9,2	8,1	8,6	9,6
400.000	7,8	6,9	7,3	8,2
500.000	6,9	6,1	6,5	7,2
750.000	5,5	4,9	5,2	5,7
1.000.000	4,7	4,2	4,4	4,8
2.000.000	3,2	2,9	3,0	3,2
3.000.000	2,5	2,3	2,4	2,6
4.000.000	2,1	2,0	2,1	2,2
5.000.000	1,9	1,7	1,8	1,9
7.500.000	1,5	1,4	1,5	1,5
10.000.000	1,3	1,2	1,3	1,3
15.000.000	1,0	1,0	1,0	1,0
20.000.000	-	-	-	0,9

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 2009

AREE TERRITORIALI	Notti – lavoro			R ²
	a	b	c	
	Gennaio-marzo			
Nord	46.278,941	0,388383	-0,000000041012	90,4
Centro	22.908,792	0,740351	-0,000000135191	88,8
Mezzogiorno	21.303,085	0,452075	-0,000000091117	96,8
Italia	107.662,367	0,290412	-0,000000015398	89,0
	Luglio-settembre			
Nord	60.603,576	0,442747	-0,000000024543	84,5
Centro	28.116,187	0,680537	-0,000000059790	96,2
Mezzogiorno	-4.998,707	0,654988	-0,000000051825	88,6
Italia	202.238,291	0,342060	-0,000000009804	90,6
	Anno			
Nord	165.459,273	0,239302	-0,000000005817	86,5
Centro	148.413,822	0,342507	-0,000000011905	94,2
Mezzogiorno	141.162,827	0,328868	-0,000000010706	77,8
Italia	408.893,027	0,164836	-0,000000001683	88,7
	Notti – vacanze			
	Gennaio-marzo			
Nord	196.473,072	0,169611	-0,000000002113	82,9
Centro	81.725,431	0,196765	-0,000000007121	92,7
Mezzogiorno	94.596,462	0,248432	-0,000000007754	83,4
Italia	280.336,767	0,116994	-0,000000000818	86,9
	Luglio-settembre			
Nord	652.004,620	0,071720	-0,000000000163	90,3
Centro	375.819,176	0,122078	-0,000000000640	91,1
Mezzogiorno	332.696,386	0,166479	-0,000000000630	86,9
Italia	900.579,528	0,060542	-0,000000000065	89,3
	Anno			
Nord	801.598,885	0,060641	-0,000000000088	91,5
Centro	440.078,385	0,095080	-0,000000000336	91,9
Mezzogiorno	483.590,066	0,139504	-0,000000000279	80,9
Italia	1.197.727,936	0,050099	-0,000000000032	88,2
	Notti - totali			
	Gennaio-marzo			
Nord	233.142,328	0,153683	-0,000000001654	85,1
Centro	107.519,870	0,216804	-0,000000007328	90,9
Mezzogiorno	113.434,097	0,217069	-0,000000005433	86,0
Italia	339.254,553	0,106587	-0,000000000638	89,1
	Luglio-settembre			
Nord	682.824,482	0,072587	-0,000000000159	91,4
Centro	436.257,404	0,120238	-0,000000000608	91,5
Mezzogiorno	411.788,839	0,158461	-0,000000000570	87,7
Italia	1.021.421,836	0,059489	-0,000000000061	90,6
	Anno			
Nord	877.447,938	0,060159	-0,000000000084	92,2
Centro	526.625,457	0,093173	-0,000000000309	92,2
Mezzogiorno	570.655,707	0,131846	-0,000000000254	80,7
Italia	1.353.885,335	0,048722	-0,000000000030	89,2

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 2009

AREE TERRITORIALI	Notti – lavoro						
	1%	5%	10%	15%	20%	25%	30%
	Gennaio-marzo						
Nord	9.346.756	8.385.288	7.188.554	6.000.496	4.827.072	3.680.739	2.590.607
Centro	5.433.554	5.139.461	4.772.152	4.405.262	4.038.904	3.673.234	3.308.473
Mezzogiorno	4.899.405	4.465.055	3.923.541	3.384.296	2.848.548	2.318.569	1.798.955
Italia	18.587.026	16.048.797	12.907.664	9.830.100	6.886.916	4.264.197	2.351.169
	Luglio-settembre						
Nord	17.771.420	16.155.499	14.140.012	12.131.648	10.134.485	8.156.306	6.213.683
Centro	11.256.647	10.590.268	9.757.796	8.926.026	8.095.173	7.265.546	6.437.610
Mezzogiorno	12.437.656	11.665.320	10.699.796	9.734.124	8.768.254	7.802.114	6.835.588
Italia	34.469.456	30.467.951	25.499.735	20.592.392	15.796.410	11.227.680	7.168.097
	Anno						
Nord	40.127.940	33.394.503	25.081.363	17.022.779	9.691.369	4.492.446	2.243.555
Centro	28.369.693	25.067.617	20.964.986	16.907.728	12.934.287	9.135.175	5.741.760
Mezzogiorno	30.220.588	26.544.750	21.977.651	17.462.456	13.047.624	8.855.669	5.221.599
Italia	94.551.635	71.612.022	44.033.285	20.603.361	8.316.929	4.415.828	2.919.047
	Notti – vacanze						
	Gennaio-marzo						
Nord	76.734.367	58.193.677	35.552.662	15.339.814	4.837.721	2.304.431	1.471.720
Centro	26.657.744	21.152.635	14.386.372	8.001.496	3.168.192	1.306.764	752.578
Mezzogiorno	31.140.387	26.058.325	19.759.481	13.591.574	7.808.285	3.393.138	1.497.301
Italia	133.331.978	85.866.671	31.608.637	7.206.214	3.271.797	2.081.062	1.521.497
	Luglio-settembre						
Nord	387.787.813	158.074.549	20.601.336	8.189.025	5.050.144	3.645.001	2.850.337
Centro	178.417.712	117.617.298	46.993.374	10.790.653	4.645.732	2.895.913	2.096.455
Mezzogiorno	250.600.662	187.785.030	110.357.893	39.533.492	8.551.361	3.870.452	2.463.108
Italia	796.398.974	224.344.114	22.025.899	9.994.593	6.438.416	4.745.734	3.757.080
	Anno						
Nord	589.188.365	173.049.017	19.512.578	8.892.445	5.731.241	4.224.904	3.344.815
Centro	257.952.786	143.129.112	29.586.437	7.654.147	4.139.462	2.823.365	2.140.040
Mezzogiorno	468.631.155	326.638.988	153.156.237	26.887.394	7.719.339	4.329.277	2.997.497
Italia	1.279.227.040	194.775.480	23.642.628	11.943.314	7.976.497	5.985.848	4.789.860
	Notti – totali						
	Gennaio-marzo						
Nord	88.444.131	64.845.570	36.328.492	13.036.405	4.355.911	2.327.524	1.565.689
Centro	28.732.071	23.390.116	16.812.316	10.512.135	5.144.968	2.185.016	1.171.486
Mezzogiorno	38.650.973	31.413.449	22.475.288	13.851.184	6.402.353	2.452.324	1.263.254
Italia	154.727.029	94.274.031	28.781.451	7.077.913	3.545.850	2.341.181	1.744.004
	Luglio-settembre						
Nord	404.465.820	167.746.894	22.081.783	8.666.315	5.323.775	3.835.598	2.996.295
Centro	185.322.989	121.518.566	48.206.523	11.810.640	5.258.828	3.310.660	2.407.273
Mezzogiorno	263.426.487	194.169.601	109.268.143	35.324.824	8.841.471	4.379.177	2.876.075
Italia	826.746.377	227.754.998	24.317.840	11.200.028	7.246.410	5.352.260	4.242.291
	Anno						
Nord	610.775.266	178.461.063	21.081.325	9.678.575	6.250.998	4.612.543	3.653.752
Centro	274.925.775	150.778.966	31.669.355	8.841.379	4.861.223	3.336.032	2.536.580
Mezzogiorno	483.542.008	328.514.486	141.066.193	23.616.828	8.126.459	4.780.536	3.376.393
Italia	1.330.373.541	192.569.453	26.008.578	13.315.657	8.933.865	6.719.729	5.384.542

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 2009

AREE TERRITORIALI	Viaggi – lavoro			R ²
	a	b	c	
	Gennaio-marzo			
Nord	9.991,399	0,379222	-0,000000102527	96,2
Centro	7.144,712	0,404212	-0,000000326081	91,8
Mezzogiorno	6.403,075	0,336043	-0,000000162706	96,7
Italia	14.994,385	0,241772	-0,000000036016	96,0
	Luglio-settembre			
Nord	14.012,432	0,262007	-0,000000073923	96,2
Centro	3.001,417	0,542770	-0,000000302046	98,1
Mezzogiorno	9.342,502	0,251971	-0,000000136869	91,8
Italia	17.129,509	0,201169	-0,000000028728	96,3
	Anno			
Nord	26.345,881	0,176567	-0,000000012493	95,9
Centro	14.774,554	0,238639	-0,000000038181	96,1
Mezzogiorno	17.248,126	0,179465	-0,000000024422	93,8
Italia	35.716,077	0,114900	-0,000000004147	95,8
	Viaggi – vacanza			
	Gennaio-marzo			
Nord	33.472,254	0,171865	-0,000000011298	92,2
Centro	16.621,993	0,234839	-0,000000042776	95,0
Mezzogiorno	11.180,742	0,298238	-0,000000033487	91,5
Italia	45.641,593	0,129539	-0,000000004341	91,3
	Luglio-settembre			
Nord	57.846,716	0,090279	-0,000000002187	91,9
Centro	32.554,138	0,132493	-0,000000009194	92,1
Mezzogiorno	33.768,427	0,142364	-0,000000009143	92,0
Italia	77.293,871	0,066338	-0,000000000932	91,4
	Anno			
Nord	87.530,329	0,063309	-0,000000000828	89,4
Centro	54.332,207	0,092920	-0,000000002840	91,1
Mezzogiorno	46.536,576	0,102470	-0,000000003240	89,7
Italia	118.867,739	0,046497	-0,000000000331	89,5
	Viaggi – totali			
	Gennaio-marzo			
Nord	31.809,562	0,170867	-0,000000010060	92,1
Centro	16.741,247	0,219735	-0,000000035283	94,3
Mezzogiorno	11.587,725	0,270474	-0,000000026794	91,6
Italia	40.662,844	0,125655	-0,000000003811	92,0
	Luglio-settembre			
Nord	50.343,809	0,091122	-0,000000002176	91,2
Centro	30.451,312	0,138614	-0,000000009300	90,3
Mezzogiorno	28.464,610	0,138202	-0,000000008392	91,6
Italia	64.735,974	0,067481	-0,000000000929	90,7
	Anno			
Nord	58.473,330	0,076491	-0,000000000840	87,2
Centro	37.925,596	0,112051	-0,000000002937	88,1
Mezzogiorno	33.254,651	0,124476	-0,000000003187	85,2
Italia	74.113,038	0,057106	-0,000000000342	86,8

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 2008

AREE TERRITORIALI	Viaggi - lavoro						
	1%	5%	10%	15%	20%	25%	30%
	Gennaio-marzo						
Nord	3.628.091	3.241.155	2.758.734	2.278.501	1.802.128	1.333.456	883.053
Centro	1.226.800	1.106.081	955.858	806.758	659.486	515.436	377.614
Mezzogiorno	2.023.331	1.780.146	1.477.374	1.176.873	880.809	594.971	337.967
Italia	6.499.334	5.401.727	4.039.440	2.702.167	1.447.445	541.037	225.936
	Luglio-settembre						
Nord	3.463.774	2.932.583	2.274.891	1.631.379	1.023.931	524.102	248.584
Centro	1.769.486	1.637.508	1.472.650	1.307.962	1.143.517	979.435	815.930
Mezzogiorno	1.805.705	1.520.544	1.168.743	827.513	512.819	268.563	139.256
Italia	6.742.795	5.372.985	3.683.445	2.069.275	792.788	298.403	165.371
	Anno						
Nord	13.489.148	10.335.072	6.455.468	2.863.115	790.858	339.201	209.021
Centro	6.052.231	5.017.772	3.734.715	2.477.728	1.307.872	490.830	212.645
Mezzogiorno	7.039.384	5.431.221	3.458.082	1.637.747	519.269	226.734	139.172
Italia	25.628.642	16.180.422	5.237.024	917.990	411.447	262.257	192.129
	Viaggi - vacanze						
	Gennaio-marzo						
Nord	14.530.305	11.054.085	6.796.549	2.942.169	879.249	404.706	255.472
Centro	5.329.088	4.409.201	3.270.994	2.162.975	1.151.814	470.850	222.571
Mezzogiorno	8.646.106	7.457.762	5.975.744	4.500.931	3.043.338	1.643.644	552.119
Italia	27.916.052	18.881.092	8.102.869	1.651.825	623.786	373.855	265.953
	Luglio-settembre						
Nord	37.415.876	19.757.059	3.380.218	936.499	521.789	360.395	275.038
Centro	13.584.374	9.351.506	4.348.568	1.156.776	454.144	271.282	192.315
Mezzogiorno	14.728.373	10.455.760	5.327.003	1.549.090	539.688	305.785	211.620
Italia	61.767.566	21.397.927	2.166.230	914.566	575.966	419.954	330.359
	Anno						
Nord	65.975.568	21.084.441	2.269.373	1.000.127	637.887	467.880	369.331
Centro	29.839.708	16.287.937	3.301.589	910.601	500.747	343.752	261.435
Mezzogiorno	29.037.170	17.038.817	4.190.355	921.250	469.819	313.283	234.689
Italia	113.480.826	14.385.923	2.191.978	1.144.259	773.078	583.553	468.614
	Viaggi - totali						
	Gennaio-marzo						
Nord	16.185.735	12.271.983	7.467.677	3.095.614	845.192	383.298	241.777
Centro	6.023.184	4.907.396	3.528.073	2.192.841	1.023.108	382.548	192.316
Mezzogiorno	9.765.700	8.280.762	6.429.699	4.590.548	2.785.499	1.142.632	306.957
Italia	30.695.043	20.375.209	8.056.113	1.374.500	532.414	323.802	232.054
	Luglio-settembre						
Nord	37.882.702	20.047.703	3.184.563	829.617	458.193	315.508	240.418
Centro	14.062.499	9.860.580	4.830.009	1.298.101	463.513	267.414	186.678
Mezzogiorno	15.494.945	10.823.191	5.203.797	1.268.349	434.921	249.919	174.350
Italia	62.981.332	21.986.719	1.888.801	777.690	486.842	354.043	278.103
	Anno						
Nord	79.996.594	33.596.776	2.298.481	788.358	471.920	336.457	261.359
Centro	35.111.108	21.719.751	6.188.942	932.124	425.183	273.333	201.154
Mezzogiorno	36.209.514	23.808.034	8.858.242	1.140.470	432.427	263.168	188.812
Italia	139.219.831	28.395.133	1.704.648	795.495	518.016	383.955	304.994

Prospetto 11 - 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 e tipologie comunali - Stime del numero di viaggi - Anno 2009

TIPOLOGIE COMUNALI	Viaggi – lavoro			R ²
	a	b	c	
Comuni centro dell'area metropolitana	21.332,668	0,210534	-0,000000030649	94,9
Comuni periferia dell'area	15.959,537	0,306182	-0,000000078275	94,4
Comuni con oltre 50.000 abitanti	11.471,282	0,274664	-0,000000055543	97,4
Comuni da 10.001 a 50.000 abitanti	16.556,792	0,263002	-0,000000040650	95,0
Comuni da 2.001 a 10.000 abitanti	19.256,117	0,225709	-0,000000040515	92,7
Comuni aventi fino a 2.000 abitanti	3.694,949	0,581343	-0,000000557430	88,0
Viaggi – vacanze				
Comuni centro dell'area metropolitana	72.579,901	0,101191	-0,000000002566	90,4
Comuni periferia dell'area	48.799,543	0,129198	-0,000000007481	87,4
Comuni con oltre 50.000 abitanti	51.547,327	0,097388	-0,000000003662	90,0
Comuni da 10.001 a 50.000 abitanti	51.191,409	0,091400	-0,000000003108	90,0
Comuni da 2.001 a 10.000 abitanti	51.420,985	0,104643	-0,000000004479	84,8
Comuni aventi fino a 2.000 abitanti	20.787,452	0,225195	-0,000000042401	87,4
Viaggi - totali				
Comuni centro dell'area metropolitana	50.246,946	0,121724	-0,000000002728	85,3
Comuni periferia dell'area	34.951,823	0,152104	-0,000000007077	85,8
Comuni con oltre 50.000 abitanti	34.049,979	0,119955	-0,000000003905	86,6
Comuni da 10.001 a 50.000 abitanti	34.893,741	0,112902	-0,000000003081	86,7
Comuni da 2.001 a 10.000 abitanti	33.720,740	0,129019	-0,000000004272	85,0
Comuni aventi fino a 2.000 abitanti	15.451,865	0,267904	-0,000000039409	88,2

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 2009

TIPOLOGIE COMUNALI	Viaggi – lavoro						
	1%	5%	10%	15%	20%	25%	30%
Comuni centro dell'area metropolitana	6.647.644	5.367.513	3.790.104	2.280.317	1.023.653	409.994	221.619
Comuni periferia dell'area	3.837.016	3.334.012	2.709.336	2.092.734	1.493.086	935.664	492.756
Comuni con oltre 50.000 abitanti	4.808.023	4.095.329	3.209.047	2.333.003	1.483.487	727.822	280.401
Comuni da 10.001 a 50.000 abitanti	6.288.699	5.316.529	4.109.025	2.919.401	1.778.839	817.861	328.756
Comuni da 2.001 a 10.000 abitanti	5.412.027	4.443.868	3.249.082	2.095.497	1.076.201	451.991	230.280
Comuni aventi fino a 2.000 abitanti	1.031.386	960.105	871.113	782.280	693.665	605.362	517.523
Viaggi - vacanze							
Comuni centro dell'area metropolitana	36.316.144	21.278.530	5.555.501	1.386.029	721.048	483.705	363.370
Comuni periferia dell'area	16.332.027	11.169.973	5.165.497	1.517.591	645.244	394.332	282.219
Comuni con oltre 50.000 abitanti	24.441.452	13.950.609	3.412.197	920.752	493.654	335.072	253.254
Comuni da 10.001 a 50.000 abitanti	26.802.791	14.458.484	2.904.184	836.464	465.183	320.755	244.514
Comuni da 2.001 a 10.000 abitanti	21.659.038	13.076.863	3.945.818	1.029.099	526.237	349.981	261.645
Comuni aventi fino a 2.000 abitanti	5.170.091	4.247.313	3.110.285	2.016.553	1.057.718	466.322	244.111
Viaggi - totali							
Comuni centro dell'area metropolitana	41.396.878	26.973.021	9.835.554	1.546.336	628.171	388.501	280.645
Comuni periferia dell'area	20.321.625	14.761.344	7.980.811	2.375.862	664.493	348.261	233.713
Comuni con oltre 50.000 abitanti	28.460.790	18.386.511	6.459.418	1.002.642	416.908	259.806	188.350
Comuni da 10.001 a 50.000 abitanti	33.735.108	20.956.777	6.057.304	876.737	395.102	253.077	185.930
Comuni da 2.001 a 10.000 abitanti	28.141.344	18.914.690	7.804.435	1.275.805	462.211	276.037	196.257
Comuni aventi fino a 2.000 abitanti	6.603.719	5.599.366	4.350.710	3.117.602	1.926.597	893.269	339.724

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:

$$\hat{\sigma}({}_C\hat{Y}_3) = 28.116,19 + 0,680537{}_C\hat{Y}_3 - 0,0000000598({}_C\hat{Y}_3)^2.$$

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 22.025.899 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 21.266.000 (si veda il prospetto 1.1).

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 21.266.000 risulta essere compreso tra i valori 25.499.735 e 20.592.392.

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

Dal prospetto 2.1, risulta che il totale delle persone che hanno viaggiato nel trimestre gennaio-marzo è pari a 11.173.000.

Si cerca il livello di stima che più si avvicina a 11.173.000 nella prima colonna del prospetto 5, che è pari a 10.000.000. 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 2,4%.

L'errore assoluto sarà:

$$\sigma(11.173.000) = 2,4/100 \times 11.173.000 = 268.152$$

L'intervallo di confidenza avrà come estremi:

$$11.173.000 - (2 \times 268.152) = 10.636.696$$

$$11.173.000 + (2 \times 268.152) = 11.709.304$$

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 1,9 percentuali.

L'errore corrispondente a 11.173.000 è pari a:

$$\sigma(11.173.000) = 2,4 - ((2,4 - 1,9) / (15.000.000 - 10.000.000)) \times (11.173.000 - 10.000.000) = 1,28\%$$

Il corrispondente errore assoluto è $1,28/100 \times 11.173.000 = 143.014$ e l'intervallo di confidenza avrà come estremi:

$$11.173.000 - (2 \times 143.014) = 10.886.972$$

$$11.173.000 + (2 \times 143.014) = 11.459.028$$

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 11.173.000 si ha:

$$\log \varepsilon^2 = 11,77505 - 1,20427 \times \log(11.173.000) = -7,76906.$$

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

Infatti, da $\log(\varepsilon^2) = -7,76906$ si ha:

$$\varepsilon^2 = \exp(-7,76906)$$

e dunque:

$$\varepsilon = 0,021$$

L'errore relativo percentuale sarà dunque pari a 2,1%.