



**File a uso pubblico
micro.STAT**

Indagine Viaggi e Vacanze

Periodo di riferimento: anno 2021

Aspetti metodologici dell'indagine

INDICE

1. Introduzione	3
2. La popolazione di riferimento.....	4
3. Il disegno campionario.....	4
4. La rilevazione e il trattamento dei dati	9
5. La metodologia di calcolo dei pesi e degli errori campionari.....	10
6. La diffusione dei risultati dell'indagine	24
7. Glossario	33
8. Riferimenti bibliografici.....	35
9. Contatti.....	35

1. Introduzione

L'indagine ha la finalità di ottenere informazioni sui movimenti turistici della popolazione (domanda turistica). Le stime prodotte riguardano il numero di turisti, di viaggi e di pernottamenti in viaggio, le spese (per il viaggio e giornaliero) e il numero di escursioni sul territorio nazionale o all'estero.

L'indagine viene svolta in conformità alle definizioni concettuali e metodologiche espresse dal Regolamento Europeo per le Statistiche del Turismo 692/2011, in vigore dal 1 gennaio 2012¹.

Il turismo è definito come l'insieme delle attività e dei servizi riguardanti le persone che si spostano al di fuori del loro 'ambiente abituale', per vacanza o per motivi di lavoro. Rientrano pertanto nei flussi turistici tutti gli spostamenti non abituali, con pernottamento (viaggi) o senza (escursioni). L'individuazione dell'ambiente abituale di una persona permette di distinguere correttamente il fenomeno turistico dalla mobilità, che non rientra nel campo di osservazione della domanda turistica.

Ad esempio, i viaggi e le escursioni abituali, quelli cioè effettuati settimanalmente nella stessa località, diversa dal luogo in cui si vive, sono comunque assimilabili all'ambiente abituale e non rientrano nei flussi turistici; si presuppone, infatti, che tali spostamenti siano riconducibili alla vita quotidiana e alle abitudini dell'individuo. Sono altresì esclusi dalla definizione di 'turista' le persone che si spostano giornalmente o settimanalmente per lavoro, per studio o per motivi personali, quando cioè lo spostamento rientra nell'ambito di attività di routine.

I viaggi turistici (non abituali) sono classificati, secondo gli standard internazionali, distinguendo i viaggi per motivi di lavoro da quelli per motivi di vacanza e le vacanze 'brevi' (da 1 a 3 notti) da quelle 'lunghe' (più di 3 notti). Tra le vacanze rientrano i viaggi per svago, piacere, relax, per visitare parenti o amici, per trattamenti di salute o per motivi religiosi.

In questa nota sono riportati in sintesi gli aspetti salienti dell'indagine. Per maggiori informazioni sulla rilevazione, è possibile consultare la pagina informativa dell'indagine sul sito dell'Istituto <http://www.istat.it/it/archivio/123949>

¹ E dalle successive modifiche introdotte dal Regolamento Delegato (UE) 2019/1681 del 1 agosto 2019

2. La popolazione di riferimento

La popolazione di riferimento è costituita dalle famiglie residenti e dagli individui che le compongono.

Per famiglia si intende la famiglia di fatto, definita come l'insieme delle persone coabitanti, legate da vincoli di matrimonio o parentela, affinità, adozione, tutela o affettivi e che partecipano alla spesa familiare e/o condividono il reddito familiare.

Sono quindi escluse dalla popolazione di riferimento le persone che vivono permanentemente nelle comunità (caserme, ospedali, brefotrofi, istituti religiosi, convitti, eccetera) e quelle presenti ma non residenti sul territorio nazionale.

La popolazione utilizzata per l'indagine 2021 è quella stimata precedentemente al rilascio dei dati di censimento 2018 e 2019 e della ricostruzione intercensuaria.

3. Il disegno campionario

Le informazioni sui viaggi e sulle escursioni sono rilevate attraverso un focus incluso nell'indagine sulle Spese delle Famiglie. 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 cinque grandi ripartizioni geografiche (Nord-ovest: Piemonte, Valle d'Aosta, Lombardia, Liguria; Nord-est: Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna; Centro: Toscana, Umbria, Marche, Lazio; Sud: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria; Isole: Sicilia, Sardegna);

Il disegno d'indagine è stato progettato, inoltre, per fornire stime a livello regionale e a livello di tipologia comunale sull'intero anno. Tuttavia, poiché per alcune regioni l'esiguità del fenomeno non permette un'adeguata robustezza delle stime, non sono diffuse stime per tale dominio territoriale. Le stime sulla tipologia comunale, invece, non sono ottenibili dal file mlcro.STAT poiché l'informazione sulla tipologia comunale è stata soppressa per motivi di tutela della riservatezza dei dati.

La base di campionamento adottata, ossia la lista di selezione delle unità campionarie, è costituita dalle LAC, un archivio unico delle famiglie residenti nei comuni italiani, proveniente dalle liste anagrafiche comunali.

3.1. Descrizione generale

Il tipo di disegno adottato è a due stadi di selezione (comuni, famiglie). Ciò è determinato dal fatto che l'indagine utilizza l'intervista diretta con tecnica CAPI, e pertanto è necessario, per ragioni sia di costo e sia soprattutto organizzative, che il campione sia concentrato in un numero limitato di comuni.

Le unità di primo stadio, i comuni, sono stratificati all'interno di ciascuna regione geografica secondo le due variabili tipologia comunale e dimensione demografica. La prima è definita sulla base dei domini illustrati nel precedente paragrafo, la seconda variabile di stratificazione – espressa in termini di popolazione residente – viene utilizzata per costruire una soglia in base alla quale suddividere i comuni in due sottoinsiemi: (i) gli Auto Rappresentativi o AR, sono i comuni con popolazione al di sopra della soglia e costituiscono ognuno uno strato a sé; (ii) i Non Auto Rappresentativi o NAR, sono i restanti comuni. Per i comuni AR si adotta un disegno ad uno stadio stratificato, mentre per i NAR si usa un disegno a due stadi con stratificazione delle unità di primo stadio.

Da ogni comune campione si estrae un campione di famiglie anagrafiche che costituiscono le unità finali di campionamento.

Ogni trimestre si intervistano circa 8.100 famiglie, per un totale di 32.500 famiglie all'anno². Il disegno di campionamento, definito con riferimento a un generico trimestre dell'anno, è replicato in modo identico per i quattro trimestri e viene anche effettuata una stratificazione mensile del campione trimestrale. Di conseguenza, la dimensione temporale può essere considerata un'ulteriore variabile di stratificazione del campione.

Essendo l'indagine basata su una rilevazione trimestrale, il disegno di campionamento viene definito con riferimento al generico trimestre di un anno e il campione di comuni risultante è coinvolto quattro volte in modo identico nella rilevazione nell'anno di riferimento. Inoltre, viene attuata una stratificazione mensile suddividendo il campione di ciascuno strato della popolazione nei tre mesi che compongono ciascun trimestre. In particolare, la partecipazione dei comuni inclusi nel campione nell'arco dell'anno di riferimento è la seguente:

² Nel triennio 2020-2022

- ogni comune AR è coinvolto nell'indagine tutti i mesi (e il numero di famiglie campione è suddiviso nei mesi);
- ciascuno dei tre comuni campione di ogni strato NAR partecipa all'indagine quattro mesi nell'anno a distanza di tre mesi secondo lo schema seguente:

comuni	mese di rilevazione			
1	gennaio	aprile	luglio	ottobre
2	febbraio	maggio	agosto	novembre
3	marzo	giugno	settembre	dicembre

3.2. Stratificazione e selezione del campione

L'obiettivo della stratificazione è quello di formare gruppi (o strati) di unità caratterizzate, relativamente alle variabili oggetto d'indagine, da massima omogeneità interna agli strati e massima eterogeneità fra gli strati. Il raggiungimento di tale obiettivo si traduce in termini statistici in un guadagno nella precisione delle stime, ossia in una riduzione dell'errore campionario a parità di numerosità campionaria.

Nell'indagine in esame, i comuni vengono stratificati in base alla loro dimensione demografica e nel rispetto delle seguenti condizioni:

- auto-ponderazione del campione al livello di regione;
- definizione di un numero minimo di famiglie da intervistare per comune;
- stratificazione dei comuni sulla base dell'ampiezza demografica;
- formazione di strati di comuni di ampiezza approssimativamente costante in termini di popolazione residente.

Il procedimento di stratificazione è attuato all'interno di ogni dominio territoriale individuato, per ciascuna regione geografica, dalle tre tipologie comunali considerate.

Con riferimento al generico incrocio di regione e tipologia comunale, indicato nel seguito con il termine *dominio*, si denoti con: h ($h = 1, \dots, H$), l'indice di strato di comuni; i ($i = 1, \dots, N$), l'indice di comune; M_h , il numero di famiglie residenti nello strato h ; M_{hi} il numero di famiglie residenti nel comune i dello strato h ; P_h , il numero di individui residenti nello strato h ; P_{hi} il numero di individui residenti nel comune i dello strato h ; n , il numero di comuni campione in ogni strato.

Per la definizione del campione di comuni e di famiglie relativo ad un trimestre sono state dapprima effettuate le seguenti scelte:

- (a) definizione del numero complessivo di famiglie campione a livello nazionale, pari a 32.500 famiglie annue;
- (b) definizione del numero m_r di famiglie campione per ciascuna regione, definito mediante una procedura di allocazione ottima multivariata del campione (Bethel, 1989³; Falorsi et al., 1998⁴) che consente di tenere sotto controllo gli errori campionari attesi dei principali capitoli di spesa (di interesse dell'indagine sulla spesa);
- (c) distribuzione delle numerosità regionali tra i domini definiti dalla tipologia comunale in misura proporzionale alla popolazione residente;
- (d) scelta del numero minimo di famiglie, \bar{m}_r , da intervistare in ciascun comune campione.

Dalla scelta di \bar{m}_r e m_r dipende la suddivisione dei comuni in AR e NAR e la formazione degli strati attraverso i seguenti passi:

- (1) calcolo della frazione di campionamento regionale $f_r = m_r/M_r$, essendo M_r il numero di famiglie residenti nella regione r ;
- (2) determinazione del valore della *soglia*⁵ λ_r , mediante la relazione

$$\lambda_r = \frac{\bar{m}_r \delta_r}{f_r},$$

in cui δ_r è il numero medio di componenti per famiglia a livello regionale; risulta evidente da tale espressione che la soglia per la definizione dei comuni AR cresce al crescere di \bar{m}_r ;

- (3) suddivisione dei comuni in AR e NAR sulla base della soglia λ_r ;
- (4) ordinamento decrescente dei comuni NAR all'interno di ogni dominio in funzione della loro dimensione demografica;
- (5) suddivisione dei comuni NAR in strati la cui dimensione è approssimativamente uguale al prodotto $n \times \lambda_r$;

³ Bethel J. (1989). Sample Allocation in Multivariate Surveys, Survey Methodology, 15, 47-57.

⁴ Falorsi P.D., Ballin M., De Vitiis C., Scepti G. (1998) "Principi e metodi del software generalizzato per la definizione del disegno di campionamento nelle indagini sulle imprese condotte dall'ISTAT", Statistica Applicata Vol. 10, n.2.

⁵ Dal momento che il campione di famiglie di ogni regione è distribuito in modo proporzionale tra i domini e, di conseguenza, la frazione di campionamento è costante all'interno della regione, la soglia risulta anch'essa definita a livello regionale, così come la suddivisione dei comuni in AR e NAR.

(6) selezione di n comuni campione da ciascuno strato h ($h=1,\dots,H$) con probabilità proporzionale all'ampiezza; per il generico comune c tale probabilità è espressa dalla formula:

$$z_{hi} = n \cdot P_{hi} / P_h ;$$

(7) definizione del numero m_{hi} di famiglie da intervistare in ogni comune; dalla condizione di autoponderazione a livello regionale

$$\frac{n \cdot P_{hi}}{P_h} \frac{m_{hi}}{M_{hi}} = f_r ,$$

in cui il primo membro rappresenta la probabilità d'inclusione delle famiglie del comune c dello strato h (essendo le due frazioni rispettivamente la probabilità d'inclusione di primo e di secondo stadio), si ottiene

$$m_{hi} = \frac{f_r}{n} \frac{P_h \cdot M_{hi}}{P_{hi}} .$$

Al fine di raggiungere la numerosità campionaria desiderata in termini di famiglie e di comuni nel rispetto della condizione di auto-ponderazione a livello regionale e degli errori campionari attesi per le stime di spesa, il numero minimo di interviste per comune è stato posto pari a 8.

Relativamente alla quantità m , per consentire la stratificazione mensile del campione di comuni, si è scelta una numerosità per ogni strato NAR pari a 3.

Nella tabella seguente sono riportate le numerosità campionarie teoriche in termini di comuni.

	Comuni campione		
	AR	NAR	TOTALI
Mese	65	159	224
Trimestre	65	477	542
Anno	65	477	542

L'estrazione delle famiglie campione viene effettuata, mediante selezione casuale, dalla lista delle anagrafi comunali; per ciascun comune campione vengono selezionate quindi complessivamente $4m_{hi}$ famiglie che vengono ripartite nell'anno come segue:

- il campione di famiglie di ciascun comune AR viene suddiviso in 12 campioni mensili;
- il campione di famiglie di ciascun comune NAR viene suddiviso in 4 campioni mensili;
- ciascun campione mensile viene suddiviso in due gruppi, uno per ciascun dei due periodi di rilevazione di 14 giorni di ogni mese.

4. La rilevazione e il trattamento dei dati

Come già accennato (par.3), le informazioni vengono rilevate attraverso un apposito focus incluso nell'intervista dell'indagine sulle Spese delle Famiglie. L'intervista è di tipo diretto, condotta mediante tecnica Capi (intervista faccia a faccia assistita da computer). Ogni individuo della famiglia viene intervistato sui viaggi e sulle escursioni effettuate nel periodo di riferimento. L'indagine è continua su tutti i mesi dell'anno, consentendo così di cogliere la stagionalità del fenomeno del turismo. Le famiglie del campione annuale vengono suddivise in dodici sotto-campioni, ciascuno dei quali partecipa alla rilevazione in uno specifico mese di riferimento. Ogni famiglia riceve tre visite del rilevatore, secondo un preciso calendario. I quesiti sulla domanda turistica sono somministrati durante la prima e la terza visita (intervista iniziale e finale). La rilevazione dei viaggi e delle escursioni fa riferimento al mese, tuttavia nella rilevazione sono inseriti anche quesiti aventi periodi di riferimento diversi. In particolare, il numero di viaggiatori per vacanza viene rilevato anche con riferimento all'ultimo anno, così da poter soddisfare le richieste incluse nel nuovo Regolamento Europeo.

L'utilizzo della tecnica Capi consente di anticipare alla fase di acquisizione dell'informazione una parte considerevole dei controlli, ottenendo dati grezzi meno affetti da errore; le risposte fornite vengono sottoposte ad un piano di verifica implementato nel questionario elettronico e quindi risolte sin dal momento della rilevazione. Tale piano si basa sul controllo dei domini delle variabili, dei percorsi del questionario e delle incongruenze logiche tra le informazioni raccolte.

Il questionario elettronico prevede sia controlli di tipo hard, sia controlli di tipo soft. Nel primo caso, a fronte di un'informazione incoerente, la registrazione del dato errato viene inibita e quindi l'intervista non può proseguire finché esso non viene corretto; nel secondo caso, la registrazione del dato errato viene consentita previa segnalazione dell'errore.

Generalmente, un controllo di tipo hard viene attivato se l'informazione risulta assolutamente inaccettabile, un controllo di tipo soft quando essa è altamente improbabile.

La strategia di controllo e correzione successiva cerca di salvaguardare il più possibile l'informazione rilevata, partendo dal presupposto che l'insieme dei controlli inseriti in fase di acquisizione fornisca dati grezzi di elevata qualità.

5. La metodologia di calcolo dei pesi e degli errori campionari

Dato il carattere campionario dell'indagine, per poter ottenere stime per l'intera popolazione di riferimento (le famiglie residenti) è necessario utilizzare un opportuno coefficiente di riporto all'universo. Il valore di tale coefficiente indica il numero di unità nella popolazione rappresentate dall'unità campionaria. L'indagine produce stime trimestrali e stime annuali.

Il calcolo del coefficiente di riporto all'universo utilizza una procedura generalizzata basata sull'uso di una famiglia di stimatori noti in letteratura come *calibration estimator* (stimatori di ponderazione vincolata). La metodologia alla base di tali stimatori consente la determinazione di un unico coefficiente di riporto all'universo in grado di produrre stime coerenti a totali noti, desunti da fonti esterne, che siano correlati alle principali variabili oggetto di indagine.

La famiglia di stimatori di ponderazione vincolata coincide asintoticamente con lo stimatore di regressione generalizzato e, per campioni sufficientemente grandi, ha approssimativamente le stesse proprietà, ovvero si tratta di stimatori corretti, consistenti e con la stessa varianza campionaria⁶.

La strategia adottata per la costruzione dei coefficienti di riporto all'universo è simile a quella utilizzata dalle varie indagini campionarie sulle famiglie dell'Istituto e si sviluppa in tre fasi:

- la determinazione della probabilità di inclusione di ogni unità statistica e del relativo peso diretto, pari all'inverso della probabilità di inclusione;
- calcolo dei coefficienti di correzione per mancata risposta totale;
- determinazione dei coefficienti di riporto all'universo finali vincolati ai totali noti desunti da fonti esterne all'indagine.

⁶ La metodologia è illustrata da Deville, J.C. e Särndal, C.E. in *Calibration Estimators in Survey Sampling*, Journal of the American Statistical Association, Vol. 87, n.418, 1992.

5.1 La probabilità di inclusione e il peso diretto

Il principio su cui è basato ogni metodo di stima campionaria è che le unità appartenenti al campione rappresentino anche le unità della popolazione non incluse nel campione stesso. Per questo motivo ad ogni unità campionaria è possibile attribuire un peso, il coefficiente di riporto all'universo, che indica quante unità della popolazione essa rappresenta.

Senza perdere di generalità, definiamo la seguente simbologia:

U popolazione di riferimento oggetto di indagine;

y_k valore della variabile Y assunto dalla k-esima osservazione della popolazione;

y_j valore della variabile Y assunto dalla j-esima osservazione della popolazione;

π_j probabilità, assegnata dal disegno di campionamento, che l'unità j-esima sia inclusa nel campione S;

Il totale di una generica variabile Y, calcolato sull'intera popolazione, assume la seguente forma:

$$Y = \sum_{k \in U} y_k \quad (1)$$

Il disegno di campionamento assegna le probabilità di inclusione ad ogni unità del campione in modo tale che

$$\hat{Y} = \sum_{j \in S} y_j \frac{1}{\pi_j} \quad (2)$$

sia uno stimatore corretto della (1).

Nel disegno di campionamento dell'indagine, la probabilità di inclusione di un generico individuo è data dalla probabilità di estrazione del suo comune di residenza (direttamente proporzionale all'ampiezza demografica dei comuni all'interno dello strato) e dalla probabilità di estrazione della sua famiglia tra le famiglie eleggibili del comune. Per la natura del disegno campionario, la probabilità di inclusione, e dunque il peso base, è uguale per tutte le famiglie appartenenti allo stesso comune.

Per una generica famiglia eleggibile j , nel comune i dello strato h , il peso diretto d_{hij} , inverso della probabilità di inclusione π_{hij} , assume la seguente forma:

$$d_{hij} = \frac{1}{\pi_{hij}} = \frac{P_h}{P_{hi}} \frac{M_{hi}}{m_{hi}} \quad (3)$$

h denota l'indice di strato;

i è l'indice di comune;

j denota l'indice della famiglia;

P_h indica il totale della popolazione residente nello strato h ;

P_{hi} indica il totale della popolazione residente nel comune i dello strato h ;

M_{hi} indica il totale di famiglie eleggibili nel comune i dello strato h ;

m_{hi} indica il numero di famiglie campione nel comune i dello strato h .

I pesi base vengono di volta in volta calcolati sui dati anagrafici più aggiornati per popolazione e famiglie residenti.

5.2 La correzione per mancata risposta

Nel corso della fase di raccolta delle informazioni, come sempre accade nelle indagini, alcune unità campionarie non partecipano (per irreperibilità, rifiuto, interruzione definitiva o altri motivi). Tale fenomeno, noto come mancata risposta totale, implica che la numerosità campionaria teorica e quella effettivamente raggiunta siano diverse. Nella fase di calcolo dei coefficienti di riporto all'universo si tiene conto di tale fenomeno e viene introdotto un correttore per mancata risposta totale: sotto l'ipotesi che in determinati gruppi (ottenuti come aggregazioni di strati appartenenti alla stessa regione e tipologia comunale) il comportamento dei rispondenti sia simile a quello dei non rispondenti, il correttore assume la forma dell'inverso del tasso di risposta (δ_g):

$$\frac{1}{\delta_g} = \frac{m_g}{m_g^r} \quad (4)$$

in cui m_g^r rappresenta il numero di famiglie rispondenti nel determinato gruppo g .

Il coefficiente di riporto all'universo, corretto per mancata risposta, per una generica famiglia eleggibile j , nel comune i dello strato h , del gruppo g (dato come intersezione tra regione e tipologia comunale), risulta pertanto essere:

$$k_{ghij} = d_{hij} \frac{1}{\delta_g} = \frac{P_h M_{hi} m_g}{P_{hi} m_{hi} m_g^r} \quad (5)$$

5.3 La calibrazione a fonti esterne

Per il calcolo dei coefficienti di riporto all'universo finali si adottano gli stimatori di ponderazione vincolata (*calibration estimator*). La metodologia si basa sull'utilizzo di informazioni ausiliarie, sintetizzate in totali noti, che, essendo correlate con le variabili principali oggetto di indagine, hanno la funzione di aumentare l'accuratezza delle stime. I pesi finali si ottengono risolvendo un problema di minimo vincolato, in cui la funzione da minimizzare è una funzione di distanza tra i pesi diretti corretti per la mancata risposta (\mathbf{k}) e i pesi finali (\mathbf{w}) delle famiglie del campione rispondente (S^r); i vincoli sono proprio le condizioni di uguaglianza tra le stime campionarie delle variabili ausiliarie e i rispettivi totali noti desunti da fonti esterne all'indagine⁷. Formalmente:

$$\begin{cases} \text{Min} \left\{ \sum_{j \in S^r} \text{dist}(k_j, w_j) \right. \\ \left. \sum_{j \in S^r} x_j * w_j = t \right. \end{cases}$$

dove t è il vettore dei totali noti e x_j è il vettore delle variabili ausiliarie osservate sulla j -esima unità campionaria appartenente al campione rispondente (S^r). La funzione di distanza utilizzata è la logaritmica troncata.

I totali noti, introdotti come vincoli nel calcolo dei pesi finali, consentono di migliorare l'accuratezza delle stime: quanto più le variabili ausiliarie considerate sono correlate con le variabili oggetto d'indagine, tanto più si riduce la distorsione delle stime. Nello specifico dell'Indagine sulle Spese delle Famiglie, le stime campionarie sono state vincolate ai seguenti totali noti:

⁷ La calibrazione è una calibrazione integrata, vale a dire che i coefficienti di riporto all'universo sono stati determinati in modo da poter essere utilizzati, indifferentemente, per la costruzione delle stime relative alle famiglie e agli individui. Tutti gli individui di una stessa famiglia hanno il medesimo coefficiente di riporto della famiglia stessa.

- Popolazione residente per ripartizione, sesso e classi di età (0-14, 15-29, 30-44, 45-59, 60-74, 75 e più);
- Popolazione residente per regione;
- Famiglie residenti per regione;
- Popolazione residente per ripartizione e tipologia di comune (Comuni metropolitani, comuni periferia dei comuni metropolitani e comune oltre 50.000 abitanti, altri comuni);
- Famiglie residenti per ripartizione e tipologia di comune (Comuni metropolitani, comuni della cintura metropolitana e comuni con più di 50.000 abitanti, altri comuni);
- Popolazione straniera residente per ripartizione;
- Popolazione di 15 anni e più per ripartizione, condizione e posizione nella professione (Dirigenti, quadri e impiegati; operai e assimilati; imprenditori e liberi professionisti; lavoratori in proprio e altri autonomi; disoccupati; inattivi);
- Popolazione 15-89 anni per titolo di studio (fino alle medie; secondario; post secondario);
- Popolazione per ripartizione e mese;
- Famiglie per ripartizione e mese.

I vincoli imposti sono gli stessi sia per le stime trimestrali che per le stime annuali. Si è adottata una procedura di individuazione delle osservazioni anomale, vale a dire osservazioni con livelli e comportamenti di spesa che, seppur veri, sono molto lontani o dissimili dal collettivo socio-economico al quale appartengono. Tali unità campionarie rappresentano quindi un numero ristretto di famiglie dell'universo ed è quindi opportuno che abbiano associato un coefficiente di riporto abbastanza basso per non influenzare la stima dal punto di vista sia della distorsione che della varianza. I gruppi socio-economici all'interno dei quali individuare comportamenti di spesa anomali sono stati definiti dall'incrocio tra ripartizione di appartenenza, tipologia di comune e numerosità familiare; all'interno di ogni gruppo, le famiglie anomale sono individuate tramite una *cluster analysis* che considera tutti i dati di spesa rilevati dall'indagine, sia a livello complessivo che rispetto alla loro composizione tra le diverse voci o capitoli.

Alle famiglie individuate come anomale è attribuito, nella fase precedente alla calibrazione finale, un peso corretto per la mancata risposta più basso rispetto a quello che avrebbero

avuto tramite la procedura generalizzata. Questa procedura permette di mantenere nel campione le osservazioni anomale e rende possibile l'uso degli stimatori classici.

Per non avere famiglie con coefficienti troppo elevati è stata adottata una procedura di calibrazione ricorsiva, individuando le famiglie con valori dei coefficienti esterni a un intervallo di accettazione (inferiori al quinto percentile o superiori al 95esimo) e imponendo a queste valori dei coefficienti pari ai limiti del suddetto intervallo. La procedura di calibrazione viene nuovamente eseguita in modo da rispettare i vincoli imposti.

5.4 I dati campionari e il loro utilizzo

Il file dei microdati per la ricerca si compone di tre diversi archivi: individui, viaggi ed escursioni. I file sono riferiti all'intero anno e sono ottenuti mettendo in sequenza i quattro file trimestrali.

Il file individui è composto da un record per ciascun individuo componente la famiglia intervistata, include quindi sia gli individui che hanno viaggiato (turisti), sia quelli che non hanno viaggiato. Per l'anno 2021, il file contiene n. 64.556 records ed è di lunghezza pari a 106.

Il file viaggi è composto da un record per ciascun viaggio realizzato da un individuo. Per distinguere i viaggi di lavoro dai viaggi di vacanza è presente una variabile (TIPOVGG) che individua la tipologia del viaggio. Per una corretta utilizzazione dei dati, si rammenta che l'indagine rileva i viaggi che si concludono in ciascun mese di riferimento, indipendentemente dalla data di inizio. Per l'anno 2021, il file contiene n. 3.517 records ed è di lunghezza pari a 140.

Il file escursioni è composto da un record per ciascuna escursione realizzata da un individuo nel mese di riferimento. Per l'anno 2021, il file contiene n. 4.284 records ed è di lunghezza pari a 90.

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

- individui: ogni individuo è identificato da: anno di riferimento (ANNRIF), mese di riferimento (MESE), numero progressivo univoco (PROGIND);
- viaggi: ogni viaggio è individuato da: anno di riferimento (ANNRIF), mese di riferimento (MESE), numero progressivo univoco dell'individuo che ha viaggiato (PROGIND); numero progressivo del viaggio tra quelli dello stesso individuo (PROGVIA);

- notti: le notti trascorse in viaggio sono espresse dalla durata del viaggio (DURATA) presente nel file viaggi.
- escursioni: ogni escursione è individuata da: anno di riferimento (ANNRIF), mese di riferimento (MESE), numero progressivo univoco dell'individuo che ha effettuato l'escursione (PROGIND); numero progressivo dell'escursione tra quelle dello stesso individuo (PROGESC).
- spese per viaggio e spese giornaliere (in euro): sono informazioni presenti nel file viaggi.

Il coefficiente di riporto all'universo è stato determinato in modo da poter essere utilizzato indifferentemente per costruire stime relative agli individui, ai viaggi, alle notti e alle escursioni.

Il coefficiente può avere fino a sette cifre decimali separati dal punto (es. "37377.259718"). Per l'elaborazione dei dati si consiglia di utilizzarlo arrotondato all'unità.

Ad esempio, se in un record individuale del primo trimestre dell'anno il coefficiente di riporto all'universo vale "37377.259718", significa che l'individuo rappresenta 37.377,259718 individui nella popolazione residente in Italia in quel trimestre. Ponderando quindi per questo valore i dati presenti nel suo record, le caratteristiche relative a tale individuo saranno stimate per 37.377,259718 individui.

Per quanto appena detto, è importante sottolineare che, mentre i viaggi, le notti e le escursioni rilevati in ciascun trimestre possono essere sommati per ottenere le stime annuali, non è corretto sommare i record trimestrali degli individui, per non quadruplicare la popolazione. Inoltre, poiché, una stessa persona può essere stata turista in mesi diversi, anche un'eventuale somma dei turisti mensili comporterebbe una sovrastima del numero dei turisti nel trimestre o nell'anno. Pertanto, per gli individui (turisti e non turisti), non è possibile ottenere stime di ammontare, ma solo medie.

Le stime ottenibili dal file dei microdati per la ricerca sono:

- ammontare dei viaggi ed escursioni nel mese, nel trimestre e nell'anno; spese medie per viaggi nel trimestre e nell'anno. A questo fine si deve utilizzare il coefficiente denominato COEV;
- medie mensili (turisti, non turisti, viaggi ed escursioni); in tal caso si deve utilizzare il coefficiente denominato W;
- spese medie giornaliere nel trimestre e nell'anno; a questo fine si deve utilizzare il coefficiente denominato COEV moltiplicato per la DURATA del viaggio.

Per completezza, si precisa che le stime dell'ammontare dei turisti e dei non turisti nell'anno e dei turisti nel trimestre vengono fornite esclusivamente nel datawarehouse dell'Istituto I.Stat (vedere Par. 6) e pertanto non sono ottenibili a partire dal presente file.

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, di seguito, le informazioni relative alla valutazione degli errori di stima, con esempi di calcolo.

5.5 Valutazione del livello di precisione delle stime

Le principali statistiche di interesse per valutare la variabilità campionaria delle stime prodotte da un'indagine sono l'errore di campionamento assoluto e l'errore di campionamento relativo. Indicando con $\hat{Var}(\hat{Y}_d)$ la stima della varianza della generica stima \hat{Y}_d , la stima dell'errore di campionamento assoluto di \hat{Y}_d si può ottenere mediante la seguente espressione

$$\hat{\sigma}(\hat{Y}_d) = \sqrt{\hat{Var}(\hat{Y}_d)}; \quad (1)$$

la stima dell'errore di campionamento relativo di \hat{Y}_d è invece definita dall'espressione

$$\hat{\varepsilon}(\hat{Y}_d) = \frac{\hat{\sigma}(\hat{Y}_d)}{\hat{Y}_d}. \quad (2)$$

Come è stato descritto nel paragrafo 5, le stime prodotte dall'indagine sono state ottenute mediante uno stimatore di ponderazione vincolata definito in base ad una funzione di distanza di tipo logaritmico troncato. Poiché, lo stimatore adottato non è funzione lineare dei dati campionari, per la stima della varianza $\hat{Var}(\hat{Y}_d)$ si è utilizzato il metodo proposto da Woodruff; in base a tale metodo, che ricorre all'espressione linearizzata in serie di Taylor, è possibile ricavare la varianza di ogni stimatore non lineare (funzione regolare di totali) calcolando la varianza dell'espressione linearizzata ottenuta. In particolare, per la definizione dell'espressione linearizzata dello stimatore ci si è riferiti allo stimatore di regressione generalizzata, sfruttando la convergenza asintotica di tutti gli stimatori di ponderazione vincolata a tale stimatore, poiché nel caso di stimatori di ponderazione

vincolata che utilizzano funzioni distanza differenti dalla distanza euclidea (che conduce allo stimatore di regressione generalizzata) non è possibile derivare l'espressione linearizzata dello stimatore. L'espressione linearizzata dello stimatore è data, quindi, da

$$\hat{Y}_d \cong \hat{Z}_d = \sum_{h=1}^{H_d} \hat{Z}_h, \quad \text{essendo} \quad \hat{Z}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} Z_{hij} W_{hij} \quad (3)$$

dove Z_{hij} è la variabile linearizzata espressa come $Z_{hij} = Y_{hij} - \mathbf{X}'_{hij}\beta$, essendo $\mathbf{X}_{hij} = (X_{hij,1}, \dots, X_{hij,K})'$ il vettore contenente i valori delle variabili ausiliarie, osservati per la generica famiglia hij e $\hat{\beta}$, il vettore dei coefficienti di regressione del modello lineare che lega la variabile di interesse y alle K variabili ausiliarie x. In base alla (3), si ha, quindi, che la stima della varianza della generica stima \hat{Y}_d è ottenuta mediante la seguente relazione

$$\hat{\text{Var}}(\hat{Y}_d) \cong \hat{\text{Var}}(\hat{Z}_d) = \sum_{h=1}^{H_d} \hat{\text{Var}}(\hat{Z}_h). \quad (4)$$

Dalla (4) risulta che la stima della varianza della stima \hat{Y}_d viene calcolata come somma della stima delle varianze dei singoli strati, AR e NAR, appartenenti al dominio d. La formula di calcolo della varianza, $\hat{\text{Var}}(\hat{Z}_h)$, della stima \hat{Z}_h è differente a seconda che lo strato sia AR oppure NAR. Possiamo, quindi scomporre come segue

$$\hat{\text{Var}}(\hat{Y}_d) \cong \hat{\text{Var}}(\hat{Z}_d) = \sum_{h=1}^{H_{AR}} \hat{\text{Var}}(\hat{Z}_h) + \sum_{h=1}^{H_{NAR}} \hat{\text{Var}}(\hat{Z}_h), \quad (5)$$

in cui H_{AR} e H_{NAR} indicano rispettivamente il numero di strati AR e NAR appartenenti al dominio d.

Negli strati AR (in cui ciascun comune fa strato a sé e $N_h = n_h = 1$, l'indice i di comune diviene superfluo e viene omesso) la varianza è stimata mediante la seguente espressione

$$\sum_{h=1}^{H_{AR}} \hat{\text{Var}}(\hat{Z}_h) = \sum_{h=1}^{H_{AR}} M_h \frac{(M_h - m_h)}{m_h(m_h - 1)} \sum_{j=1}^{m_h} (Z_{hj} - \bar{Z}_h)^2, \quad (6)$$

dove si è posto $M_h = M_{hi}$, $m_h = m_{hi}$, $Z_{hj} = Z_{hij}$ e $\bar{Z}_h = \frac{1}{m_h} \sum_{j=1}^{m_h} Z_{hj}$.

Per l'insieme degli strati NAR la varianza viene stimata invece mediante la formula seguente

$$\sum_{h=1}^{H_{NAR}} \hat{\text{Var}}(\hat{Z}_h) = \sum_{h=1}^{H_{NAR}} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\hat{Z}_{hi} - \frac{\hat{Z}_h}{n_h} \right)^2 \quad (7)$$

dove le quantità sono espresse come

$$\hat{Z}_{hi} = \sum_{j=1}^{m_{hi}} Z_{hij} W_{hij} \quad \text{e} \quad \hat{Z}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} Z_{hij} W_{hij} .$$

Utilizzando le espressioni (6) e (7) è possibile, infine, calcolare la varianza di campionamento, $\hat{\text{Var}}(\hat{Y}_d)$, in base alla (5) e calcolare, quindi, in base alla (1) ed alla (2) rispettivamente l'errore di campionamento assoluto e l'errore di campionamento relativo.

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

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

in cui si è indicato 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, ottenuta secondo il procedimento sopra descritto.

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

$$\hat{\varepsilon}_{(d\hat{R})} = \sqrt{\hat{\varepsilon}^2_{(d\hat{Y}_1)} - \hat{\varepsilon}^2_{(d\hat{Y}_2)}} . \quad (9)$$

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:

$$\{\hat{Y} - k\hat{\sigma}(\hat{Y}), \hat{Y} + k\hat{\sigma}(\hat{Y})\} \quad (10)$$

Nella (10) il valore di k dipende dal valore fissato per P; ad esempio, per P=0,95 si ha k=2.

5.6 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.

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

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

$$\log \hat{\varepsilon}^2({}_d\hat{Y}_t) = a + b \log({}_d\hat{Y}_t) \quad (11)$$

in cui i parametri a e b vengono stimati utilizzando il metodo dei minimi quadrati.

Diversamente dagli anni passati, nel 2021 l'analisi dei coefficienti di regressione ha mostrato significative differenze tra i trimestri, causate dall'andamento molto difforme dei fenomeni indagati nei 4 periodi dell'anno. Pertanto, per le stime degli individui è stato studiato un modello di interpolazione per ogni trimestre.

Il prospetto 2 riporta 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, nel prospetto 3 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 (11).

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 3 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:

$$\hat{\varepsilon}({}_d\hat{Y}_t) = \hat{\varepsilon}({}_d\hat{Y}_t^{k-1}) + \frac{\hat{\varepsilon}({}_d\hat{Y}_t^k) - \hat{\varepsilon}({}_d\hat{Y}_t^{k-1})}{{}_d\hat{Y}_t^k - {}_d\hat{Y}_t^{k-1}} ({}_d\hat{Y}_t - {}_d\hat{Y}_t^{k-1}) \quad (12)$$

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.

5.6.2 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 , è del tipo seguente:

$$\hat{\sigma}_d(\hat{Y}) = a + b {}_d\hat{Y} + c {}_d\hat{Y}^2 \quad (13)$$

dove i parametri a , b e c vengono stimati utilizzando il metodo dei minimi quadrati, adattando il modello (13) a una nuvola di punti costituita da un consistente numero di coppie $(\hat{\sigma}_d(\hat{Y}), {}_d\hat{Y})$. In considerazione del fatto che il modello (13) è di tipo empirico, l'insieme delle stime ${}_d\hat{Y}$ 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 ai trimestri e/o all'intero anno a seconda del dominio territoriale - 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;
- numero totale di escursioni.

Diversamente dagli anni passati, nel 2021 l'analisi dei coefficienti di regressione ha mostrato significative differenze tra i trimestri, causate dall'andamento molto disforme dei fenomeni indagati nei 4 periodi dell'anno. Pertanto, anche per le stime dei viaggi e delle notti è stato studiato un modello di interpolazione per ogni trimestre, analogamente a quanto avviene solitamente per le stime delle escursioni.

I prospetti 4, 6 (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 ad ogni 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 (13) e dividendo ambo i membri del modello per il valore della stima, ${}_d\hat{Y}$, si perviene alla seguente equazione di secondo grado:

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

la cui radice positiva è espressa dalla seguente formula:

$${}_d\hat{Y} = \frac{-[b - \varepsilon({}_d\hat{Y})] - \sqrt{[b - \varepsilon({}_d\hat{Y})]^2 - 4ac}}{2c} \quad (14)$$

Nei prospetti 5 ed 7 (per le stime relative alle notti e ai viaggi) vengono riportati i valori delle stime ${}_d\hat{Y}$ ottenuti sulla base della relazione (14) in corrispondenza di alcuni valori tipici prefissati dell'errore relativo percentuale, indicati come $\varepsilon^* = \varepsilon({}_d\hat{Y})$; le stime con valori superiori a ${}_d\hat{Y}$ presentano valori dell'errore relativo inferiori a ε^* , mentre le stime che assumono valori inferiori a ${}_d\hat{Y}$ 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.

I prospetti 8 e 9 riportano rispettivamente 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 dei viaggi ed i valori delle stime ${}_d\hat{Y}^*$ ottenuti in corrispondenza di alcuni valori tipici prefissati dell'errore relativo percentuale. I viaggi vengono distinti in viaggi di lavoro, di vacanza e totali, facendo inoltre riferimento alla tipologia comunale, espressa nelle 3 seguenti modalità: comuni centro dell'area metropolitana, comuni periferici della stessa o con comunque più di 50.000 abitanti e comuni al di sotto di quest'ultima soglia.

I prospetti 10 e 11 riportano invece 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 delle escursioni totali, con riferimento ad ogni trimestre e all'intero anno, e i valori delle stime ${}_d\hat{Y}$ ottenuti in corrispondenza di alcuni valori tipici prefissati dell'errore relativo percentuale, analogamente a quanto visto nei prospetti 6 e 7.

Prospetto 2 - Valori dei coefficienti a, b e di R² delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di frequenze assolute riferite agli individui per aree territoriali - Trimestri – Anno 2021

AREE TERRITORIALI	Trimestre 1			Trimestre 2		
	a	b	R ²	a	b	R ²
Nord-Ovest	9,62469	-1,11685	0,96712	9,09657	-1,06832	0,94966
Nord-Est	8,66209	-1,05195	0,94409	8,92544	-1,07247	0,93349
Centro	9,63585	-1,14733	0,96993	9,35075	-1,11369	0,95182
Sud	9,41159	-1,10044	0,94942	8,87667	-1,05010	0,92427
Isole	9,45740	-1,11769	0,94927	9,41713	-1,11127	0,93905
Italia	10,23780	-1,15006	0,97927	9,70593	-1,10729	0,95238

AREE TERRITORIALI	Trimestre 3			Trimestre 4		
	a	b	R ²	a	b	R ²
Nord-Ovest	9,58235	-1,09866	0,93553	8,74069	-1,03709	0,93734
Nord-Est	9,17529	-1,07656	0,91296	8,64411	-1,03696	0,90849
Centro	9,24194	-1,10222	0,95430	9,95438	-1,16211	0,96233
Sud	8,66769	-1,03593	0,94763	9,30647	-1,08985	0,95653
Isole	9,31660	-1,09130	0,89432	10,06503	-1,17504	0,93736
Italia	10,02267	-1,12254	0,95011	9,77464	-1,11183	0,97275

6.

Prospetto 3 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite agli individui per aree territoriali- Trimestri – Anno 2021

STIME	Trimestre 1						Trimestre 2					
	Aree territoriali						Aree territoriali					
	Nord-Ovest	Nord-Est	Centro	Sud	Isole	Italia	Nord-Ovest	Nord-Est	Centro	Sud	Isole	Italia
20000	48,8	41,6	42,2	47,6	44,7	56,2	47,6	42,8	43,2	46,7	45,2	53,3
30000	38,9	33,6	33,4	38,0	35,6	44,5	38,4	34,5	34,5	37,7	36,1	42,5
40000	33,1	28,9	28,3	32,5	30,3	37,7	32,9	29,5	29,4	32,5	30,8	36,3
50000	29,2	25,7	24,9	28,7	26,8	33,2	29,2	26,2	25,9	28,9	27,2	32,1
60000	26,4	23,3	22,5	26,0	24,2	29,9	26,5	23,8	23,4	26,2	24,5	29,0
70000	24,2	21,5	20,6	23,9	22,2	27,4	24,4	21,9	21,5	24,2	22,5	26,6
80000	22,5	20,0	19,0	22,2	20,6	25,3	22,7	20,4	20,0	22,6	20,9	24,7
90000	21,1	18,8	17,8	20,8	19,3	23,7	21,3	19,1	18,7	21,2	19,6	23,2
100000	19,9	17,8	16,8	19,6	18,2	22,3	20,2	18,1	17,6	20,1	18,5	21,8
200000	13,5	12,4	11,3	13,4	12,3	15,0	13,9	12,5	12,0	13,9	12,6	14,9
300000	10,8	10,0	8,9	10,7	9,8	11,8	11,2	10,0	9,6	11,3	10,0	11,9
400000	9,2	8,6	7,6	9,1	8,4	10,0	9,6	8,6	8,1	9,7	8,6	10,1
500000	8,1	7,6	6,7	8,1	7,4	8,8	8,5	7,6	7,2	8,6	7,6	9,0
750000	6,4	6,2	5,3	6,5	5,9	7,0	6,9	6,1	5,7	7,0	6,0	7,2
1000000	5,5	5,3	4,5	5,5	5,0	5,9	5,9	5,3	4,9	6,0	5,1	6,1
2000000	3,7	3,7	3,0	3,8	3,4	4,0	4,1	3,6	3,3	4,2	3,5	4,2
3000000	3,0	3,0	2,4	3,0	2,7	3,2	3,3	2,9	2,7	3,4	2,8	3,3
4000000	2,5	2,6	2,0	2,6	2,3	2,7	2,8	2,5	2,3	2,9	2,4	2,8
5000000	2,2	2,3	1,8	2,3	2,0	2,3	2,5	2,2	2,0	2,6	2,1	2,5
7500000	1,8	1,8	1,4	1,8	1,6	1,9	2,0	1,8	1,6	2,1	1,7	2,0
10000000	1,5	1,6	1,2	1,6	1,4	1,6	1,7	1,5	1,4	1,8	1,4	1,7

STIME	Trimestre 3						Trimestre 4					
	Aree territoriali						Aree territoriali					
	Nord-Ovest	Nord-Est	Centro	Sud	Isole	Italia	Nord-Ovest	Nord-Est	Centro	Sud	Isole	Italia
20000	52,3	47,6	43,3	45,1	47,4	57,9	46,5	44,4	46,0	47,5	45,6	53,9
30000	41,8	38,2	34,6	36,6	38,0	46,1	37,7	36,0	36,3	38,1	35,9	43,0
40000	35,7	32,7	29,6	31,5	32,5	39,2	32,5	31,0	30,7	32,6	30,3	36,7
50000	31,6	29,0	26,1	28,1	28,8	34,6	28,9	27,6	27,0	28,9	26,6	32,4
60000	28,6	26,3	23,6	25,5	26,1	31,2	26,3	25,1	24,3	26,1	23,9	29,3
70000	26,3	24,2	21,7	23,6	24,0	28,6	24,3	23,2	22,2	24,0	21,8	26,9
80000	24,4	22,6	20,2	22,0	22,3	26,6	22,7	21,6	20,5	22,3	20,2	24,9
90000	22,9	21,2	18,9	20,7	20,9	24,9	21,3	20,3	19,2	21,0	18,8	23,4
100000	21,6	20,0	17,8	19,6	19,7	23,4	20,2	19,3	18,0	19,8	17,7	22,0
200000	14,7	13,8	12,2	13,7	13,5	15,9	14,1	13,4	12,1	13,6	11,8	15,0
300000	11,8	11,1	9,7	11,1	10,8	12,7	11,4	10,9	9,5	10,9	9,3	12,0
400000	10,1	9,5	8,3	9,6	9,3	10,8	9,8	9,4	8,1	9,3	7,8	10,2
500000	8,9	8,4	7,3	8,5	8,2	9,5	8,8	8,4	7,1	8,2	6,9	9,0
750000	7,1	6,8	5,9	6,9	6,6	7,6	7,1	6,8	5,6	6,6	5,4	7,2
1000000	6,1	5,8	5,0	5,9	5,6	6,4	6,1	5,8	4,7	5,6	4,6	6,1
2000000	4,2	4,0	3,4	4,2	3,8	4,4	4,3	4,1	3,2	3,9	3,0	4,2
3000000	3,3	3,2	2,7	3,4	3,1	3,5	3,5	3,3	2,5	3,1	2,4	3,3
4000000	2,8	2,7	2,3	2,9	2,6	3,0	3,0	2,8	2,1	2,7	2,0	2,8
5000000	2,5	2,4	2,1	2,6	2,3	2,6	2,7	2,5	1,9	2,3	1,8	2,5
7500000	2,0	2,0	1,7	2,1	1,9	2,1	2,2	2,1	1,5	1,9	1,4	2,0
10000000	1,7	1,7	1,4	1,8	1,6	1,8	1,9	1,8	1,2	1,6	1,2	1,7

Prospetto 4 - 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 2021

NOTTI – LAVORO				
AREE TERRITORIALI	a	b	c	R ²
		Gennaio-marzo		
Italia	16.029,46	0,642589895	-0,0000001599	0,96
		Aprile-giugno		
Italia	30.131,87	0,389414273	-0,0000000797	0,82
		Luglio-settembre		
Italia	35.447,59	0,354771585	-0,0000000396	0,93
		Ottobre-dicembre		
Italia	71.327,76	0,361430707	-0,0000000397	0,91
		Anno		
Nord-Ovest	54.707,37	0,440380849	-0,0000000424	0,98
Nord-Est	51.377,29	0,306761143	-0,0000000458	0,96
Centro	24.338,95	0,460046604	-0,0000001231	0,97
Sud	19.475,79	0,558922409	-0,0000001667	0,89
Isole	13.248,28	0,791077743	-0,0000002036	0,86
Italia	129.696,49	0,221317083	-0,0000000084	0,93
NOTTI – VACANZE				
AREE TERRITORIALI	a	b	c	R ²
		Gennaio-marzo		
Italia	78.429,79	0,230110336	-0,0000000159	0,92
		Aprile-giugno		
Italia	191.601,12	0,173714250	-0,0000000021	0,70
		Luglio-settembre		
Italia	487.966,11	0,0737896655	-0,0000000001	0,96
		Ottobre-dicembre		
Italia	128.744,33	0,1322728926	-0,0000000015	0,84
		Anno		
Nord-Ovest	329.414,07	0,143657020	-0,0000000003	0,98
Nord-Est	227.978,11	0,130342924	-0,0000000006	0,96
Centro	231.533,44	0,156196796	-0,0000000008	0,90
Sud	149.955,41	0,270844763	-0,0000000026	0,87
Isole	53.950,88	0,358210126	-0,0000000048	0,94
Italia	721.527,25	0,078467130	-0,0000000001	0,94
NOTTI – TOTALI				
AREE TERRITORIALI	a	b	c	R ²
		Gennaio-marzo		
Italia	97.571,86	0,231089320	-0,0000000138	0,91
		Aprile-giugno		
Italia	298.297,86	0,164384093	-0,0000000019	0,67
		Luglio-settembre		
Italia	522.442,99	0,078074569	-0,0000000002	0,93
		Ottobre-dicembre		
Italia	127.892,81	0,183137345	-0,0000000050	0,85
		Anno		
Nord-Ovest	340.969,65	0,141378530	-0,0000000003	0,97
Nord-Est	226.291,81	0,131776903	-0,0000000006	0,95
Centro	231.240,82	0,158158036	-0,0000000008	0,89
Sud	147.345,91	0,267589670	-0,0000000025	0,88
Isole	77.341,84	0,338303321	-0,0000000043	0,93
Italia	720.024,88	0,077824482	-0,0000000001	0,93

Prospetto 5 - 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 2021

NOTTI – LAVORO							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Italia	3.981.131	3.732.664	3.422.419	3.112.655	2.803.527	2.495.265	2.188.223
			Gennaio-marzo				
Italia	4.841.305	4.348.009	3.734.598	3.126.595	2.527.568	1.944.722	1.393.889
			Aprile-giugno				
Italia	8.816.273	7.818.184	6.576.004	5.343.599	4.129.090	2.951.807	1.864.891
			Luglio-settembre				
Italia	9.055.823	8.071.879	6.851.374	5.647.149	4.470.737	3.345.762	2.322.362
			Ottobre-dicembre				
			Anno				
Nord-Ovest	10.285.243	9.353.501	8.192.749	7.038.292	5.893.617	4.765.194	3.666.128
Nord-Est	6.642.289	5.794.417	4.746.421	3.720.805	2.738.193	1.845.482	1.134.962
Centro	3.710.730	3.390.690	2.992.122	2.595.874	2.203.120	1.815.924	1.438.194
Sud	3.328.353	3.091.069	2.795.098	2.500.057	2.206.306	1.914.407	1.625.291
Isole	3.852.340	3.656.825	3.412.575	3.168.520	2.924.708	2.681.205	2.438.102
Italia	25.705.286	21.084.400	15.412.987	10.011.970	5.390.903	2.575.384	1.429.688
NOTTI – VACANZE							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Italia	14.219.547	11.770.951	8.763.900	5.888.215	3.366.069	1.683.157	927.069
			Gennaio-marzo				
Italia	79.937.107	61.045.010	37.905.705	16.875.637	5.173.360	2.359.919	1.481.106
			Aprile-giugno				
Italia	529.512.587	213.376.614	17.232.679	6.338.461	3.851.926	2.763.927	2.154.626
			Luglio-settembre				
Italia	81.647.512	55.760.255	24.710.536	5.066.271	1.826.231	1.078.593	762.326
			Ottobre-dicembre				
			Anno				
Nord-Ovest	420.112.772	296.143.166	143.591.864	23.669.207	5.664.353	3.069.308	2.097.987
Nord-Est	190.643.945	128.801.153	54.194.044	8.981.541	3.180.296	1.886.305	1.337.041
Centro	189.735.268	138.828.618	76.237.693	21.705.096	4.865.809	2.419.790	1.596.304
Sud	99.728.355	84.624.886	65.810.606	47.146.646	28.902.962	12.488.394	3.823.962
Isole	72.062.447	63.822.127	53.530.337	43.254.431	33.009.025	22.834.074	12.885.438
Italia	851.793.658	373.564.604	30.087.205	9.973.497	5.913.475	4.197.989	3.253.089
NOTTI – TOTALI							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Italia	16.501.535	13.682.368	10.223.115	6.919.671	4.023.036	2.063.161	1.151.310
			Gennaio-marzo				
Italia	82.966.881	62.575.198	37.941.834	16.850.910	6.272.174	3.249.325	2.135.546
			Aprile-giugno				
Italia	451.732.073	200.212.769	20.803.142	7.154.605	4.262.105	3.030.590	2.350.323
			Luglio-settembre				
Italia	35.340.881	27.547.909	18.040.059	9.358.049	3.644.603	1.697.282	1.047.432
			Ottobre-dicembre				
			Anno				
Nord-Ovest	499.108.776	349.044.372	164.230.953	23.129.882	5.671.290	3.115.420	2.141.928
Nord-Est	197.913.778	134.380.225	57.500.454	9.403.882	3.222.407	1.895.238	1.338.573
Centro	189.131.605	139.045.028	77.416.013	23.037.540	5.045.913	2.465.535	1.615.734
Sud	104.658.496	88.597.997	68.589.505	48.738.534	29.341.595	12.049.286	3.572.023
Isole	77.375.525	68.009.166	56.316.264	44.652.175	33.046.720	21.590.143	10.698.651
Italia	929.792.404	401.444.672	29.561.854	9.876.318	5.872.539	4.174.456	3.237.313

Prospetto 6 - 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 2021

VIAGGI – LAVORO				
AREE TERRITORIALI	a	b	c	R ²
		Gennaio-marzo		
Italia	6.749,12	0,436768538	-0,0000005499	0,81
		Aprile-giugno		
Italia	13.658,85	0,216510540	-0,0000001631	0,84
		Luglio-settembre		
Italia	11.000,65	0,307958306	-0,0000002123	0,88
		Ottobre-dicembre		
Italia	13.427,05	0,245357033	-0,0000001479	0,90
		Anno		
Nord-Ovest	6.895,973	0,442487223	-0,0000004698	0,89
Nord-Est	11.389,827	0,314887036	-0,0000001683	0,91
Centro	3.455,219	0,336673942	-0,0000004215	0,96
Sud	2.350,159	0,445196050	-0,0000010400	0,94
Isole	2.971,545	0,507773627	-0,0000010232	0,94
Italia	25.558,819	0,148208717	-0,0000000257	0,91
VIAGGI – VACANZE				
AREE TERRITORIALI	a	b	c	R ²
		Gennaio-marzo		
Italia	13.742,86	0,227025809	-0,0000000969	0,91
		Aprile-giugno		
Italia	44.714,88	0,110061451	-0,0000000090	0,95
		Luglio-settembre		
Italia	83.195,09	0,072149496	-0,0000000016	0,94
		Ottobre-dicembre		
Italia	41.935,12	0,107314486	-0,0000000070	0,96
		Anno		
Nord-Ovest	48.197,916	0,117899977	-0,0000000029	0,98
Nord-Est	41.902,409	0,117965576	-0,0000000046	0,98
Centro	48.551,664	0,106308066	-0,0000000038	0,94
Sud	26.573,662	0,214493650	-0,0000000258	0,92
Isole	10.127,063	0,321315189	-0,0000000496	0,96
Italia	109.051,126	0,060117943	-0,0000000005	0,97
VIAGGI – TOTALI				
AREE TERRITORIALI	a	b	c	R ²
		Gennaio-marzo		
Italia	17.872,25	0,201183781	-0,0000000641	0,90
		Aprile-giugno		
Italia	45.331,44	0,105317689	-0,0000000077	0,95
		Luglio-settembre		
Italia	86.612,77	0,067921459	-0,0000000014	0,94
		Ottobre-dicembre		
Italia	46.678,80	0,096188939	-0,0000000056	0,96
		Anno		
Nord-Ovest	54.498,948	0,105862819	-0,0000000023	0,99
Nord-Est	48.480,521	0,104478377	-0,0000000036	0,98
Centro	51.179,059	0,097227322	-0,0000000032	0,95
Sud	29.784,103	0,184303646	-0,0000000199	0,93
Isole	14.485,025	0,268683650	-0,0000000386	0,94
Italia	118.695,516	0,053232314	-0,0000000004	0,98

Prospetto 7 - 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 2021

VIAGGI – LAVORO							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Italia	791.636	720.424	Gennaio-marzo 631.880	544.083	457.426	372.603	290.921
Italia	1.329.145	1.097.218	Aprile-giugno 816.858	557.891	344.392	204.393	130.387
Italia	1.439.196	1.256.069	Luglio-settembre 1.029.669	808.002	595.424	401.863	247.119
Italia	1.646.802	1.386.637	Ottobre-dicembre 1.068.034	763.766	491.491	286.046	168.705
			Anno				
Nord-Ovest	936.217	852.615	748.584	645.298	543.152	442.849	345.735
Nord-Est	1.848.309	1.615.889	1.327.865	1.044.579	770.516	516.585	308.118
Centro	785.488	691.996	575.760	460.688	347.834	239.821	143.957
Sud	423.776	385.840	338.582	291.582	244.981	199.035	154.256
Isole	492.376	453.786	405.679	357.772	310.153	262.968	216.474
Italia	5.549.765	4.060.973	2.304.402	962.406	409.981	236.902	163.831
VIAGGI – VACANZE							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Italia	2.301.422	1.901.565	Gennaio-marzo 1.411.445	945.021	541.056	276.273	156.008
Italia	11.577.446	7.367.203	Aprile-giugno 2.861.631	926.614	474.683	313.238	232.855
Italia	40.721.778	17.131.614	Luglio-settembre 2.603.537	1.046.477	645.583	465.857	364.212
Italia	14.233.456	8.812.193	Ottobre-dicembre 3.013.765	860.289	437.874	289.754	215.931
			Anno				
Nord-Ovest	37.635.848	24.095.448	8.197.177	1.339.370	575.366	361.982	263.572
Nord-Est	23.800.146	15.334.763	5.537.999	1.125.678	496.914	313.919	228.863
Centro	25.649.585	15.524.258	4.478.967	1.020.052	507.674	334.899	249.435
Sud	8.048.382	6.528.856	4.655.672	2.858.103	1.333.322	537.973	286.070
Isole	6.306.307	5.504.898	4.505.470	3.510.655	2.525.675	1.567.425	715.005
Italia	96.359.476	26.702.365	2.641.304	1.204.682	777.296	573.389	454.146
VIAGGI – TOTALI							
AREE TERRITORIALI	1%	5%	10%	15%	20%	25%	30%
Italia	3.073.130	2.471.239	Gennaio-marzo 1.738.779	1.061.179	537.331	270.224	163.518
Italia	12.869.031	7.945.239	Aprile-giugno 2.800.161	881.087	461.498	308.272	230.748
Italia	41.761.273	16.196.908	Luglio-settembre 2.434.576	1.036.441	651.155	473.917	372.346
Italia	15.804.169	9.098.460	Ottobre-dicembre 2.558.693	800.319	439.172	300.177	227.596
			Anno				
Nord-Ovest	42.718.170	25.506.166	6.351.601	1.164.856	571.054	375.876	279.807
Nord-Est	26.770.582	15.989.189	4.346.283	987.888	498.189	330.451	246.834
Centro	27.754.582	15.729.843	3.584.730	918.498	490.471	332.677	251.396
Sud	8.941.117	6.975.009	4.571.312	2.361.487	891.518	404.002	246.960
Isole	6.750.151	5.724.914	4.449.702	3.189.028	1.967.984	900.029	328.990
Italia	107.878.555	21.376.918	2.483.773	1.220.279	806.908	602.469	480.616

Prospetto 8 - 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 2021

TIPOLOGIE COMUNALI	VIAGGI – LAVORO			R ²
	a	b	c	
Comuni centro dell'area metropolitana	4.929,82	0,312271709	-0,000000248	0,96
Comuni periferici dell'area metropolitana o con oltre 50.000 abitanti	7.901,15	0,283739779	-0,000000191	0,94
Comuni fino a 50.000 abitanti	16.217,04	0,260513339	-0,000000103	0,87
VIAGGI – VACANZE				
Comuni centro dell'area metropolitana	40.628,95	0,110893080	-0,000000007	0,96
Comuni periferici dell'area metropolitana o con oltre 50.000 abitanti	61.270,15	0,097462046	-0,000000003	0,96
Comuni fino a 50.000 abitanti	70.687,29	0,106218232	-0,000000002	0,97
VIAGGI - TOTALI				
Comuni centro dell'area metropolitana	42.262,44	0,104276321	-0,000000006	0,9560401343
Comuni periferici dell'area metropolitana o con oltre 50.000 abitanti	62.774,95	0,093272610	-0,000000002	0,957233125
Comuni fino a 50.000 abitanti	75.676,86	0,098269639	-0,000000002	0,97032093

Prospetto 9 - 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 2021

TIPOLOGIE COMUNALI	VIAGGI – LAVORO						
	1%	5%	10%	15%	20%	25%	30%
Comuni centro dell'area metropolitana	1.233.942	1.075.160	877.863	682.875	492.656	314.122	167.806
Comuni periferici dell'area metropolitana o con oltre 50.000 abitanti	1.461.770	1.256.919	1.003.406	755.133	518.335	310.105	165.242
Comuni fino a 50.000 abitanti	2.496.697	2.119.354	1.654.516	1.204.381	787.822	451.250	249.028
VIAGGI – VACANZE							
Comuni centro dell'area metropolitana	14.338.003	9.038.516	3.239.410	891.797	440.223	287.763	213.109
Comuni periferici dell'area metropolitana o con oltre 50.000 abitanti	35.540.900	20.126.899	4.461.408	1.107.611	589.045	399.052	301.387
Comuni fino a 50.000 abitanti	49.029.482	29.429.813	7.719.195	1.510.706	742.047	488.326	363.420
VIAGGI – TOTALI							
Comuni centro dell'area metropolitana	15.611.739	9.456.454	2.975.078	830.574	429.531	286.518	214.469
Comuni periferici dell'area metropolitana o con oltre 50.000 abitanti	37.003.929	20.198.652	3.965.060	1.061.037	580.920	398.213	302.643
Comuni fino a 50.000 abitanti	53.051.800	30.039.663	6.198.108	1.398.947	734.920	496.017	373.967

Prospetto 10 - Valori dei coefficienti a, b, c e di R2 delle funzioni utilizzate per l'interpolazione degli errori campionari delle stime di totali per trimestre e anno - Stime del numero di escursioni - Anno 2021

TRIMESTRI	ESCURSIONI			R ²
	a	b	c	
I trimestre	27.619,93	0,1715507402	-0,0000000090	0,95
II trimestre	30.918,98	0,1173096937	-0,0000000029	0,99
III trimestre	55.640,45	0,1155065806	-0,0000000012	0,98
IV trimestre	38.331,95	0,1121636267	-0,0000000030	0,97
Anno	83.963,97	0,0580756653	-0,0000000005	0,98

Prospetto 11 - Valori dei totali corrispondenti ad alcuni valori tipici degli errori relativi percentuali per gruppi di stime, trimestri e anno - Stime del numero di escursioni – Anno 2021

TRIMESTRI	ESCURSIONI						
	1%	5%	10%	15%	20%	25%	30%
I trimestre	18.067.086	13.689.761	8.295.747	3.311.557	778.539	338.862	211.872
II trimestre	36.922.045	23.430.439	7.346.494	876.914	369.088	231.830	168.786
III trimestre	87.751.959	54.994.145	15.742.223	1.530.892	652.425	412.176	300.991
IV trimestre	34.649.528	21.456.791	6.166.768	943.044	430.124	276.445	203.414
Anno	107.417.455	25.107.584	1.961.029	909.312	590.493	437.032	346.841

5.6.3. Esempi di calcolo degli errori campionari

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 4 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 6 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 l'anno, ${}_c\hat{Y}_a$, si possono leggere i valori dei coefficienti a , b e c dal prospetto 4 e si può utilizzare la seguente espressione:

$$\hat{\sigma}({}_c\hat{Y}_a) = 24.338,95 + 0,460046604({}_c\hat{Y}_a) - 0,0000001231({}_c\hat{Y}_a)^2$$

Esempio 2

La valutazione degli errori di campionamento delle stime mediante un metodo approssimato si può ottenere utilizzando i dati dei prospetti 5 e 7. Il prospetto 5 fa riferimento alle stime del numero di notti; il prospetto 7 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 5, 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 17.232.679 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 5, si consideri la stima del numero di notti per lavoro, riferita al totale Italia nell'anno, pari a 14.194.577.

Dal prospetto 5, considerando le notti-lavoro ed esaminando la riga corrispondente al totale Italia per l'anno, si può osservare che il valore di tale stima risulta essere compreso tra i valori 15.412.987 e 10.011.970.

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

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

Esempio 1

L'ammontare dei turisti medi mensili nel trimestre aprile-giugno è pari a 2.362.261.

Si cerca il livello di stima che più si avvicina a 2.362.261 nella prima colonna del prospetto 3, che è pari a 2.000.000. Dalla colonna del prospetto, riferita all'Italia per il secondo trimestre, si ricava il corrispondente errore relativo percentuale della stima considerata, che è pari a 4,2%.

L'errore assoluto sarà:

$$\sigma(2.362.261) = 4,2/100 \times 2.362.261 = 99.215$$

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

$$2.362.261 - (2 \times 99.215) = 2.163.831$$

$$2.362.261 + (2 \times 99.215) = 2.560.691$$

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 2.000.000 e 3.000.000 ai quali corrispondono i valori 4,2 e 3,3 percentuali.

L'errore corrispondente a 2.362.261 è pari a:

$$\sigma(2.362.261) = 4,2 - ((4,2 - 3,3) / (3.000.000 - 2.000.000)) \times (2.362.261 - 2.000.000) = 3,87\%$$

Il corrispondente errore assoluto è $3,87/100 \times 2.362.261 = 91.419$ e l'intervallo di confidenza avrà come estremi:

$$2.362.261 - (2 \times 91.419) = 2.179.423$$

$$2.362.261 + (2 \times 91.419) = 2.545.099$$

Esempio 3

Il calcolo dell'errore può essere effettuato direttamente tramite la funzione interpolante (11):

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

Per la stima di 2.362.261 si ha:

$$\begin{aligned} \log \varepsilon^2 &= 9,70593 - (1,10729 \times \log 2.362.261) = \\ &= 9,70593 - (1,10729 \times 14,675129) = \\ &= 9,70593 - 16,2496 = -6,54369 \end{aligned}$$

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

Infatti, da $\log(\varepsilon^2) = -6,54369$ si ha:

$$\varepsilon^2 = \exp(-6,54369)$$

e dunque:

$$\varepsilon = 0,0379$$

L'errore relativo percentuale sarà dunque pari a 3,79%.

6. La diffusione dei risultati dell'indagine

Le stime provvisorie riferite all'anno precedente sono diffuse tra febbraio e aprile di ogni anno, nella Statistica Report "Viaggi e vacanze in Italia e all'estero".

Tutti i risultati definitivi dell'indagine sono diffusi attraverso il datawarehouse dell'Istituto I.Stat (<http://dati.istat.it>) nella sezione Cultura, comunicazione, viaggi --> Viaggi.

In adempimento alle richieste del Regolamento Europeo per le Statistiche del Turismo 692/2011, entro il 30 giugno di ogni anno sono trasmessi a Eurostat i dati sulla partecipazione al turismo nell'anno precedente dei residenti di 15 anni e più e i microdati sui viaggi effettuati dai residenti di 15 anni e più. Con cadenza triennale, inoltre, sono trasmessi i dati sulle escursioni.

Tutte queste informazioni sono successivamente archiviate nel data base di Eurostat, consultabile al link <http://ec.europa.eu/eurostat/web/tourism/>.

7. Glossario

Destinazione, mezzo di trasporto, tipo di alloggio, motivo della vacanza e del viaggio di lavoro: sono informazioni che vengono rilevate sulla base del concetto di “prevalenza”. In particolare, la destinazione del viaggio e il tipo di alloggio vengono associate rispettivamente alla località e 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.

Durata media del viaggio: rapporto tra il numero di notti trascorse in viaggio e il numero di viaggi.

Escursione: visita senza pernottamento effettuata fuori dal Comune dove la famiglia vive abitualmente, diretta in località italiane o estere, con una durata di almeno tre ore nel luogo di destinazione, esclusi i tempi per gli spostamenti di andata e ritorno. Sono escluse le eventuali escursioni che si effettuano durante i soggiorni di vacanza/di lavoro, poiché il luogo di partenza e di ritorno dell'escursione in questi casi non è il Comune dove vive la persona intervistata, bensì il luogo di destinazione del viaggio personale o di lavoro. Sono altresì esclusi gli spostamenti che hanno carattere di periodicità/regolarità nell'arco del mese di riferimento (per seguire un corso di studi, frequentare una palestra, fare la spesa). Le escursioni possono essere effettuate sia per motivi personali che per motivi di lavoro. Tra i motivi personali, vi sono: piacere, svago, vacanza, visita a parenti o amici, motivi religiosi o di pellegrinaggio, formazione/cultura, cure termali o trattamenti di salute, visite e cure mediche, shopping; accompagnare un familiare/parente/amico.

Organizzazione: L'organizzazione del viaggio si riferisce alla presenza o meno di una prenotazione dell'alloggio e/o del trasporto ed, eventualmente, di altri servizi acquistati in agenzia o da un tour operator. Se c'è stata una prenotazione dell'alloggio e/o del trasporto, questa può essere avvenuta in modo diretto o tramite l'agenzia/tour operator. Nel caso dell'alloggio, per prenotazione diretta si intende la prenotazione effettuata direttamente presso la struttura ricettiva, recandosi fisicamente sul posto o accedendo mediante internet al sito web della struttura alloggiativa. Nel caso del trasporto, la prenotazione diretta consiste nell'acquisto del servizio direttamente presso uffici fisici o virtuali che operano su internet, come le biglietterie ferroviarie, aeree, ecc. Per

prenotazione tramite agenzia/tour operator si intende sia quella effettuata recandosi presso gli uffici dell'agenzia/tour operator dislocati sul territorio, sia quella effettuata mediante un'agenzia/tour operator virtuale su internet (sito web dell'agenzia/tour operator, i portali quali Booking, Expedia, Tripadvisor, Trivago, Kayak o le piattaforme di prenotazione online utilizzate prevalentemente per gli alloggi privati come Airbnb, HomeAway, Scambiocasa, HomeToGo). Nel caso della prenotazione del mezzo di trasporto mediante i servizi di car rental (Hertz, Avis, ecc.), questi non sono classificati come agenzie/tour operator, pertanto la prenotazione deve essere considerata una prenotazione diretta. Per prenotazione tramite internet si intende l'utilizzo di internet per prenotare direttamente o tramite agenzia/tour operator on line l'alloggio e/o il trasporto.

Spesa del viaggio: include tutte le spese finalizzate al viaggio, sostenute prima della partenza, durante il viaggio o al ritorno, in conseguenza del viaggio stesso. E' composta dalle spese per trasporto, per l'alloggio e dalle restanti spese (pasti, acquisti vari, attività ricreative, culturali e sportive, souvenir etc..).

Spesa media per viaggio: ammontare di spesa del viaggio riferito al singolo partecipante al viaggio

Spesa media giornaliera: rapporto tra spesa media per viaggio, riferita al singolo partecipante, e durata del viaggio (numero di notti)

Tipo di luogo: l'informazione connota le caratteristiche del luogo visitato (città, mare, campagna, montagna, altro), oppure se si è trattato di una crociera.

Turista: persona che ha effettuato uno o più viaggi nel periodo di riferimento

Viaggio: spostamento realizzato, per vacanza o per ragioni di lavoro, fuori dal comune in cui si vive e che comporta almeno un pernottamento nel luogo visitato; sono esclusi i viaggi e gli spostamenti effettuati nelle località frequentate tutte le settimane (ritenuti abituali secondo la definizione di turismo), nonché i viaggi di durata superiore a un anno: in questi casi, infatti, il viaggio non costituisce flusso turistico poiché la località visitata viene associata al luogo in cui si vive. Sono altresì esclusi i viaggi effettuati per coercizione (viaggi per traslocare, viaggi per andare a votare per elezioni/referendum, spostamenti dovuti a calamità naturali, ecc.).

Viaggio abituale: viaggio con almeno un pernottamento fuori dal comune in cui si vive effettuato per vacanza o per lavoro tutte le settimane nella stessa località.

Viaggio di vacanza: viaggio svolto per motivi prevalenti di piacere, svago o riposo, per visita 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.

Viaggio per motivi di lavoro o professionali: viaggio svolto per motivi prevalenti di lavoro quali missioni, partecipazione a congressi, riunioni d'affari o esercizio di attività di rappresentanza, docenza o altre attività professionali. E' escluso l'esercizio di un'attività alle dipendenze di un soggetto residente nel luogo di destinazione (lavori stagionali, supplenze o altri lavori a tempo determinato).

8. Riferimenti bibliografici

Deville J. C., Särndal C. E., (1992), "Calibration Estimators in Survey Sampling", Journal of the American Statistical Association, vol. 87, n.418, 1992.

Di Torrice M., 2018 (a cura di) "La nuova indagine sulla domanda turistica"

ISBN 978-88-458-1964-3, Istat

Grassi D., Pannuzi N., 2015 (a cura di) "La nuova indagine sulle spese per consumi in Italia" ISBN 978-88-458-1856-1, Istat

Regolamento per le Statistiche sul Turismo 692/2011

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:192:0017:0032:EN:PDF>

Regolamento Delegato (UE) 2019/1681 della Commissione

<https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32019R1681&from=IT>

9. Contatti

"Indagine su viaggi e vacanze e informazione integrata sulla domanda turistica"

Mascia Di Torrice

e-mail: maditorr@istat.it

Barbara Dattilo

e-mail: dattilo@istat.it

Curatori dei capitoli

Il documento è stato curato da Mascia Di Torrice.

Si devono a:

Andrea Cutillo i paragrafi 5.1, 5.2, 5.3

Claudia De Vitiis, Monica Russo e Marco D. Terribili i paragrafi 3, 5.5, 5.6