

Strategia di campionamento e valutazione degli errori campionari¹

Introduzione

La *popolazione di interesse* dell'indagine è costituita dalle famiglie residenti in Italia e dagli individui di 14 anni e più che le compongono. Sono esclusi gli individui che sono membri permanenti delle convivenze. Per famiglia si intende la *famiglia di fatto*, ovvero 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 degli abbonati ad utenze di telefonia fissa; le *unità di campionamento* sono, pertanto, i numeri telefonici appartenenti a detto archivio.

L'indagine è di tipo trasversale a cadenza quinquennale e ha la finalità di fornire stime di parametri di diversa natura (totali, medie, rapporti, frequenze assolute e relative) con diversi riferimenti territoriali:

- l'intero territorio nazionale;
- le cinque ripartizioni geografiche (Nord-est, Nord-ovest, Centro, Sud e Isole);
- le regioni geografiche;
- sei aree basate sulla tipologia socio-demografica dei comuni, così definite:

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 mila abitanti;

B₂, comuni con 2.001-10 mila abitanti;

B₃, comuni con 10.001-50 mila abitanti;

B₄, comuni con oltre 50 mila abitanti.

La lista di selezione

La *base di campionamento* adottata, ovvero la lista di selezione delle unità campionarie, è l'archivio informatizzato ufficiale delle famiglie abbonate a un'utenza 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

¹ A cura di Nicoletta Cibella, Claudia De Vitiis, Alessandro Martini

censimento, la via, il numero civico, l'ampiezza del comune di appartenenza, in termini demografici e 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:

- a) *sovracopertura*, quando alcuni elementi dell'insieme L non appartengono all'insieme U;
- b) *sottocopertura*, nel caso in cui alcuni elementi di U non sono contenuti in L;
- c) *duplicazione di alcune unità*, se alcuni elementi di U sono presenti più volte in L;
- d) *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 varia misura tutte le situazioni elencate; tuttavia, i fenomeni che appaiono più rilevanti sono quelli della *sovracopertura* e della *sottocopertura*. Rientrano, infatti, nel caso della *sovracopertura* tutti i numeri telefonici che corrispondono a seconde case o ad attività professionali oppure a pubblici esercizi, eccetera; mentre rientrano nel caso della *sottocopertura* tutte le famiglie non abbonate al telefono o intestatarie di numeri riservati. Rientrano, poi, nel caso c) le famiglie intestatarie di più di un numero riferito all'abitazione principale e nel caso d) i numeri telefonici a cui corrispondono più famiglie.

L'effetto di ciascuna delle situazioni appena elencate è quella 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 le unità di U che appartengono ad L sono differenti dalle unità di U non presenti in L relativamente alle variabili d'interesse dell'indagine. È possibile in parte attenuare tale effetto distorsivo ricorrendo a opportune procedure di stima, note in letteratura come stimatori di *ponderazione vincolata*, che tengono conto di totali noti (sulla base di fonti esterne all'indagine) sulla popolazione relativi a variabili ausiliarie correlate a quelle di interesse.

Un ulteriore effetto dovuto alla *sovracopertura* è il fatto che la numerosità effettiva del campione sia minore di quella definita nella fase di progettazione del disegno e ciò comporta un aumento della variabilità delle stime. Si tratta allora di decidere se accettare una dimensione campionaria inferiore a quella prevista o, alternativamente, ricorrere a un sovradimensionamento del campione o a una lista di unità sostitutive. Nelle indagini di tipo telefonico, in cui la rilevazione è affidata ad una società esterna, si ricorre in genere alla sostituzione in quanto la procedura di quest'ultima per le indagini telefoniche è meno complessa e anche per garantire che la società svolga un numero prefissato di interviste utili.

Il disegno di campionamento

Il disegno di campionamento è a *due stadi* con stratificazione delle unità di primo stadio. Le unità di primo stadio sono gli indirizzi telefonici dell'archivio di selezione, e quindi le famiglie ad essi corrispondenti. Le unità di secondo stadio sono gli individui: per ciascuna famiglia selezionata al primo stadio si seleziona un individuo campione tra i componenti della famiglia aventi età pari o superiore a 14 anni.

Gli indirizzi telefonici sono stati stratificati per regione geografica e per tipologia di comune. La tipologia di comune non è una variabile presente sull'*Archivio telefonico informatizzato delle famiglie*, in quanto viene definita dall'Istat sulla base delle caratteristiche socio-demografiche dei comuni italiani rilevate al *Censimento della popolazione*. La stratificazione mediante tale variabile ha comportato, quindi, la necessità di integrare l'archivio con tali informazioni.

Il disegno campionario ora descritto presenta caratteristiche molto differenti rispetto ai disegni che vengono correntemente utilizzati per le indagini sulle famiglie condotte mediante intervista diretta. In primo luogo, l'indagine in esame prevede la selezione diretta degli indirizzi campione senza estrarre al

primo stadio un campione di comuni; ciò come è noto determina un aumento dell'efficienza delle stime prodotte e tale guadagno è tanto più elevato quanto più è alto il coefficiente di correlazione intra-classe, ovvero il grado di omogeneità delle variabili di interesse all'interno dei comuni. Inoltre, l'indagine prevede la selezione di un solo individuo campione da ciascuna famiglia; anche questa scelta conduce a stime tanto più efficienti quanto più è alto il coefficiente di correlazione intra-classe delle famiglie, ossia quanto più è alta la somiglianza (relativamente alle variabili d'indagine) tra gli individui all'interno delle famiglia.

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 a tutte le stime d'interesse, considerando anche il fatto che le stime vengono prodotte con diversi riferimenti territoriali. L'allocazione ottimale delle unità del campione con riferimento ad un dato tipo di dominio può risultare in contrasto con l'allocazione ottimale con riferimento ad un altro tipo di dominio. In particolare, per quanto riguarda le stime riferite all'intero territorio nazionale l'allocazione ottimale risulta vicina a quella proporzionale tra le diverse regioni; per quanto riguarda, invece, le stime riferite alle regioni, l'allocazione ottimale risulta prossima a quella che assegna a tutte le regioni un campione di uguale numerosità. È necessario quindi 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, viene definita la numerosità n complessiva del campione; analogamente all'edizione precedente dell'indagine, svolta nel 2002, la numerosità del campione è stata fissata a 60 mila unità, al fine di garantire adeguati livelli di attendibilità delle stime a livello regionale. 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 *uguale* e quella *proporzionale*. Infine, le numerosità campionarie regionali sono state ripartite tra le diverse tipologie di comune in modo proporzionale alla popolazione residente.

Nel prospetto 1 sono riportate le numerosità campionarie per regione.

Prospetto 1 - Distribuzione regionale del campione

Regioni	Indirizzi Campione
Piemonte	3.796
Valle d'Aosta	2.679
Lombardia	4.622
Bolzano	2.140
Trento	2.179
Veneto	3.241
Friuli-Venezia Giulia	2.675
Liguria	2.835
Emilia-Romagna	3.525
Toscana	3.030
Umbria	2.210
Marche	2.459
Lazio	3.137
Abruzzo	2.726
Molise	1.939
Campania	2.984
Puglia	3.132
Basilicata	2.205
Calabria	2.692
Sicilia	3.229
Sardegna	2.566
ITALIA	60.001

La selezione del campione

L'estrazione degli indirizzi campione da ciascuno strato è stata effettuata con probabilità uguali e senza reimmissione mediante tecnica di selezione sistematica.

La selezione del campione di indirizzi è avvenuta nel seguente modo:

- per ciascuna regione, con riferimento ad una data tipologia comunale, sono stati ordinati gli indirizzi dell'Archivio telefonico delle famiglie per ampiezza demografica del comune, strada, numero civico e piano;
- definito il passo di estrazione regionale, sono stati selezionati 6 punti di partenza casuali, uno per ciascuna tipologia comunale. Si è proceduto, poi, alla selezione degli indirizzi campione sulla base del passo di estrazione regionale a partire dal punto di partenza casuale;
- ogni indirizzo selezionato dall'Archivio è entrato a far parte dell'elenco base, e per ognuno di essi si sono selezionati i tre indirizzi successivi, che sono entrati 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.

La selezione casuale del componente da intervistare è stata effettuata nel seguente modo:

- a) è stato chiesto alla persona rispondente al telefono il numero dei componenti della famiglia, e per ciascuno di essi il sesso, l'età, lo stato civile e la relazione con la persona che ha risposto al telefono (che ha relazione di parentela = 01);
- b) è stata formata la graduatoria per età dei componenti, assegnando un numero progressivo ai componenti risultati intervistabili (età maggiore di 13 anni);
- c) è stato estratto un numero casuale compreso tra 1 ed il numero totale dei componenti intervistabili;
- d) è stato intervistato il componente il cui numero in graduatoria coincide con il numero casuale estratto.

Si fa presente che ai fini dell'individuazione del componente da intervistare, da effettuare per le famiglie con 2 o più individui eleggibili, l'Istat ha fornito alla ditta che si occupa della rilevazione un file costituito da 60001 record così formati: codice di regione, tipologia di comune, numero progressivo da 1 al numero di indirizzi campione nel dominio; 9 campi successivi così formati: il primo campo contiene un numero casuale pari a 1 o a 2; il secondo campo contiene un numero casuale pari a 1, 2 o 3 e così via fino all'ultimo campo che contiene un numero casuale pari a 1, 2, 3,...10. Per selezionare l'individuo da intervistare, viene utilizzato il primo campo per le famiglie di due componenti, il secondo per le famiglie di tre componenti e così via. Ciascun numero casuale è stato estratto indipendentemente dagli altri e pertanto, nell'ambito di ciascun campo k ($k=1,...,9$), ogni numero i ($i=1,..., k+1$) compare lo stesso numero di volte nei 60001 record. In tal modo, ciascun componente delle famiglie di ampiezza $k+1$ ($k=1,...,9$) ha la stessa probabilità di essere selezionato.

Qualora non sia possibile intervistare il componente selezionato, si procede alla sua sostituzione assegnando un nuovo numero progressivo ai componenti eleggibili e si utilizza il componente indicato per lo stesso record in corrispondenza della dimensione familiare immediatamente inferiore.

Il calcolo delle stime

Le stime prodotte dall'indagine sono stime del numero di vittime di eventi delittuosi subiti dagli individui e del numero di vittime e di eventi delittuosi subiti dalle famiglie, nonché le caratteristiche strutturali degli individui e delle famiglie che hanno subito tali fatti delittuosi.

Le stime sono ottenute mediante uno stimatore di ponderazione vincolata, 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 ad ogni unità campionaria un peso che indica il numero di unità della popolazione rappresentate dall'unità medesima. Se, ad esempio, ad 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; h, indice di strato; j, indice di famiglia; q indice di individuo all'interno della famiglia j; y, generica variabile oggetto di indagine; Y_{hjq} valore di y osservato sull'individuo q della famiglia j dello strato h; Q_{hj} , numero di individui di 14 anni e oltre appartenenti alla famiglia j dello strato h; M_h , numero di famiglie residenti nello strato h; m_h , campione di famiglie nello strato h; p_h , numero di individui campione nello strato h (dal momento che si intervista un unico individuo in ciascuna famiglia campione si ha $p_h = m_h$); H_d , numero di strati nel dominio d.

Ipotizziamo di voler stimare, con riferimento ad un generico dominio d (ad esempio una regione geografica) il totale della variabile y oggetto di indagine (ad esempio il numero totale di scippi subiti dagli individui di 14 anni e oltre), espresso dalla seguente relazione:

$${}_dY = \sum_{h=1}^{H_d} \sum_{j=1}^{M_h} \sum_{q=1}^{Q_{hj}} Y_{hjq} \quad (1)$$

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

$${}_d\hat{Y} = \sum_{h=1}^{H_d} \hat{Y}_h = \sum_{h=1}^{H_d} \sum_{j=1}^{m_h} Y_{hj} \cdot W_{hj},$$

in cui Y_{hj} e W_{hj} rappresentano rispettivamente il valore assunto dalla variabile y e il peso finale da attribuire all'individuo campione della famiglia j dello strato h.

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.

Per questa indagine sono stati calcolati due sistemi di pesi differenti per la produzione delle stime riferite agli individui e alle famiglie. Per i pesi individuali la popolazione di riferimento è costituita dalla popolazione di 14 anni e più al netto delle convivenze e i totali noti imposti a livello regionale sono i seguenti:

- a. popolazione per sesso e classi di età (14-24, 25-34, 35-44, 45-54, 55-64, 65 e oltre);
- b. popolazione per tipologia comunale (aree A1, A2, B1, B2, B3, B4 definite nel paragrafo 1);
- c. popolazione per due classi di età (14-54, 55 e oltre) e titolo di studio (fino a licenza elementare, licenza media o diploma superiore fino a tre anni, titolo superiore);
- d. popolazione per dimensione familiare (famiglie monocomponenti per sesso ed età (14-64, 65 e oltre), 2 componenti, 3-4 componenti, 5 e più componenti).

Per i pesi a livello familiare la popolazione di riferimento è costituita invece dalla popolazione residente al netto delle convivenze e i totali noti sono i seguenti:

- a livello regionale

- a. popolazione per sesso e classi di età (0-13, 14-24, 25-34, 35-44, 45-54, 55-64, 65 e oltre);

- b. popolazione per tipologia comunale (aree A1, A2, B1, B2, B3, B4 definite nel paragrafo 1);
- a livello ripartizionale
 - a. popolazione per due classi di età (14-54, 55 e oltre) e titolo di studio² (fino a licenza elementare, licenza media o diploma superiore fino a tre anni, titolo superiore);
 - b. popolazione per dimensione familiare (famiglie monocomponenti per sesso ed età - 14-64, 65 e oltre - , 2 componenti, 3-4 componenti, 5 e più componenti).

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 3. 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, ed i vincoli sono definiti dalla condizione che le stime campionarie dei totali di popolazione sopra definiti coincidano con i valori noti degli stessi.

E' utile osservare che i vincoli c. e d. (relativi al titolo di studio e alla dimensione delle famiglie) sono stati introdotti nonostante il fatto che non si basino su totali noti da fonte censuaria o anagrafica, ma solo su stime prodotte da un'altra indagine campionaria. Si è comunque ritenuto opportuno utilizzarli per correggere, almeno in parte, la distorsione dovuta alla sottocopertura della lista di selezione, considerando anche il fatto che l'indagine a cui si ricorre è basata su un campione di dimensione sufficiente per garantire stime affidabili.

Valutazione del livello di precisione delle stime

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

Indicando con $\hat{Var}(\hat{Y}_d)$ la varianza della stima $\hat{Y}_d$, riferita al dominio d, la stima dell'errore di campionamento assoluto di $\hat{Y}_d$ si può ottenere mediante la seguente espressione:

$$\hat{\sigma}(\hat{Y}_d) = \sqrt{\hat{Var}(\hat{Y}_d)} \quad (3)$$

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

$$\hat{\varepsilon}(\hat{Y}_d) = \frac{\sqrt{\hat{Var}(\hat{Y}_d)}}{\hat{Y}_d} \quad (4)$$

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

$$\hat{Var}(\hat{Y}_d) = \sum_{h=1}^{H_d} \hat{Var}(\hat{Y}_h) = \sum_{h=1}^{H_d} \frac{m_h}{m_h - 1} \sum_{j=1}^{m_h} \frac{(\hat{Y}_{hj} - \hat{Y}_h)^2}{m_h - 1} \quad (5)$$

² I totali noti relativi al titolo di studio e alla dimensione familiare derivano da stime dell'indagine Rilevazione Continua sulle Forze di Lavoro, condotta con tecnica di rilevazione diretta (faccia a faccia)

dove

$$\hat{Y}_{hj} = Y_{hj} W_{hj} \quad e \quad \hat{\bar{Y}}_h = \frac{1}{m_h} \sum_{j=1}^{m_h} \hat{Y}_{hj}.$$

Gli errori campionari delle espressioni (3) e (4), consentono di valutare il grado di precisione delle stime; inoltre, l'errore assoluto permette di costruire l'intervallo di confidenza, che, con una certa probabilità, contiene il parametro d'interesse. Con riferimento alla generica stima $\hat{Y}$ tale intervallo assume la seguente forma:

$$\Pr\{\hat{Y} - k \hat{\sigma}(\hat{Y}) \leq Y \leq \hat{Y} + k \hat{\sigma}(\hat{Y})\} = P \quad (6)$$

Nella (6) il valore di k dipende dal valore fissato per la probabilità P ; ad esempio, per $P=0,95$ si ha $k=1,96$.

Presentazione sintetica degli errori campionari

Ad ogni stima ${}_d\hat{Y}$ è associato un errore campionario relativo $\hat{\varepsilon}({}_d\hat{Y})$; quindi, per consentire un uso corretto delle stime fornite dall'indagine, sarebbe necessario presentare, per ogni stima pubblicata, anche il corrispondente errore di campionamento relativo.

Ciò, tuttavia, non è possibile, sia per limiti di tempo e di costi di elaborazione, sia perché le tavole della pubblicazione risulterebbero eccessivamente appesantite e di non agevole consultazione per l'utente finale. Inoltre, non sarebbero in ogni caso disponibili gli errori delle stime non pubblicate, che l'utente può ricavare in modo autonomo.

Per questi motivi, generalmente, si ricorre ad una presentazione sintetica degli errori relativi, basata sul *metodo dei modelli regressivi*. Tale metodo si fonda sulla determinazione di una funzione matematica che mette in relazione ciascuna stima con il proprio errore relativo.

L'approccio utilizzato per la costruzione di questi modelli è diverso a seconda che si tratti di variabili qualitative o quantitative. Infatti, nel caso delle stime di frequenze assolute (o relative) riferite alle modalità di variabili qualitative, è possibile utilizzare dei modelli che hanno un fondamento teorico e secondo cui gli errori relativi delle stime di frequenze assolute sono funzione decrescente dei valori delle stime stesse. Per quanto riguarda, invece, le stime di totali di variabili quantitative, si tratta di un problema di notevole complessità, che può essere risolto in maniera empirica adattando diversi modelli regressivi ai dati osservati e scegliendo tra i modelli stimati quello che conduce ad un R^2 maggiore.

Per calcolare gli errori di campionamento è stato utilizzato un software generalizzato, messo a punto presso l'Istat, che consente di calcolare gli errori campionari e gli intervalli di confidenza e permette di costruire dei modelli regressivi per la presentazione sintetica degli errori di campionamento.

Presentazione sintetica degli errori campionari per stime di frequenze

Il modello utilizzato per le stime di frequenze assolute, con riferimento al generico dominio d , è il seguente:

$$\log \hat{\varepsilon}^2({}_d\hat{Y}) = a + b \log({}_d\hat{Y}) \quad (7)$$

dove i parametri a e b vengono stimati mediante il metodo dei minimi quadrati.

Il prospetto 2 riporta i valori dei coefficienti a e b e dell'indice di determinazione R^2 del modello utilizzato per l'interpolazione degli errori campionari delle stime di frequenze riferite alle famiglie e alle persone, per aree territoriali.

Sulla base delle informazioni contenute nel suddetto prospetto è possibile calcolare l'errore relativo di una determinata stima di frequenza assoluta ${}_d\hat{Y}^*$, riferita ai diversi domini, mediante la formula:

$$\hat{\varepsilon}({}_d\hat{Y}^*) = \sqrt{\exp(a + b \log({}_d\hat{Y}^*))} \quad (8)$$

e costruire l'intervallo di confidenza al 95% come:

$$\left\{ {}_d\hat{Y}^* - 1,96 \cdot \hat{\varepsilon}({}_d\hat{Y}^*) \cdot {}_d\hat{Y}^* ; {}_d\hat{Y}^* + 1,96 \cdot \hat{\varepsilon}({}_d\hat{Y}^*) \cdot {}_d\hat{Y}^* \right\}.$$

Allo scopo di facilitare il calcolo degli errori campionari, nei prospetti 3 e 4 sono riportati, gli errori relativi percentuali corrispondenti a valori crescenti di stime di frequenze assolute riferite, rispettivamente, alle famiglie e alle persone, calcolati introducendo nella (8) i valori di a e b riportati nel prospetto 2.

Le informazioni contenute in tali prospetti consentono di calcolare l'errore relativo di una generica stima di frequenza assoluta mediante due procedimenti di facile applicazione che, tuttavia, conducono a risultati meno precisi di quelli ottenibili applicando direttamente la formula (8).

Il primo metodo consiste nell'approssimare l'errore relativo della stima di interesse ${}_d\hat{Y}^*$ con quello, riportato nei prospetti, corrispondente al livello di stima che più si avvicina a ${}_d\hat{Y}^*$.

Il secondo metodo, più preciso del primo, si basa sull'uso di una formula di interpolazione lineare per il calcolo degli errori di stime non comprese tra i valori forniti nei prospetti. In tal caso, l'errore campionario della stima ${}_d\hat{Y}^*$, si ricava mediante l'espressione:

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

dove ${}_d\hat{Y}^{k-1}$ e ${}_d\hat{Y}^k$ sono i valori delle stime entro i quali è compresa la stima ${}_d\hat{Y}^*$, mentre $\hat{\varepsilon}({}_d\hat{Y}^{k-1})$ e $\hat{\varepsilon}({}_d\hat{Y}^k)$ sono i corrispondenti errori relativi.

Prospetto 2 - Valori dei coefficienti a, b e dell'indice di determinazione R2 (%) delle funzioni utilizzate per le interpolazioni degli errori campionari delle stime di frequenze assolute riferite alle FAMIGLIE e alle PERSONE per aree territoriali

	Famiglie			Individui		
	A	B	R2 (%)	A	B	R2 (%)
ITALIA	10.7165	-1.295	96,7	10.5176	-1.228	96,8
RIPARTIZIONI GEOGRAFICHE						
Nord	10.5035	-1.290	97,0	10.7420	-1.257	95,7
Nord-Ovest	10.3528	-1.278	96,9	10.4517	-1.234	96,1
Nord-Est	10.0975	-1.299	97,5	10.0409	-1.239	95,9
Centro	9.7432	-1.271	96,7	9.8171	-1.208	97,5
Sud	9.0926	-1.217	92,2	9.9441	-1.208	97,0
Isole	9.4592	-1.260	94,6	9.1921	-1.162	96,5
Sud-Isole	9.6628	-1.249	94,2	9.7260	-1.187	96,9
TIPI DI COMUNE						
A1	9.1804	-1.214	96,3	9.9696	-1.206	97,3
A2	9.8651	-1.267	95,9	9.9908	-1.210	96,7
B1	7.9002	-1.157	94,0	9.3875	-1.208	97,1
B2	10.4195	-1.308	96,7	9.7607	-1.198	97,4
B3	11.4475	-1.376	96,3	10.1259	-1.225	96,7
B4	9.9387	-1.284	96,9	9.6389	-1.202	95,9
REGIONE						
Piemonte	9.1582	-1.234	94,9	10.0679	-1.255	96,0
Valle d'Aosta	6.0047	-1.329	97,0	5.8856	-1.244	94,8
Lombardia	10.5577	-1.289	97,0	10.7117	-1.246	95,9
Bolzano	8.8629	-1.425	95,9	7.7067	-1.250	96,0
Trento	8.3119	-1.360	93,3	7.4191	-1.203	96,7
Veneto	9.2968	-1.246	97,0	9.4986	-1.193	97,1
Friuli-Venezia Giulia	7.5276	-1.203	96,4	8.3232	-1.212	97,3
Liguria	7.9238	-1.208	96,2	9.2052	-1.263	96,2
Emilia-Romagna	10.3398	-1.328	97,5	9.6193	-1.217	95,7
Toscana	9.4493	-1.264	95,3	9.9744	-1.252	96,3
Umbria	8.6378	-1.331	95,2	8.1623	-1.220	96,1
Marche	8.8236	-1.312	97,5	8.5928	-1.208	96,7
Lazio	9.4169	-1.249	96,9	9.8544	-1.203	97,2
Abruzzo	8.3716	-1.289	94,9	8.4019	-1.212	96,9
Molise	6.5399	-1.261	95,6	6.6821	-1.176	96,7
Campania	9.0337	-1.189	92,0	9.6454	-1.170	97,1
Puglia	8.9900	-1.252	96,4	9.2664	-1.180	96,6
Basilicata	7.7296	-1.327	95,5	7.2030	-1.175	95,4
Calabria	8.0159	-1.208	93,0	8.5850	-1.178	97,1
Sicilia	9.6088	-1.267	94,1	9.4231	-1.172	96,3
Sardegna	8.8179	-1.300	97,2	8.0458	-1.139	95,5

(a) Italia nord-occidentale: Piemonte, Valle d'Aosta, Lombardia, Liguria; Italia nord-orientale: Bolzano, Trento, Veneto, Friuli-Venezia Giulia, Emilia Romagna; Italia centrale: Toscana, Umbria, Marche, Lazio; Italia meridionale: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria; Italia insulare: Sicilia, Sardegna.

(b) Comuni tipo A1: Area urbana centro; Tipo A2: Area urbana periferia; Tipo B1: comuni fino a 2 mila abitanti; Tipo B2: da 2.001 a 10 mila abitanti; Tipo B3: da 10.001 a 50 mila abitanti; Tipo B4: oltre 50 mila abitanti.

Prospetto 2 segue - Valori dei coefficienti a, b e dell'indice di determinazione R2 (%) delle funzioni utilizzate per le interpolazioni degli errori campionari delle stime di frequenze assolute riferite alle FAMIGLIE e alle PERSONE per aree territoriali

	Famiglie			Persone		
	A	B	R2 (%)	A	B	R2 (%)
Nord-ovest						
A1	8.9289	-1.209	96,1	10.0559	-1.241	96,0
A2	10.2474	-1.300	95,0	10.3546	-1.255	95,1
B1	9.0462	-1.247	95,9	9.9078	-1.249	96,4
B2	9.7440	-1.260	96,9	10.3279	-1.245	95,7
B3	10.1734	-1.294	96,0	9.6122	-1.201	94,6
B4	8.5948	-1.181	95,6	8.5470	-1.126	93,8
Nord-est						
A1	8.8260	-1.236	95,1	8.1529	-1.118	94,2
A2	8.1664	-1.186	95,8	8.3529	-1.131	95,0
B1	7.0302	-1.194	94,4	8.1665	-1.186	93,7
B2	9.5774	-1.289	97,3	8.8061	-1.163	96,5
B3	9.5293	-1.284	97,7	9.6409	-1.226	96,5
B4	10.4169	-1.355	97,9	9.2170	-1.200	95,5
Centro						
A1	9.4720	-1.260	97,2	10.2347	-1.249	96,3
A2	8.9497	-1.251	96,9	8.7054	-1.140	96,3
B1	7.7233	-1.240	96,1	9.0983	-1.229	96,0
B2	9.4973	-1.294	92,1	9.2323	-1.209	96,1
B3	9.7414	-1.326	97,6	9.3851	-1.210	96,8
B4	9.5699	-1.302	96,6	9.8117	-1.261	95,3
Sud						
A1	7.3865	-1.091	93,5	9.1493	-1.146	97,0
A2	7.6429	-1.121	92,7	9.7781	-1.206	96,1
B1	6.3835	-1.047	84,8	7.5823	-1.094	96,0
B2	7.6215	-1.136	93,2	9.1316	-1.183	96,6
B3	8.7861	-1.206	89,4	9.7658	-1.213	96,5
B4	9.0113	-1.256	93,1	8.9339	-1.164	96,0
Isole						
A1	8.3742	-1.209	92,7	9.4484	-1.191	96,9
A2	8.7013	-1.252	94,9	8.6613	-1.153	93,6
B1	7.0195	-1.200	89,7	8.2031	-1.157	95,0
B2	8.2402	-1.199	92,7	8.5490	-1.140	96,2
B3	11.9967	-1.492	96,4	8.7413	-1.142	95,8
B4	7.6987	-1.155	94,1	8.7660	-1.146	95,2

(a) Italia nord-occidentale: Piemonte, Valle d'Aosta, Lombardia, Liguria; Italia nord-orientale: Bolzano, Trento, Veneto, Friuli-Venezia Giulia, Emilia Romagna; Italia centrale: Toscana, Umbria, Marche, Lazio; Italia meridionale: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria; Italia insulare: Sicilia, Sardegna.

(b) Comuni tipo A1: Area urbana centro; Tipo A2: Area urbana periferia; Tipo B1: comuni fino a 2 mila abitanti; Tipo B2: da 2.001 a 10 mila abitanti; Tipo B3: da 10.001 a 50 mila abitanti; Tipo B4: oltre 50 mila abitanti.

Prospetto 3 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle FAMIGLIE per aree territoriali

STIME	Italia	Nord-Ovest	Nord-Est	Centro	Sud	Isole	A1	A2	B1	B2	B3	B4
20.000	32,5	28,8	25,1	25,8	24,5	25,0	27,2	25,6	17,3	26,7	28,1	25,0
30.000	25,1	22,4	19,3	19,8	18,9	19,2	21,1	19,8	13,5	20,6	21,6	19,3
40.000	20,9	18,8	16,0	16,4	15,8	16,0	17,6	16,5	11,3	17,1	17,9	16,1
50.000	18,2	16,3	13,9	14,2	13,7	13,8	15,3	14,4	9,9	14,8	15,5	13,9
60.000	16,2	14,6	12,3	12,6	12,2	12,3	13,6	12,8	8,8	13,2	13,7	12,4
70.000	14,7	13,3	11,2	11,4	11,0	11,1	12,4	11,6	8,0	11,9	12,4	11,2
80.000	13,5	12,2	10,2	10,4	10,1	10,2	11,4	10,7	7,4	10,9	11,4	10,3
90.000	12,5	11,4	9,5	9,6	9,4	9,5	10,5	9,9	6,9	10,1	10,6	9,6
100.000	11,7	10,6	8,9	9,0	8,8	8,8	9,9	9,3	6,5	9,5	9,9	9,0
200.000	7,5	6,9	5,7	5,7	5,6	5,7	6,4	6,0	4,2	6,1	6,3	5,8
300.000	5,8	5,4	4,4	4,4	4,3	4,4	4,9	4,6	3,3	4,7	4,8	4,4
400.000	4,8	4,5	3,6	3,6	3,6	3,6	4,1	3,9	2,8	3,9	4,0	3,7
500.000	4,2	3,9	3,1	3,1	3,1	3,1	3,6	3,4	2,4	3,4	3,5	3,2
750.000	3,2	3,1	2,4	2,4	2,4	2,4	2,8	2,6	1,9	2,6	2,7	2,5
1.000.000	2,7	2,6	2,0	2,0	2,0	2,0	2,3	2,2	1,6	2,1	2,2	2,1
2.000.000	1,7	1,7	1,3	1,3	1,3	1,3	1,5	1,4	1,0	1,4	1,4	1,3
3.000.000	1,3	1,3	1,0	1,0	1,0	1,0	1,2	1,1	0,8	1,1	1,1	1,0
4.000.000	1,1	1,1	0,8	0,8	0,8	0,8	1,0	0,9	0,7	0,9	0,9	0,8
5.000.000	1,0	0,9	0,7	0,7	0,7	0,7	0,8	0,8	0,6	0,8	0,8	0,7
7.500.000	0,7	0,7	0,5	0,5	0,6	0,5	0,6	0,6	0,5	0,6	0,6	0,6
10.000.000	0,6	0,6	0,4	0,4	0,5	0,5	0,5	0,5	0,4	0,5	0,5	0,5
15.000.000	0,5	0,5	0,3	0,3	0,4	0,3	0,4	0,4	0,3	0,4	0,4	0,4
20.000.000	0,4	0,4	0,3	0,3	0,3	0,3	0,3	0,3	0,2	0,3	0,3	0,3

Prospetto 3 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle FAMIGLIE per aree territoriali

STIME	Piemonte	Valle D'Aosta	Lombardia	Bolzano	Trento	Veneto	Friuli-Venezia Giulia	Liguria	Emilia-Romagna	Toscana	Umbria
20.000	21,3	2,7	30,3	6,3	6,9	23,6	11,6	13,5	23,0	20,6	9,4
30.000	16,6	2,0	23,6	4,8	5,3	18,2	8,9	10,5	17,7	15,9	7,2
40.000	13,8	1,7	19,7	4,0	4,3	15,2	7,4	8,7	14,7	13,2	5,9
50.000	12,0	1,5	17,2	3,4	3,7	13,2	6,4	7,6	12,7	11,5	5,1
60.000	10,7	1,3	15,4	3,0	3,3	11,7	5,7	6,7	11,3	10,2	4,5
70.000	9,7	1,2	14,0	2,7	3,0	10,6	5,2	6,1	10,2	9,3	4,0
80.000	9,0	1,1	12,9	2,5	2,7	9,8	4,8	5,6	9,4	8,5	3,7
90.000	8,3	1,0	12,0	2,3	2,5	9,0	4,4	5,2	8,7	7,9	3,4
100.000	7,8	0,9	11,2	2,2	2,3	8,5	4,1	4,9	8,1	7,4	3,2
200.000	5,0	0,6	7,3	1,4	1,4	5,4	2,6	3,2	5,2	4,8	2,0
300.000	3,9	0,5	5,7	1,0	1,1	4,2	2,0	2,4	4,0	3,7	1,5
400.000	3,3	0,4	4,8	0,9	0,9	3,5	1,7	2,0	3,3	3,1	1,2
500.000	2,8	0,3	4,2	0,7	0,8	3,0	1,5	1,8	2,9	2,7	1,1
600.000	2,5	0,3	3,7	0,6	0,7	2,7	1,3	1,6	2,6	2,4	1,0
700.000	2,3	0,3	3,4	0,6	0,6	2,4	1,2	1,4	2,3	2,1	0,9
800.000	2,1	0,2	3,1	0,5	0,6	2,2	1,1	1,3	2,1	2,0	0,8
900.000	2,0	0,2	2,9	0,5	0,5	2,1	1,0	1,2	2,0	1,8	0,7
1.000.000	1,8	0,2	2,7	0,5	0,5	1,9	0,9	1,1	1,8	1,7	0,7
1.250.000	1,6	0,2	2,4	0,4	0,4	1,7	0,8	1,0	1,6	1,5	0,6
1.500.000	1,4	0,2	2,1	0,4	0,4	1,5	0,7	0,9	1,4	1,3	0,5
1.750.000	1,3	0,2	1,9	0,3	0,3	1,4	0,7	0,8	1,3	1,2	0,5
2.000.000	1,2	0,1	1,8	0,3	0,3	1,2	0,6	0,7	1,2	1,1	0,4
2.500.000	1,0	0,1	1,5	0,2	0,3	1,1	0,5	0,6	1,0	1,0	0,4

Prospetto 3 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle FAMIGLIE per aree territoriali

STIME	Marche	Lazio	Abruzzo	Molise	Campania	Puglia	Basilicata	Calabria	Sicilia	Sardegna
20.000	12,8	26,0	12,1	5,3	26,2	21,5	7,2	14,4	26,3	14,9
30.000	9,7	19,9	9,1	4,1	20,2	16,6	5,5	11,1	20,1	11,3
40.000	8,0	16,5	7,5	3,4	16,7	13,7	4,6	9,2	16,6	9,3
50.000	6,9	14,3	6,4	2,9	14,5	11,9	4,0	8,0	14,4	8,0
60.000	6,1	12,6	5,7	2,6	12,9	10,5	3,5	7,1	12,7	7,1
70.000	5,5	11,4	5,1	2,4	11,6	9,5	3,2	6,4	11,5	6,4
80.000	5,0	10,5	4,7	2,2	10,7	8,7	2,9	5,9	10,5	5,8
90.000	4,6	9,7	4,3	2,0	9,9	8,1	2,7	5,5	9,7	5,4
100.000	4,3	9,1	4,0	1,9	9,2	7,6	2,5	5,1	9,1	5,0
200.000	2,7	5,8	2,5	1,2	5,9	4,8	1,6	3,3	5,7	3,1
300.000	2,0	4,4	1,9	0,9	4,5	3,7	1,2	2,5	4,4	2,4
400.000	1,7	3,7	1,5	0,8	3,8	3,1	1,0	2,1	3,6	1,9
500.000	1,4	3,2	1,3	0,7	3,3	2,7	0,9	1,8	3,1	1,7
600.000	1,3	2,8	1,2	0,6	2,9	2,4	0,8	1,6	2,8	1,5
700.000	1,2	2,5	1,0	0,5	2,6	2,1	0,7	1,5	2,5	1,3
800.000	1,1	2,3	1,0	0,5	2,4	2,0	0,6	1,3	2,3	1,2
900.000	1,0	2,1	0,9	0,5	2,2	1,8	0,6	1,2	2,1	1,1
1.000.000	0,9	2,0	0,8	0,4	2,1	1,7	0,6	1,2	2,0	1,0
1.250.000	0,8	1,7	0,7	0,4	1,8	1,5	0,5	1,0	1,7	0,9
1.500.000	0,7	1,5	0,6	0,3	1,6	1,3	0,4	0,9	1,5	0,8
1.750.000	0,6	1,4	0,6	0,3	1,4	1,2	0,4	0,8	1,4	0,7
2.000.000	0,6	1,3	0,5	0,3	1,3	1,1	0,4	0,7	1,3	0,7
2.500.000	0,5	1,1	0,4	0,2	1,1	0,9	0,3	0,6	1,1	0,6

Prospetto 3 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle FAMIGLIE per aree territoriali

STIME	Nord-ovest						Nord-est					
	A1	A2	B1	B2	B3	B4	A1	A2	B1	B2	B3	B4
20.000	24,1	25,2	18,2	24,6	26,2	20,5	19,2	17,8	9,1	21,1	20,8	21,3
30.000	18,7	19,3	14,3	19,1	20,1	16,3	14,7	13,7	7,0	16,2	16,0	16,3
40.000	15,6	16,0	12,0	15,9	16,7	13,8	12,1	11,3	5,8	13,3	13,3	13,4
50.000	13,6	13,8	10,5	13,8	14,4	12,1	10,4	9,8	5,1	11,5	11,5	11,6
60.000	12,1	12,3	9,4	12,4	12,8	10,9	9,2	8,7	4,5	10,2	10,2	10,3
70.000	11,0	11,1	8,5	11,2	11,6	10,0	8,3	7,9	4,1	9,2	9,2	9,3
80.000	10,1	10,2	7,9	10,3	10,6	9,3	7,6	7,2	3,7	8,4	8,4	8,5
90.000	9,4	9,4	7,3	9,6	9,8	8,7	7,1	6,7	3,5	7,8	7,8	7,9
100.000	8,8	8,8	6,9	9,0	9,2	8,2	6,6	6,3	3,3	7,3	7,3	7,3
200.000	5,7	5,6	4,5	5,8	5,8	5,5	4,2	4,0	2,1	4,6	4,7	4,6
300.000	4,4	4,3	3,5	4,5	4,5	4,3	3,2	3,1	1,6	3,5	3,6	3,5
400.000	3,7	3,6	3,0	3,8	3,7	3,7	2,6	2,5	1,3	2,9	3,0	2,9
500.000	3,2	3,1	2,6	3,3	3,2	3,2	2,3	2,2	1,2	2,5	2,6	2,5
600.000	2,9	2,7	2,3	2,9	2,8	2,9	2,0	2,0	1,0	2,2	2,3	2,2
700.000	2,6	2,5	2,1	2,7	2,6	2,7	1,8	1,8	0,9	2,0	2,1	2,0
800.000	2,4	2,3	2,0	2,4	2,4	2,5	1,7	1,6	0,9	1,8	1,9	1,8
900.000	2,2	2,1	1,8	2,3	2,2	2,3	1,5	1,5	0,8	1,7	1,8	1,7
1.000.000	2,1	2,0	1,7	2,1	2,0	2,2	1,4	1,4	0,8	1,6	1,6	1,6
1.250.000	1,8	1,7	1,5	1,8	1,8	1,9	1,2	1,2	0,7	1,4	1,4	1,4
1.500.000	1,6	1,5	1,3	1,6	1,6	1,7	1,1	1,1	0,6	1,2	1,3	1,2
1.750.000	1,5	1,4	1,2	1,5	1,4	1,6	1,0	1,0	0,5	1,1	1,1	1,1
2.000.000	1,3	1,2	1,1	1,4	1,3	1,5	0,9	0,9	0,5	1,0	1,0	1,0
2.500.000	1,2	1,1	1,0	1,2	1,1	1,3	0,8	0,8	0,4	0,9	0,9	0,9

Prospetto 3 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle FAMIGLIE per aree territoriali

STIME	Centro						Sud					
	A1	A2	B1	B2	B3	B4	A1	A2	B1	B2	B3	B4
20.000	23,7	19,6	9,9	18,8	18,2	18,2	19,3	17,9	12,5	19,3	21,7	18,5
30.000	18,3	15,1	7,6	14,3	13,9	13,9	14,9	14,0	9,9	14,8	16,7	14,3
40.000	15,2	12,6	6,3	11,8	11,5	11,6	12,5	11,8	8,4	12,3	13,9	11,9
50.000	13,2	10,9	5,4	10,2	9,9	10,0	10,8	10,3	7,3	10,6	12,0	10,3
60.000	11,8	9,7	4,8	9,0	8,8	8,9	9,6	9,3	6,6	9,4	10,7	9,2
70.000	10,7	8,7	4,4	8,1	7,9	8,0	8,8	8,4	6,0	8,5	9,7	8,4
80.000	9,8	8,0	4,0	7,5	7,3	7,4	8,0	7,8	5,6	7,8	8,9	7,7
90.000	9,1	7,4	3,7	6,9	6,7	6,8	7,5	7,3	5,2	7,2	8,3	7,1
100.000	8,5	7,0	3,5	6,4	6,3	6,4	7,0	6,8	4,9	6,8	7,7	6,7
200.000	5,5	4,4	2,2	4,0	4,0	4,0	4,5	4,5	3,3	4,3	4,9	4,3
300.000	4,2	3,4	1,7	3,1	3,0	3,1	3,5	3,5	2,6	3,3	3,8	3,3
400.000	3,5	2,8	1,4	2,5	2,5	2,6	2,9	3,0	2,2	2,7	3,2	2,8
500.000	3,0	2,5	1,2	2,2	2,2	2,2	2,5	2,6	1,9	2,4	2,7	2,4
600.000	2,7	2,2	1,1	1,9	1,9	2,0	2,3	2,3	1,7	2,1	2,4	2,1
700.000	2,5	2,0	1,0	1,8	1,7	1,8	2,0	2,1	1,6	1,9	2,2	1,9
800.000	2,3	1,8	0,9	1,6	1,6	1,6	1,9	2,0	1,5	1,7	2,0	1,8
900.000	2,1	1,7	0,8	1,5	1,5	1,5	1,7	1,8	1,4	1,6	1,9	1,6
1.000.000	2,0	1,6	0,8	1,4	1,4	1,4	1,6	1,7	1,3	1,5	1,8	1,5
1.250.000	1,7	1,4	0,7	1,2	1,2	1,2	1,4	1,5	1,1	1,3	1,5	1,3
1.500.000	1,5	1,2	0,6	1,1	1,0	1,1	1,3	1,4	1,0	1,2	1,4	1,2
1.750.000	1,4	1,1	0,5	1,0	0,9	1,0	1,1	1,2	0,9	1,1	1,2	1,1
2.000.000	1,3	1,0	0,5	0,9	0,9	0,9	1,1	1,1	0,9	1,0	1,1	1,0
2.500.000	1,1	0,9	0,4	0,7	0,7	0,8	0,9	1,0	0,8	0,8	1,0	0,9

Prospetto 3 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle FAMIGLIE per aree territoriali

STIME	Isole					
	A1	A2	B1	B2	B3	B4
20.000	19,4	15,9	10,1	17,6	21,9	20,4
30.000	14,9	12,4	8,0	13,7	16,6	15,6
40.000	12,4	10,4	6,7	11,4	13,7	12,9
50.000	10,7	9,0	5,9	10,0	11,7	11,1
60.000	9,5	8,1	5,3	8,9	10,4	9,9
70.000	8,6	7,3	4,9	8,1	9,3	8,9
80.000	7,9	6,8	4,5	7,4	8,5	8,2
90.000	7,3	6,3	4,2	6,9	7,9	7,6
100.000	6,9	5,9	3,9	6,5	7,3	7,1
200.000	4,4	3,9	2,6	4,2	4,6	4,5
300.000	3,4	3,0	2,1	3,3	3,5	3,4
400.000	2,8	2,5	1,8	2,8	2,9	2,8
500.000	2,4	2,2	1,5	2,4	2,5	2,4
600.000	2,2	2,0	1,4	2,1	2,2	2,2
700.000	2,0	1,8	1,3	1,9	2,0	2,0
800.000	1,8	1,6	1,2	1,8	1,8	1,8
900.000	1,7	1,5	1,1	1,7	1,6	1,7
1.000.000	1,6	1,4	1,0	1,6	1,5	1,5
1.250.000	1,3	1,2	0,9	1,4	1,3	1,3
1.500.000	1,2	1,1	0,8	1,2	1,2	1,2
1.750.000	1,1	1,0	0,7	1,1	1,0	1,1
2.000.000	1,0	0,9	0,7	1,0	1,0	1,0
2.500.000	0,9	0,8	0,6	0,9	0,8	0,8

Prospetto 4 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle PERSONE per aree territoriali

STIME	Italia	Nord-Ovest	Nord-Est	Centro	Sud	Isole	A1	A2	B1	B2	B3	B4
20.000	43,9	41,2	32,8	34,2	36,4	31,4	37,3	36,9	27,5	34,9	36,8	32,2
30.000	34,3	32,1	25,5	26,7	28,5	24,8	29,2	28,9	21,6	27,4	28,7	25,2
40.000	28,7	26,9	21,3	22,5	24,0	21,0	24,6	24,3	18,1	23,0	24,0	21,2
50.000	25,0	23,4	18,6	19,6	20,9	18,4	21,5	21,2	15,8	20,2	21,0	18,6
60.000	22,4	20,9	16,6	17,6	18,8	16,6	19,3	19,0	14,2	18,1	18,8	16,6
70.000	20,4	19,0	15,1	16,0	17,1	15,2	17,5	17,3	12,9	16,5	17,1	15,2
80.000	18,8	17,5	13,9	14,8	15,8	14,0	16,2	16,0	11,9	15,2	15,7	14,0
90.000	17,4	16,3	12,9	13,8	14,7	13,1	15,1	14,9	11,1	14,2	14,6	13,0
100.000	16,4	15,3	12,1	12,9	13,8	12,3	14,2	13,9	10,4	13,3	13,7	12,2
200.000	10,7	9,9	7,9	8,5	9,1	8,2	9,3	9,2	6,9	8,8	9,0	8,1
300.000	8,3	7,7	6,1	6,7	7,1	6,5	7,3	7,2	5,4	6,9	7,0	6,3
400.000	7,0	6,5	5,1	5,6	6,0	5,5	6,1	6,0	4,5	5,8	5,9	5,3
500.000	6,1	5,6	4,5	4,9	5,2	4,8	5,4	5,3	3,9	5,1	5,1	4,6
750.000	4,7	4,4	3,5	3,8	4,1	3,8	4,2	4,1	3,1	4,0	4,0	3,6
1.000.000	4,0	3,7	2,9	3,2	3,4	3,2	3,5	3,5	2,6	3,4	3,4	3,1
2.000.000	2,6	2,4	1,9	2,1	2,3	2,2	2,3	2,3	1,7	2,2	2,2	2,0
3.000.000	2,0	1,9	1,5	1,7	1,8	1,7	1,8	1,8	1,3	1,7	1,7	1,6
4.000.000	1,7	1,6	1,2	1,4	1,5	1,4	1,5	1,5	1,1	1,5	1,4	1,3
5.000.000	1,5	1,4	1,1	1,2	1,3	1,3	1,3	1,3	1,0	1,3	1,3	1,2
7.500.000	1,2	1,1	0,8	1,0	1,0	1,0	1,0	1,0	0,8	1,0	1,0	0,9
10.000.000	1,0	0,9	0,7	0,8	0,9	0,8	0,9	0,9	0,6	0,8	0,8	0,8
15.000.000	0,8	0,7	0,5	0,6	0,7	0,7	0,7	0,7	0,5	0,7	0,6	0,6
20.000.000	0,6	0,6	0,5	0,5	0,6	0,6	0,6	0,6	0,4	0,6	0,5	0,5
25.000.000	0,6	0,5	0,4	0,5	0,5	0,5	0,5	0,5	0,4	0,5	0,5	0,4

Prospetto 4 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle PERSONE per aree territoriali

STIME	Piemonte	Valle D'Aosta	Lombardia	Bolzano	Trento	Veneto	Friuli-Venezia Giulia	Liguria	Emilia-Romagna	Toscana	Umbria
20.000	30,7	4,0	44,3	9,7	10,5	31,4	15,9	19,2	29,7	29,8	14,1
30.000	23,8	3,1	34,4	7,5	8,3	24,6	12,4	14,8	23,2	23,1	11,0
40.000	19,9	2,6	28,8	6,3	7,0	20,7	10,5	12,4	19,5	19,3	9,2
50.000	17,3	2,3	25,1	5,5	6,1	18,2	9,1	10,8	17,0	16,8	8,0
60.000	15,4	2,0	22,4	4,9	5,4	16,3	8,2	9,6	15,2	15,0	7,2
70.000	14,0	1,8	20,3	4,4	5,0	14,9	7,4	8,7	13,9	13,6	6,5
80.000	12,9	1,7	18,7	4,1	4,6	13,7	6,9	8,0	12,8	12,5	6,0
90.000	12,0	1,6	17,4	3,8	4,3	12,8	6,4	7,4	11,9	11,6	5,6
100.000	11,2	1,5	16,3	3,5	4,0	12,0	6,0	6,9	11,2	10,9	5,3
200.000	7,2	1,0	10,6	2,3	2,6	7,9	3,9	4,5	7,3	7,0	3,5
300.000	5,6	0,7	8,2	1,8	2,1	6,2	3,1	3,5	5,7	5,5	2,7
400.000	4,7	0,6	6,9	1,5	1,7	5,3	2,6	2,9	4,8	4,6	2,3
500.000	4,1	0,5	6,0	1,3	1,5	4,6	2,3	2,5	4,2	4,0	2,0
600.000	3,6	0,5	5,3	1,2	1,4	4,1	2,0	2,2	3,8	3,5	1,8
700.000	3,3	0,4	4,8	1,1	1,2	3,8	1,8	2,0	3,4	3,2	1,6
800.000	3,0	0,4	4,5	1,0	1,1	3,5	1,7	1,9	3,1	3,0	1,5
900.000	2,8	0,4	4,1	0,9	1,1	3,2	1,6	1,7	2,9	2,7	1,4
1.000.000	2,6	0,4	3,9	0,8	1,0	3,0	1,5	1,6	2,7	2,6	1,3
1.250.000	2,3	0,3	3,4	0,7	0,9	2,7	1,3	1,4	2,4	2,2	1,1
1.500.000	2,0	0,3	3,0	0,7	0,8	2,4	1,2	1,3	2,1	2,0	1,0
1.750.000	1,9	0,2	2,7	0,6	0,7	2,2	1,1	1,1	2,0	1,8	0,9
2.000.000	1,7	0,2	2,5	0,5	0,7	2,0	1,0	1,0	1,8	1,7	0,8
2.500.000	1,5	0,2	2,2	0,5	0,6	1,8	0,9	0,9	1,6	1,4	0,7

Prospetto 4 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle PERSONE per aree territoriali

STIME	Marche	Lazio	Abruzzo	Molise	Campania	Puglia	Basilicata	Calabria	Sicilia	Sardegna
20.000	18,6	35,7	16,5	8,3	37,8	29,8	10,9	21,4	33,6	19,9
30.000	14,5	28,0	12,9	6,6	29,8	23,5	8,6	16,8	26,5	15,8
40.000	12,2	23,5	10,8	5,5	25,2	19,8	7,3	14,2	22,4	13,4
50.000	10,7	20,6	9,5	4,9	22,1	17,4	6,4	12,5	19,6	11,8
60.000	9,6	18,4	8,5	4,4	19,9	15,6	5,7	11,2	17,6	10,6
70.000	8,7	16,8	7,7	4,0	18,2	14,2	5,2	10,2	16,1	9,7
80.000	8,0	15,5	7,1	3,7	16,8	13,2	4,8	9,4	14,9	9,0
90.000	7,5	14,4	6,6	3,4	15,7	12,3	4,5	8,8	13,9	8,4
100.000	7,0	13,5	6,2	3,2	14,7	11,5	4,2	8,3	13,1	7,9
200.000	4,6	8,9	4,1	2,2	9,8	7,7	2,8	5,5	8,7	5,4
300.000	3,6	7,0	3,2	1,7	7,7	6,0	2,2	4,3	6,9	4,2
400.000	3,0	5,9	2,7	1,4	6,5	5,1	1,9	3,7	5,8	3,6
500.000	2,7	5,1	2,3	1,3	5,7	4,5	1,6	3,2	5,1	3,2
600.000	2,4	4,6	2,1	1,1	5,2	4,0	1,5	2,9	4,6	2,9
700.000	2,2	4,2	1,9	1,0	4,7	3,7	1,4	2,6	4,2	2,6
800.000	2,0	3,9	1,8	1,0	4,4	3,4	1,2	2,4	3,9	2,4
900.000	1,9	3,6	1,6	0,9	4,1	3,2	1,2	2,3	3,6	2,3
1.000.000	1,7	3,4	1,5	0,8	3,8	3,0	1,1	2,1	3,4	2,1
1.250.000	1,5	3,0	1,3	0,7	3,4	2,6	1,0	1,9	3,0	1,9
1.500.000	1,4	2,7	1,2	0,7	3,0	2,3	0,9	1,7	2,7	1,7
1.750.000	1,2	2,4	1,1	0,6	2,8	2,1	0,8	1,5	2,4	1,6
2.000.000	1,2	2,2	1,0	0,6	2,6	2,0	0,7	1,4	2,3	1,4
2.500.000	1,0	2,0	0,9	0,5	2,2	1,7	0,6	1,2	2,0	1,3

Prospetto 4 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle PERSONE per aree territoriali

STIME	Nord-ovest						Nord-est					
	A1	A2	B1	B2	B3	B4	A1	A2	B1	B2	B3	B4
20.000	32,8	35,4	29,2	36,8	32,0	27,3	23,3	24,1	16,7	25,8	28,6	26,3
30.000	25,5	27,5	22,6	28,6	25,1	21,7	18,6	19,2	13,2	20,4	22,3	20,6
40.000	21,3	22,9	18,9	23,9	21,1	18,5	15,8	16,3	11,1	17,3	18,7	17,3
50.000	18,6	19,9	16,5	20,8	18,5	16,3	14,0	14,4	9,7	15,2	16,3	15,2
60.000	16,6	17,8	14,7	18,6	16,5	14,7	12,6	13,0	8,7	13,6	14,6	13,6
70.000	15,1	16,1	13,3	16,9	15,1	13,5	11,6	11,9	8,0	12,5	13,3	12,4
80.000	13,9	14,9	12,3	15,5	13,9	12,5	10,7	11,0	7,4	11,5	12,2	11,4
90.000	12,9	13,8	11,4	14,4	13,0	11,7	10,1	10,3	6,9	10,8	11,4	10,7
100.000	12,1	12,9	10,7	13,5	12,2	11,0	9,5	9,7	6,4	10,1	10,7	10,0
200.000	7,9	8,4	6,9	8,8	8,0	7,5	6,4	6,6	4,3	6,8	7,0	6,6
300.000	6,1	6,5	5,4	6,8	6,3	5,9	5,1	5,2	3,4	5,3	5,4	5,2
400.000	5,1	5,4	4,5	5,7	5,3	5,1	4,4	4,4	2,8	4,5	4,6	4,4
500.000	4,4	4,7	3,9	5,0	4,6	4,5	3,9	3,9	2,5	4,0	4,0	3,8
600.000	4,0	4,2	3,5	4,4	4,2	4,0	3,5	3,5	2,2	3,6	3,6	3,4
700.000	3,6	3,8	3,2	4,0	3,8	3,7	3,2	3,2	2,0	3,3	3,2	3,1
800.000	3,3	3,5	2,9	3,7	3,5	3,4	3,0	3,0	1,9	3,0	3,0	2,9
900.000	3,1	3,3	2,7	3,4	3,3	3,2	2,8	2,8	1,8	2,8	2,8	2,7
1.000.000	2,9	3,0	2,5	3,2	3,1	3,0	2,6	2,6	1,6	2,7	2,6	2,5
1.250.000	2,5	2,6	2,2	2,8	2,7	2,7	2,3	2,3	1,4	2,3	2,3	2,2
1.500.000	2,2	2,4	2,0	2,5	2,4	2,4	2,1	2,1	1,3	2,1	2,0	2,0
1.750.000	2,0	2,1	1,8	2,3	2,2	2,2	1,9	1,9	1,2	1,9	1,8	1,8
2.000.000	1,9	2,0	1,6	2,1	2,0	2,0	1,8	1,8	1,1	1,8	1,7	1,7
2.500.000	1,6	1,7	1,4	1,8	1,8	1,8	1,6	1,6	1,0	1,6	1,5	1,4

Prospetto 4 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle PERSONE per aree territoriali

STIME	Centro						Sud					
	A1	A2	B1	B2	B3	B4	A1	A2	B1	B2	B3	B4
20.000	34,4	27,5	21,5	25,5	27,2	26,2	33,3	33,8	19,7	27,4	32,4	27,4
30.000	26,7	21,8	16,7	19,9	21,3	20,3	26,4	26,5	15,7	21,6	25,4	21,6
40.000	22,3	18,5	14,0	16,7	17,9	16,9	22,4	22,2	13,5	18,2	21,3	18,3
50.000	19,4	16,3	12,2	14,6	15,6	14,7	19,7	19,4	11,9	16,0	18,6	16,1
60.000	17,3	14,7	10,9	13,1	14,0	13,1	17,7	17,4	10,8	14,3	16,7	14,5
70.000	15,7	13,4	9,9	11,9	12,8	11,9	16,2	15,9	9,9	13,1	15,2	13,2
80.000	14,5	12,5	9,2	11,0	11,8	10,9	15,0	14,6	9,2	12,1	14,0	12,2
90.000	13,4	11,6	8,5	10,3	11,0	10,2	14,1	13,6	8,6	11,3	13,0	11,4
100.000	12,6	11,0	8,0	9,6	10,3	9,5	13,2	12,8	8,1	10,6	12,2	10,7
200.000	8,2	7,4	5,2	6,3	6,8	6,1	8,9	8,4	5,6	7,0	8,0	7,2
300.000	6,3	5,9	4,1	5,0	5,3	4,8	7,1	6,6	4,5	5,5	6,3	5,7
400.000	5,3	5,0	3,4	4,2	4,4	4,0	6,0	5,5	3,8	4,7	5,3	4,8
500.000	4,6	4,4	3,0	3,6	3,9	3,4	5,3	4,8	3,4	4,1	4,6	4,2
600.000	4,1	4,0	2,7	3,3	3,5	3,1	4,7	4,3	3,1	3,7	4,1	3,8
700.000	3,7	3,6	2,4	3,0	3,2	2,8	4,3	4,0	2,8	3,3	3,8	3,5
800.000	3,4	3,4	2,2	2,7	2,9	2,6	4,0	3,6	2,6	3,1	3,5	3,2
900.000	3,2	3,1	2,1	2,6	2,7	2,4	3,8	3,4	2,4	2,9	3,2	3,0
1.000.000	3,0	3,0	1,9	2,4	2,6	2,2	3,5	3,2	2,3	2,7	3,0	2,8
1.250.000	2,6	2,6	1,7	2,1	2,2	1,9	3,1	2,8	2,0	2,4	2,6	2,5
1.500.000	2,3	2,3	1,5	1,9	2,0	1,7	2,8	2,5	1,9	2,1	2,4	2,2
1.750.000	2,1	2,1	1,4	1,7	1,8	1,6	2,6	2,3	1,7	1,9	2,2	2,0
2.000.000	1,9	2,0	1,3	1,6	1,7	1,4	2,4	2,1	1,6	1,8	2,0	1,9
2.500.000	1,7	1,8	1,1	1,4	1,5	1,2	2,1	1,8	1,4	1,6	1,7	1,7

Prospetto 4 segue - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute riferite alle PERSONE per aree territoriali

STIME	Isole					
	A1	A2	B1	B2	B3	B4
20.000	30,9	25,2	19,6	25,4	27,7	27,5
30.000	24,3	19,9	15,5	20,2	22,0	21,8
40.000	20,5	16,9	13,2	17,1	18,7	18,5
50.000	17,9	14,8	11,6	15,1	16,4	16,3
60.000	16,1	13,3	10,4	13,6	14,8	14,7
70.000	14,7	12,2	9,5	12,5	13,6	13,4
80.000	13,6	11,3	8,8	11,5	12,6	12,4
90.000	12,6	10,6	8,2	10,8	11,8	11,6
100.000	11,9	9,9	7,7	10,2	11,1	10,9
200.000	7,9	6,7	5,2	6,8	7,5	7,4
300.000	6,2	5,3	4,1	5,4	5,9	5,8
400.000	5,2	4,5	3,5	4,6	5,0	4,9
500.000	4,6	3,9	3,1	4,1	4,4	4,4
600.000	4,1	3,5	2,7	3,7	4,0	3,9
700.000	3,7	3,2	2,5	3,4	3,6	3,6
800.000	3,4	3,0	2,3	3,1	3,4	3,3
900.000	3,2	2,8	2,2	2,9	3,2	3,1
1.000.000	3,0	2,6	2,0	2,7	3,0	2,9
1.250.000	2,6	2,3	1,8	2,4	2,6	2,6
1.500.000	2,4	2,1	1,6	2,2	2,4	2,3
1.750.000	2,2	1,9	1,5	2,0	2,2	2,1
2.000.000	2,0	1,8	1,4	1,8	2,0	2,0
2.500.000	1,7	1,6	1,2	1,6	1,8	1,7

Esempi di calcolo degli errori campionari

Esempio 1

Dai dati si desume che la stima del numero di persone del Veneto che si sentono molto sicure camminando da sole per strada al buio è pari a 730 mila unità.

Nella colonna relativa alle persone del Veneto, del prospetto 4, si individua il valore corrispondente al livello di stima che più si avvicina a 730.000, ovvero 700.000. A questo valore è associato un errore relativo percentuale del 3,8%.

Pertanto, l'errore assoluto della stima sarà uguale a:

$$\sigma(730.000) = 0.025 \times 730.000 = 27.740$$

e l'intervallo di confidenza avrà come estremi :

$$730.000 - (1,96 \times 27.740) = 675.630$$

$$730.000 + (1,96 \times 27.740) = 784.370.$$

Esempio 2

Considerando la stima precedente. è possibile ottenere dei valori più precisi dell'errore di campionamento mediante interpolazione lineare dei due livelli di stima consecutivi entro i quali è compreso il valore della stessa. Tali livelli sono 700.000 e 800.000 ai quali corrispondono, rispettivamente, i valori percentuali 3,8 e 3,5.

Allora, l'errore relativo corrispondente a 730.000 è pari a:

$$\hat{\varepsilon}(730.000) = 3.8 + (3,5 - 3,8) / (800.000 - 700.000) \times (730.000 - 700.000) = 3,71 \%$$

L'errore assoluto sarà il seguente:

$$\sigma(730.000) = 0,0371 \times 730.000 = 27.083$$

e l'intervallo di confidenza avrà come estremi:

$$730.000 - (1,96 \times 27.083) = 676.917$$

$$730.000 + (1,96 \times 27.083) = 783.083.$$

Esempio 3

Il calcolo dell'errore può essere effettuato, direttamente, tramite la funzione interpolatrice $\hat{\varepsilon}(\hat{Y}) = \sqrt{\exp(a + b \log(\hat{Y}))}$ i cui parametri, riportati nel prospetto 2, per le persone della regione Veneto sono i seguenti:

$$a = 9.4986 \quad b = -1.193$$

Per $\hat{Y} = 730.000$ si ha:

$$\hat{\varepsilon}(\hat{Y}) = \sqrt{\exp(9.4986 - 1,193 \times \log(730.000))} = 0,0367.$$

L'errore relativo percentuale è quindi pari al 3.67% e il calcolo dell'errore assoluto e dell'intervallo di confidenza è del tutto analogo a quello degli esempi 1 e 2.