

Indicatori standard di qualità inclusi nei quality report

NOTA METODOLOGICA

Gli indicatori inclusi nei quality report diffusi dall'Istat sono armonizzati con gli [standard Europei di riferimento](#). In ogni report sono inclusi gli indicatori applicabili e disponibili, tra i seguenti:

- Tasso di sovra-copertura
- Tasso di mancata risposta totale
- Tasso di imputazione (per le principali variabili)
- Indicatori di qualità sulle revisioni
- Tempestività dei dati provvisori
- Tempestività dei dati definitivi
- Puntualità
- Lunghezza delle serie storiche comparabili
- Indicatore di coerenza.

Nel dettaglio, segue una sintetica spiegazione e la formula di calcolo di ogni indicatore.

TASSO DI SOVRA-COPERTURA

La lista di riferimento è l'elenco delle unità utilizzato per estrarre il campione e/o per contattare le unità stesse. Tale lista dovrebbe coincidere con l'elenco delle unità della popolazione obiettivo. La sovra-copertura rappresenta una discrepanza tra la lista e la popolazione obiettivo. Infatti, per «sovra-copertura» si intende la presenza nella lista di unità che non appartengono alla popolazione obiettivo.

$$\text{Tasso di sovra-copertura} = \frac{\text{Unità non eleggibili} + (1 - \alpha) \cdot \text{Unità non risolte}}{\text{Unità risolte} + \text{Unità non risolte}} \cdot 100$$

dove:

Unità risolte: un'unità è risolta se in corso di rilevazione è stato possibile accertare se era eleggibile o meno

Unità non risolte: un'unità è non risolta se in corso di rilevazione non è stato possibile accertare nemmeno se era eleggibile o meno (spesso chiamate anche unità con eleggibilità sconosciuta)

Unità non eleggibili: l'unità non appartiene alla popolazione oggetto di indagine pur essendo presente nell'archivio o lista di estrazione delle unità di rilevazione

α (*alpha*) = frazione delle unità non risolte che si stimano essere eleggibili. Eurostat raccomanda di porre $\alpha=1$.

TASSO DI MANCATA RISPOSTA TOTALE

La mancata risposta totale rappresenta un tentativo fallimento completo nell'ottenere l'informazione desiderata da una unità eleggibile

$$\text{Tasso di mancata risposta totale} = \frac{\text{Unità non rispondenti} + \alpha \cdot \text{Unità non risolte}}{\text{Unità rispondenti} + \text{Unità non rispondenti} + \alpha \cdot \text{Unità non risolte}} \cdot 100$$

dove:

Unità non risolte: un'unità è non risolta se in corso di rilevazione non è stato possibile accertare nemmeno se era eleggibile o meno (spesso chiamate anche unità con eleggibilità sconosciuta)

Rispondenti: unità eleggibile per la quale è stato possibile rilevare l'informazione

Non Rispondenti: unità eleggibile per la quale non è stato possibile rilevare l'informazione

α (*alpha*) = frazione delle unità non risolte che si stimano essere eleggibili. Eurostat raccomanda di porre $\alpha=1$.

TASSO DI IMPUTAZIONE

L'imputazione consiste nel sostituire i valori mancanti, errati o inconsistenti nei dati rilevati con valori opportunamente predetti sulla base dei dati osservati o di informazioni ausiliarie al fine di ottenere un insieme di dati completo e consistente.

$$\text{Tasso di imputazione (per singola variabile)} = \frac{\text{Numero valori imputati}}{\text{Totale record}} \cdot 100$$

dove:

Valori imputati: valori sui quali hanno agito le regole d'imputazione (valori modificati da codice a codice diverso + valori modificati da blank a codice + valori modificati da codice a blank)

Totale Record: numero complessivo di unità osservate

INDICATORI DI QUALITÀ SULLE REVISIONI

Per revisione si intende una modifica di un dato statistico precedentemente diffuso. L'analisi delle revisioni consente di calcolare indicatori di qualità che sull'ampiezza media, la direzione e della variabilità delle revisioni stesse, al fine di migliorare i processi statistici.

$$\text{Revisione Media} = \frac{1}{n} \sum_{t=1}^n (X_{L_t} - X_{P_t})$$

$$\text{Revisione Media Assoluta} = \frac{1}{n} \sum_{t=1}^n |X_{L_t} - X_{P_t}|$$

$$\text{Revisione Media Assoluta Relativa} = \frac{\sum_{t=1}^n |X_{L_t} - X_{P_t}|}{\sum_{t=1}^n |X_{L_t}|}$$

dove:

X_{L_t} : "Later" stima successiva, L-simo rilascio di una stima riferita al tempo t

X_{P_t} : "Previous" stima precedente, P-simo rilascio di una stima riferita al tempo t

n: numero di periodi di riferimento nella serie storica

TEMPESTIVITÀ

Per "tempestività" si intende il periodo temporale, misurato in giorni di calendario, che intercorre fra l'evento o il fenomeno descritto dai dati e quando essi vengono resi disponibili agli utenti.

Tempestività dati provvisori (gg) = *Data Pubblicazione Dati Provvisori* – *Data di Riferimento dei dati*

Tempestività dati definitivi (gg) = *Data Pubblicazione Dati Definitivi* – *Data di Riferimento dei dati*

dove:

Data di Pubblicazione Dati Provvisori: data di pubblicazione dei dati provvisori che possono essere resi disponibili in vari formati (come ad esempio volumi, dati online, etc.). Ci si riferisce alla prima versione dei dati provvisori resa disponibile

Data di Pubblicazione Dati Definitivi: data di pubblicazione effettiva dei dati definitivi che possono essere resi disponibili in vari formati (come ad esempio volumi, dati online, etc.). Ci si riferisce alla prima versione dei dati definitivi resa disponibile

Data di Riferimento dei dati: ultimo giorno del periodo di riferimento a cui sono riferiti i dati resi disponibili.

PUNTUALITÀ

Per "puntualità" si intende il periodo temporale, misurato in giorni di calendario, che intercorre fra la data prestabilita in cui avrebbero dovuto essere forniti i dati e la data della loro effettiva diffusione.

Puntualità (gg) = *Data di Pubblicazione Effettiva Dati Definitivi*
– *Data di Pubblicazione Programmata Dati Definitivi*

dove:

Data di Pubblicazione Effettiva Dati Definitivi: data della effettiva diffusione dei dati definitivi . Ci si riferisce alla prima versione dei dati definitivi resa disponibile.

Data di Pubblicazione Programmata Dati Definitivi: data ufficiale in cui i dati definitivi dovrebbero essere pubblicati, come da "[calendario delle diffusioni e degli eventi](#)"

LUNGHEZZA DELLE SERIE STORICHE COMPARABILI

La confrontabilità nel tempo di una statistica può essere inficiata da modifiche nei concetti e nelle definizioni dei fenomeni che essa misura o nelle procedure e metodologie applicate.

$$\text{Lunghezza delle serie storica comparabile} = \\ da \text{ "data di inizio della serie storica"} a \text{ "data di fine della serie storica"}$$

INDICATORE DI COERENZA

L'indicatore di coerenza permette di confrontare il risultato ottenuto dal processo statistico di riferimento con quello proveniente da una fonte esterna, qualora disponibile.

$$\text{Indicatore di Coerenza} = \\ \frac{\text{Valore della stima nel processo di riferimento} - \text{Valore della stima nella fonte di confronto}}{\text{Valore della stima nella fonte di confronto}} \cdot 100$$