



Statistical Matching

Percorso formativo: "Trattamento di archivi"

OBIETTIVI FORMATIVI

Per *statistical matching* si intende l'insieme di metodologie statistiche per l'integrazione delle informazioni statistiche contenute in due campioni diversi, selezionati dalla stessa popolazione ma non contenenti le stesse unità campionarie. L'obiettivo può essere la individuazione di un *data set* completo di tutte le variabili di interesse o la stima di parametri di variabili osservate su campioni diversi.

Il corso si propone di illustrare i principali metodi di *statistical matching* disponibili in letteratura.

Contestualmente saranno fornite indicazioni sugli strumenti *software* disponibili.

Infine, verranno illustrate alcune applicazioni condotte in Istat.

DESTINATARI

Personale con una preparazione statistica medio-alta.

PREREQUISITI

Conoscenza di: metodi di inferenza statistica per variabili normali (teoria della regressione), per variabili categoriali, stimatori di massima verosimiglianza; principi essenziali di funzionamento di R (software usato per le esercitazioni)

DOCENTI

Istat

METODI

Le lezioni saranno di natura sia teorica, sia applicativa

LETTURE CONSIGLIATE

M. D'Orazio, M. Di Zio, M. Scanu (2006). *Statistical matching: Theory and Practice*. Wiley.

IMPATTO ORGANIZZATIVO ATTESO

Arricchimento degli strumenti a disposizione degli statistici per le attività di produzione delle diverse direzioni, con l'obiettivo di un miglioramento complessivo degli *output* produttivi.

PROGRAMMA

1-a giornata – Introduzione allo *statistical matching*

9.00-13.30

- Presentazione del corso
- Definizione di *statistical matching*
 - Esempi e motivazioni
 - Fasi dello *statistical matching*
 - Preprocessamento e standardizzazione dei dati
- Metodi di *statistical matching* che usano l'indipendenza condizionata

14:30-17:30

- Descrizione di una applicazione di *statistical matching* che usa l'ipotesi di indipendenza condizionata
- Esercitazione con *StatMatch*, pacchetto R per lo *statistical matching*
- Riepilogo e chiusura

2-a giornata – Informazione ausiliaria e accuratezza

9.00-13.30

- Introduzione della giornata
- L'uso di informazione ausiliaria
- La scelta delle variabili di *matching*
- Valutazione dell'accuratezza

14:30-17:30

- Descrizione della costruzione della matrice di contabilità sociale
- Esercitazioni con *StatMatch*
- Riepilogo e chiusura

3-a giornata – Incertezza e campioni complessi

9.00-13.30

- Introduzione della giornata
- Descrizione dell'incertezza
- Metodi di *statistical matching* con dati generati da campioni generati con piani di campionamento complessi

14:30-17:30

- Descrizione di una applicazione di *statistical matching* quando si usano campioni generati tramite piani di campionamento complessi
- Esercitazioni con *StatMatch*
- Riepilogo e chiusura