Istat
Istituto Nazionale di Statistica

Rilevazione statistica sull'Innovazione nelle Imprese (CIS 2008)

Anni 2006-2008

Aspetti metodologici dell'indagine

Dicembre 2010

INDICE

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

1. Introduzione

La Cis, sviluppata congiuntamente dall'Eurostat e dagli Istituti statistici dei Paesi Ue (in collaborazione con la Commissione europea), è finalizzata a raccogliere informazioni sui processi di innovazione nelle imprese europee dell'industria e dei servizi. In particolare, la rilevazione fornisce un ampio e articolato insieme di informazioni relative alla dimensione e al settore in cui operano le imprese innovatrici, alla tipologia di innovazione introdotta e alle spese sostenute per l'introduzione di innovazioni (incluse le spese per ricerca e sviluppo), agli obiettivi dell'innovazione e al loro impatto sui risultati economici, al sostegno pubblico dell'innovazione e agli accordi di cooperazione. Raccoglie, inoltre, una serie di informazioni di carattere generale sull'appartenenza a gruppi di imprese e sul fatturato delle imprese, oltre a fornire informazioni di carattere strutturale, come l'attività economica prevalente, il numero di addetti e la regione di residenza. Tra le novità dell'edizione corrente si segnala l'inserimento di alcuni quesiti di approfondimento relativi all'impatto ambientale dell'innovazione.

L'indagine è condotta sulla base di criteri definitori e metodologie di rilevazione comuni a tutti i Paesi dell'Unione Europea ed è inserita nel quadro concettuale del cosiddetto "Manuale di Oslo", che dal 1992 rappresenta la base concettuale e metodologica per la misurazione dell'innovazione tecnologica svolta dalle imprese (Ocse/Eurostat, 1997 e Ocse/Eurostat, 2005). La rilevazione Cis viene svolta con cadenza biennale (a partire dal 2004) ed è inserita in un quadro normativo europeo (Regolamento della Commissione europea n. 1450/2004 del 13/08/2004) che ne stabilisce l'obbligatorietà per gli stati membri. Si rimanda ai capitoli successivi per una descrizione dettagliata del contenuto informativo e degli aspetti metodologici della Cis.

Oltre alla diffusione di dati aggregati a livello settoriale o dimensionale, l'Istat mette a disposizione di un'utenza specializzata - analisti e ricercatori appartenenti ad organizzazioni pubbliche o private - un file di dati elementari resi anonimi nel pieno rispetto delle norme sulla tutela della riservatezza dei dati individuali.

Scopo del presente documento è fornire alcuni elementi informativi sui principali aspetti metodologici della Cis 2008. In particolare, nel Capitolo 2 è descritta la popolazione di riferimento. Nel Capitolo 3 è illustrata il disegno di campionamento. Nel Capitolo 4 sono presentate le modalità di acquisizione e trattamento dei dati, mentre nel Capitolo 5 è indicata la procedura di calcolo delle stime campionarie. Le ultime sezioni, infine,

contengono informazioni sui principali canali di diffusione dati, sulla terminologia ricorrente e i riferimenti bibliografici fondamentali per lo svolgimento dell'indagine.

2. La popolazione di riferimento

Il campo di osservazione dell'indagine è costituito dalle imprese con almeno 10 addetti medi annui, attive nel 2008 nei seguenti settori Ateco 2007: attività estrattive (B); manifatturiero (C); fornitura di energia elettrica, gas, vapore e aria condizionata (D); fornitura di acqua, reti fognarie, attività di gestione dei rifiuti e risanamento (E); costruzioni (F); commercio all'ingrosso e al dettaglio e riparazione di autoveicoli e motocicli (G); trasporto e magazzinaggio (H); servizi di alloggio e ristorazione (I); servizi di informazione e comunicazione (J), escluse le attività di programmazione e trasmissione (divisione 60); attività finanziarie e assicurative (K); attività immobiliari (L); gli studi di architettura e ingegneria e le attività di ricerca e sviluppo (M); le attività di noleggio e leasing operativo (N).

La popolazione oggetto della rilevazione è costituita da 208.637 imprese.

3. Il disegno campionario

La rilevazione è campionaria per le imprese da 10 a 249 addetti e censuaria per quelle con almeno 250 addetti. Il disegno di campionamento è ad uno stadio stratificato con selezione delle unità a uguale probabilità. Almeno cinque imprese in ciascuno strato sono state selezionate. Nel caso di un numero totale di imprese inferiore presenti nello strato, tutte le unità sono state coinvolte. La popolazione è stata suddivisa in strati (ossia, sottoinsiemi tra loro non sovrapposti definiti sulla base di alcune caratteristiche strutturali delle unità statistiche e all'interno dei quali le unità sono fra loro omogenee riguardo alle variabili oggetto di studio). Gli strati sono definiti dalla concatenazione delle modalità identificative dei settori di attività economica (divisione Ateco 2007), delle classi di addetti (10-49 addetti, 50-249 addetti, 250 addetti e oltre) e delle regioni di localizzazione delle imprese (livello 2 della classificazione europea Nuts, disciplinata dal Regolamento Ce n.1059/2003)[1]. La stratificazione adottata costituisce la partizione minima della popolazione che permette di ottenere i domini di stima come aggregazione di strati elementari.

[1] La regione attribuita all'impresa è quella della sede legale o amministrativa come risulta dall'archivio Asia di riferimento.

La tabella seguente riporta il tipo e il numero dei domini di stima dell'Indagine:

Tipo di dominio	Descrizione	Numero di domini
DOM1	*Divisione NACE Rev.2 , ad eccezione delle sezioni F e I*	55
DOM2	Sezione NACE Rev.2 *(i servizi sono distinti in settori core e settori non core) * classe dimensionale*	48
DOM3	Regione NUTS * classe dimensionale *(PMI/grandi imprese) e macro-settore (Industria/Costruzioni/Servizi core/Servizi non core)*	168

La numerosità teorica del campione è stata ottenuta impiegando una metodologia sviluppata dall'Istat per la determinazione della dimensione campionaria ottima in presenza di molteplici vincoli e obiettivi. La metodologia impiegata adotta una generalizzazione dell'algoritmo proposto da Bethel (1989), in cui viene definito un problema di minimo vincolato con funzione obiettivo, rappresentata dal costo complessivo dell'indagine, e vincoli di tipo lineare. I vincoli sono rappresentati dai livelli massimi di errore ammessi nella stima delle principali variabili di interesse, espressi in termini di coefficiente di variazione e con riferimento al grado di dettaglio richiesto per i domini di stima finali. La ricerca dell'allocazione ottima è avvenuta assumendo numero di addetti, fatturato e spesa totale per l'innovazione come variabili d'interesse e impiegando le informazioni disponibili - in termini di medie e varianze - relative alla precedente occasione d'indagine Cis 2006 (periodo di riferimento: 2004-2006). Il dimensionamento del campione è stato, dunque, ottenuto imponendo alle stime delle tre variabili (numero di addetti, fatturato e spesa totale per innovazione) per dominio di studio un errore massimo pari al 10 per cento.

Allo scopo di limitare il carico statistico gravante sulle imprese, è stata adottata una tecnica di selezione coordinata delle unità campionarie (tecnica Jales) che consente di ridurre la probabilità di selezionare unità già estratte nell'ambito di altre indagini condotte in precedenza (Ohlsson, 1995).

La lista di riferimento per le unità campionate è l'Archivio statistico delle imprese attive (Asia).

Il campione di partenza, al netto delle imprese cessate, inattive, fallite, in liquidazione e fuori campo di osservazione e degli errori di lista, è stato di 38.200 imprese, pari al 21,5 per cento della popolazione di riferimento.

4. La rilevazione e il trattamento dei dati

La tecnica prescelta è stata quella dell'indagine postale con auto-compilazione di un questionario elettronico attraverso l'accesso personalizzato al sito web dell'Istat dedicato all'indagine: https://indata.istat.it .

A causa degli elevati tassi di non risposta, al primo invio dei questionari sono seguiti due solleciti postali rivolti a tutte le imprese non rispondenti e un terzo sollecito che ha interessato solo le imprese più grandi (con 500 addetti e oltre).

I risultati si basano su 19.904 risposte validate, pari al 52,1 per cento del campione teorico. Il campione dei rispondenti è rappresentativo della popolazione di riferimento dell'indagine, costituita da 208.637 imprese.

I dati dei rispondenti sono stati sottoposti a un processo di controllo e correzione articolato in 3 fasi principali e basato su procedure standard definite da Eurostat. La prima fase ha riguardato la localizzazione deterministica degli errori non campionari (errori di dominio, mancate risposte parziali, valori anomali e incompatibilità fra risposte, errori di codifica e di percorso), condotta sulla base degli edit specificati da Eurostat definiti a partire delle regole interne del questionario. Si è quindi proceduto all'imputazione dei valori mancanti e errati mediante l'implementazione sequenziale di procedure automatiche di tipo sia deterministico che probabilistico individuate in funzione della tipologia di variabile (quantitativa o qualitativa) e dell'errore riscontrato (incoerenze logiche, valori anomali, valori mancanti). In particolare, il processo di correzione si compone di tre passi: l'esecuzione iniziale delle procedure di imputazione logico-deduttiva che permette, sulla base di un sistema di vincoli e relazioni logiche tra le variabili, di eliminare tutte le incongruenze interne al singolo record; l'imputazione delle variabili quantitative, che viene effettuata utilizzando uno 'stimatore rapporto', basato su una serie di relazioni predefinite tra le variabili da imputare e alcune variabili ausiliarie ad esse altamente correlate; l'applicazione delle procedure di imputazione per le variabili qualitative (dicotomiche e categorico-ordinali) che utilizza il metodo del 'donatore hot deck' e consiste nell'individuare, per ogni record errato e rispetto a ciascuna specifica tipologia di variabile, il record donatore 'più vicino', i cui valori consentono al recipiente di soddisfare tutti gli edit. Il donatore è scelto in modo tale da minimizzare la distanza tra esso ed il ricevente.

Si è infine proceduto alla validazione dei dati mediante un confronto dei dati aggregati corretti e opportunamente ponderati con informazioni storiche o ausiliarie al fine di evidenziare eventuali situazioni 'sospette'.

5. La metodologia di calcolo dei pesi campionari

Data la tipologia campionaria dell'indagine, la stima dei totali delle variabili di interesse è stata calcolata attribuendo ad ogni unità rispondente un coefficiente di riporto o peso, indicante il numero di unità della popolazione rappresentate dall'impresa, inclusa se stessa. I pesi finali da associare alle unità rispondenti sono stati calcolati in base alla teoria dello stimatore di ponderazione vincolata di Deville e Särndal (1992), utilizzabile quando sono noti i totali per dominio di alcune variabili – c.d. variabili ausiliarie – correlate con quelle di interesse; le stime finali sono state prodotte per i domini di stima definiti nel paragrafo 3. Il peso finale è stato calcolato in base alla metodologia di Deville e Särndal (1992) risolvendo un sistema di minimo vincolato. Sintetizzando, si può affermare che il vettore dei pesi finali è quello che, modificando il meno possibile il vettore dei pesi diretti corretti per mancata risposta totale, rispetta per ciascun dominio di stima la condizione di uguaglianza tra le stime dei totali delle variabili ausiliarie ed i corrispondenti totali noti della popolazione. Le variabili ausiliarie utilizzate per il calcolo dei pesi finali sono state: il numero di imprese e il numero di addetti, i cui valori sono noti dall'archivio Asia 2008 per tutte le unità della popolazione obiettivo. La convergenza delle stime delle variabili ausiliarie ai corrispondenti totali noti è stata realizzata congiuntamente per i tre domini di stima già definiti.

Per il calcolo delle stime campionarie si è utilizzato Genesees, un software generalizzato sviluppato dall'Istat in linguaggio Sas e disponibile al seguente indirizzo del sito Istat: http://www.istat.it/strumenti/metodi/software/produzione_stime/genesees/index.html.

Oltre alla predisposizione del file di dati elementari per la ricerca, l'Istat ha reso disponibili i risultati dell'indagine attraverso altri canali di diffusione.

6. La diffusione dei risultati dell'indagine

In particolare, i dati e la documentazione dell'indagine sono disponibili sul sito dell'Istat all'i indirizzo: http://www.istat.it/salastampa/comunicati/non_calendario/20101209_00/

Le statistiche europee sull'innovazione sono, invece, disponibili sul sito di Eurostat: http://epp.eurostat.ec.europa.eu.

Infine, un'alternativa al file di dati elementari per la ricerca, per l'accesso ai microdati è il Laboratorio Adele dell'Istat che consente agli analisti e ai ricercatori appartenenti ad organizzazioni pubbliche o private di condurre analisi micro-economiche per esclusivi motivi di ricerca. L'accesso ai microdati avviene presso una struttura sotto costante controllo per garantire il rispetto della riservatezza dei dati individuali (per informazioni, consultare: http://www.istat.it/dati/microdati/adele.html).

7. Glossario

Il questionario dell'indagine, contenuto nel file MFR_ISTAT_CIS8_Questionario.pdf, contiene tutte le definizioni necessarie a interpretare correttamente i fenomeni rilevati. Le definizioni principali sono riportate anche di seguito e servono a distinguere le imprese sulla base delle attività innovative svolte e della tipologia di innovazione introdotta.

Innovazioni tecnologiche

Tutti i prodotti, servizi o processi introdotti dall'impresa che possono essere considerati nuovi o significativamente migliorati, rispetto a quelli precedentemente disponibili, in termini di caratteristiche tecniche e funzionali, prestazioni, facilità d'uso, ecc.. Un'innovazione tecnologica si realizza nel momento della sua introduzione sul mercato (innovazione di prodotto o servizio) o del suo utilizzo in un processo produttivo (innovazione di processo). Le innovazioni di prodotto e di processo non devono necessariamente consistere in prodotti, servizi o processi totalmente nuovi; è infatti sufficiente che risultino nuovi per l'impresa che li introduce.

Innovazioni di prodotto e servizio

Consistono nell'introduzione sul mercato di prodotti o di servizi nuovi (o significativamente migliorati) in termini di caratteristiche tecniche e funzionali, uso di materiali e componenti, prestazioni, facilità d'uso, ecc., rispetto ai prodotti e servizi correntemente realizzati e offerti sul mercato dall'impresa.

L'innovazione di prodotto o di servizio non deve necessariamente consistere in prodotti o servizi nuovi per il mercato in cui opera l'impresa; è infatti sufficiente che prodotti e servizi risultino nuovi per l'impresa che li introduce.

L'innovazione di prodotto o di servizio può essere sviluppata dall'impresa stessa o da altre imprese o istituzioni.

Innovazioni di processo

Le innovazioni di processo consistono nell'adozione di tecniche di produzione, sistemi di logistica, metodi di distribuzione di prodotti o attività di supporto alla produzione tecnologicamente nuovi (o significativamente migliorati). Tali innovazioni sono introdotte al fine di rendere l'attività aziendale economicamente più efficiente, migliorare gli standard di qualità e la flessibilità produttiva o ridurre i pericoli di danni all'ambiente e i rischi di incidenti sul lavoro.

Attività innovative

Le attività innovative sono tutte quelle che si rendono necessarie per sviluppare e introdurre prodotti, servizi o processi produttivi tecnologicamente nuovi (o significativamente migliorati). Sono da considerarsi attività innovative: la ricerca e sviluppo (R&S) svolta all'interno dell'impresa; l'acquisizione dall'esterno di servizi di R&S; l'acquisizione di macchinari, attrezzature, hardware e software finalizzati all'introduzione di innovazioni; l'acquisizione di altre tecnologie dall'esterno sotto forma di brevetti, invenzioni non brevettate, licenze, know-how, marchi connessi alla introduzione di innovazioni; le attività preliminari alla realizzazione di innovazioni, quali la progettazione e il disegno tecnico, gli studi di fattibilità, le attività di verifica e collaudo, l'ingegnerizzazione industriale, ecc.; le attività di formazione del personale necessarie per l'introduzione di innovazioni; le attività legate al lancio di innovazioni di prodotto e di servizio, quali le ricerche preliminari di mercato, i test di mercato e la pubblicità di lancio.

8. Riferimenti bibliografici

Bethel, J. (1989) "Sample allocation in multivariate surveys". *Survey Methodology*, 15: 47-57.

Deville, J.C. and Särndal, C.E. (1992) "Calibration estimators in survey sampling", Journal of the American Statistical Association 87, 367.382

OECD (2005) "Oslo-Manual. Guidelines for Collecting and Interpreting Innovation Data, 3rd Edition", Organisation for Economic Co-Operation and Development, Paris.

Ohlsson, E., B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, M. J. Kott e P. S. Colledge (eds.). "Coordination of samples using permanent random numbers". In Business Survey Methods. Wiley, New York, 1995.

9. Contatti

Per informazioni riguardanti l'indagine si prega di contattare Valeria Mastrostefano: mastrost@istat.it

Curatori dei capitoli

Il documento è stato redatto da Valeria Mastrostefano