



Appendice metodologica

Appendice Metodologica

Premessa

Gli obiettivi dell'indagine e la popolazione di riferimento sono stati descritti nel capitolo I. L'attuale rapporto conclude l'esperienza STELLA, comprendendo solo quegli atenei che a partire dal 2006, hanno sempre effettuato le rilevazioni con il questionario redatto dal Comitato Scientifico formato da delegati dei Rettori del proprio ateneo.

Per questa ultima edizione del volume, solo l'Università di Palermo ha proseguito con la rilevazione in modalità campionaria, mentre gli altri Atenei hanno scelto la modalità censuaria.

Il disegno campionario

- Modalità di rilevazione

La popolazione di riferimento è costituita dai laureati nel 2012 intervistati a circa un anno dal conseguimento del titolo di studio.

Il gruppo STELLA ha intervistato i laureati in due momenti successivi dell'anno per tenere conto delle diverse sessioni di laurea e mantenere più o meno costante e pari a 12 mesi, l'intervallo di tempo dalla laurea all'intervista.

La modalità di rilevazione adottata è stata, per questa edizione, **condotta in modalità mista CAWI e CATI.**

- Piano di campionamento

L'Università di Palermo ha scelto di intervistare tutti i loro laureati, ha adottato un piano di campionamento di tipo stratificato non proporzionale (con estrazione casuale delle unità da ogni strato). Gli strati relativi alle classi di laurea 21 – Scienze e tecnologie chimiche, 25, L-30 – Scienze e tecnologie fisiche e 32, L-35 – Scienze matematiche, mentre per i laureati magistrali alle classi 20/S, LM-17 Fisica, 45/S, LM-40 Matematica, 62/S, LM-54 Scienze chimiche e 81/S Scienze e tecnologie della chimica industriale– le cosiddette Scienze Dure – e tutti quelli con dimensione complessiva inferiore a 30 sono stati censiti.

Le variabili di stratificazione scelte sono: Ateneo, Tipo di Corso, Classe di Laurea e Corso di Studio.

L'obiettivo è stato quello di proporre stime significative anche ai livelli Ateneo, Tipo di Corso, Classe di Laurea e Corso di studio. Le modalità delle variabili di stratificazione rappresentano quindi i domini di studio.

Il numero di strati da cui è stato estratto il campione è determinato dalle modalità congiunte delle variabili di stratificazione introdotte. Il monitoraggio della numerosità degli strati, è stato

fondamentale per tenere sotto controllo la significatività dei risultati campionari.

Lo stimatore del totale del carattere y associato a questo piano di campionamento è dato da:

$$\hat{t}_y = \sum_{h=1}^L N_h \bar{y}_h$$

dove

N_h $h = 1, 2, \dots, L$ sono le dimensioni degli strati nella popolazione

obiettivo ($\sum_{h=1}^L N_h = N$).

\bar{y}_h rappresenta la media campionaria calcolata sulle n_h unità statistiche campionate all'interno dello strato h ed è quindi definita come

$$\bar{y}_h = n_h^{-1} \sum_{j=1}^{n_h} y_{hj}.$$

Possiamo quindi scrivere:

$$\hat{t}_y = \sum_{h=1}^L \sum_{j=1}^{n_h} \tilde{w}_h y_{hj}$$

$$\tilde{w}_h = \frac{N_h}{n_h}$$

Ovviamente, per gli strati censiti i pesi sono pari a 1.

In previsione della possibilità di mancate risposte, e quindi di ottenere un campione effettivo più piccolo di quello programmato, con conseguente diminuzione della qualità dei risultati, è stato estratto per ciascuno strato un campione di riserva di numerosità uguale a quella del campione "obiettivo".

Per gli strati censiti, ovviamente non esiste alcuna possibilità di individuare riserve e quindi le unità non rispondenti non possono essere sostituite. Per strati di modeste dimensioni, inoltre, in alcuni casi le riserve rappresentano semplicemente il complemento alla numerosità complessiva, meno che doppia rispetto alla dimensione complessiva del campione.

Date le scelte metodologiche effettuate dai singoli atenei, la popolazione base per l'indagine è costituita da **38.574** laureati, di cui il 59,6% composti dai laureati di I livello e il 28,7% da laureati specialistici. I residui, pari al 11,7%, costituiscono la popolazione dei laureati a ciclo unico.

- Tassi di risposta e dimensione del campione

La dimensione del campione teorico risultante dalla procedura descritta è pari a **35.527** laureati, ovvero l'92,1% della popolazione. Il tasso di risposta all'indagine è risultato pari al 72,5%. L'Università di Palermo, unica ad avere adottato la modalità di campionamento, raggiunge un tasso di risposta sul campione superiore al 95% (*Tabella VI.1.1.*).

VI.1.1 Tassi di risposta per atenei e corso di studio

STELLA				
Ateneo	LT	LM	CU	Totale
Univ. Studi di Bergamo (**)	78,5%	76,3%	100,0%	77,6%
Univ. Studi di Brescia (**)	83,7%	81,0%	81,7%	82,7%
Univ. Studi di Milano-Bicocca (**)	78,4%	78,4%	100,0%	78,5%
Univ. Studi di Milano (**)	77,1%	76,7%	70,7%	76,2%
Univ. Studi di Palermo (***)	96,9%	79,2%	100,0%	96,9%
Univ. di Pisa (**)	76,1%	73,3%	71,2%	74,7%
Univ. Studi di Pavia (**)	73,7%	70,6%	67,1%	71,6%
Totale STELLA	80,2%	76,1%	78,2%	72,5%

(**) Viene effettuato il censimento dei laureati

(***) Tasso calcolato sulla dimensione del campione

Per una informazione più completa, utile soprattutto per gli atenei che hanno scelto di effettuare il censimento si riporta anche il tasso di contatto, ovvero il numero di laureati effettivamente raggiunti (contattati) - indipendentemente dall'esito dell'intervista stessa - sul numero totale di laureati. Il tasso di contatto, che in alcuni casi raggiunge quasi il 100% dei laureati, è rimasto invariato rispetto allo scorso anno. La tabella successiva mostra quindi quanta parte della popolazione di laureati è stata comunque raggiunta:

VI.1.2 Tassi di contatto per atenei e corso di studio

STELLA

Ateneo	LT	LM	CU	Totale
Univ. Studi di Bergamo	93,4%	94,3%	95,7%	93,8%
Univ. Studi di Brescia	96,6%	97,1%	98,7%	97,1%
Univ. Studi di Milano-Bicocca	93,5%	93,5%	95,6%	93,6%
Univ. Studi di Milano	94,5%	93,3%	93,3%	94,0%
Univ. Studi di Palermo (*)	96,5%	94,1%	96,9%	95,4%
Univ. di Pisa	94,2%	94,1%	94,8%	94,2%
Univ. Studi di Pavia	95,1%	94,5%	95,6%	95,0%
TOTALE	94,8%	94,0%	95,5%	94,7%

(*) Viene effettuato il campionamento dei laureati

Rileviamo, rispetto alla scorsa rilevazione, una riduzione del tasso di risposta. Inoltre, per la prima volta, l'indagine è stata condotta in modalità mista CAWI, CATI.

Le due sessione di indagine CATI (luglio-dicembre 2013), sono state precedute da altrettante indagini CAWI (giugno-novembre 2013), organizzate con 4 invii di invito alla compilazione del questionario reso disponibile on-line dal servizio VULCANO. .

Durante la sessione CATI è stato richiesto di tracciare, per quei laureati raggiunti telefonicamente, la 'modalità di risposta', secondo la seguente tabelle:

Chiuso	Intervista parziale. Dopo essere riusciti a contattare il laureato e iniziata l'intervista, non si riesce a concluderla
Non collabora	Il laureato rifiuta di rilasciare l'intervista
Non reperibile	Il laureato vive o studia all'estero.
Numero errato	Numero di telefono errato
Abbandono	Dopo 5 tentativi non si riesce a parlare con il laureato

VI.1.3 Analisi risposte per ateneo

	Ateneo							Totale
	BGS	BSS	MIB	MIS	PAL (*)	PIS	PVS	
Rispondente - CAWI	150	273	368	615	371	1.012	469	3.258
Rispondente - CATI	1.763	1.710	3.756	6.964	4.275	3.803	2.443	24.714
Rifiuto	116	81	137	330	236	197	113	1.210
Abbandono/ Chiuso	283	263	659	1.444	2.793	1.056	838	7.336
Non reperibile	85	53	189	414	150	177	99	1.167
Numero errato	69	17	147	184	170	197	105	889
% Rispondenti CAWI	7,8%	13,8%	8,9%	8,1%	8,0%	21,0%	16,1%	11,6%
% Rispondenti CATI	92,2%	86,2%	91,1%	91,9%	92,0%	79,0%	83,9%	88,4%
% popolazione rispondenti CAWI	6,1%	11,4%	7,0%	6,2%	4,6%	15,7%	11,5%	8,4%
% popolazione rispondenti CATI	71,5%	71,3%	71,5%	70,0%	53,5%	59,0%	60,1%	64,1%
% rifiuti	4,7%	3,4%	2,6%	3,3%	3,0%	3,1%	2,8%	3,1%
% Abbandono/ Chiuso	11,5%	11,0%	12,5%	14,5%	34,9%	16,4%	20,6%	19,0%
% Non reperibile	3,4%	2,2%	3,6%	4,2%	1,9%	2,7%	2,4%	3,0%
% Numero errato	2,8%	0,7%	2,8%	1,8%	2,1%	3,1%	2,6%	2,3%
Popolazione	2.466	2.397	5.256	9.951	7.995	6.442	4.067	38.574

Il tasso di risposta delle interviste condotte via WEB è di poco superiore al 10%.

Spicca l'Università di Pisa con un sorprendente 21% di risposte ai questionari in modalità CAWI..

Solo il 3% dei laureati rifiuta esplicitamente di rispondere al sondaggio, anche se per il 19% della popolazione non si riesce a chiudere l'intervista, dopo per 5 tentativi di chiamata nel corso delle tre settimane della campagna.

La qualità dei dati amministrativi degli atenei VULCANO/STELLA è buona: solo il 2,3% dei contatti non ha un recapito telefonico o ce l'ha errato.

Classificazioni

La variabile *Livello sociale della famiglia*

Il livello sociale è stato ricavato tenendo conto sia della professione svolta che del titolo di studio di entrambi i genitori.

Dopo avere assegnato il livello sociale per ciascuno dei genitori in base al criterio esplicitato nello schema sottostante è stato assegnato come livello della famiglia il livello massimo tra quello assegnato al padre e alla madre.

Il titolo di studio dei genitori è stato trasformato in anni di scolarità totali (8 anni per scuola dell'obbligo, 13 per diploma scuola media superiore, 17 per laurea), con valori che andavano quindi da 16 (entrambi i genitori con la sola scuola dell'obbligo) a 34 (entrambi i genitori laureati).

Professioni dichiarate	Livello assegnato
Imprenditori, dirigenti e legislatori Professioni intellettuali, scientifiche ad elevata specializzazione	Alto
Professioni tecniche Impiegati	Medio
Professioni qualificate Artigiani, operai spec e agricoltori Conduttori impianti e operai semiqualeficati	Basso
Professioni non qualificate Forze Armate Nessuna Non risposta Non voglio dirlo	Non assegnabile

Gli anni di scolarità di entrambi i genitori sono stati utilizzati per attribuire i livelli sociali intermedi o assegnare il livello sociale, secondo lo schema seguente:

Max Livello tra padre e madre	anni di scolarità	Livello assegnato
Alto	26-31	Medio alto
	< 26	Medio
Medio	16	Medio basso
	> 26	Medio alto
Basso	> 16	Medio basso
	>26	Medio
Non assegnabile	27-31	Medio
	16	Basso
	17-25	Medio basso

Indicatori CI per la sintesi delle valutazioni espresse dai laureati

Come nelle passate edizioni dell'indagine STELLA, anche in questo rapporto vengono proposti gli indicatori CI per la sintesi delle valutazioni soggettive espresse dai laureati occupati in merito ai livelli di soddisfazione del lavoro svolto.

Nel questionario le opinioni dei laureati vengono raccolte utilizzando una scala a quattro gradi, con modalità 1="molto", 2="abbastanza", 3="poco", 4="per nulla".

La considerazione che una sintesi di tali valori attraverso la media aritmetica dei punteggi implicherebbe o l'assunzione di equipartizione tra i quattro gradi o l'attribuzione ad ogni modalità di risposta di punteggi scelti più o meno arbitrariamente, unitamente ad altre più strettamente attinenti al problema della misurazione, ha condotto all'utilizzo di **indici basati sulle distribuzioni osservate delle risposte**.

Questa famiglia di indici assegna un punteggio numerico di sintesi alla dimensione valutata da parte di ogni gruppo d'analisi. Esso è ottenuto come somma algebrica di due indici; il primo esprime il punteggio conseguito nel semipiano delle valutazioni positive (modalità 1 e 2), il secondo, il punteggio conseguito nel semipiano delle valutazioni negative (modalità 3 e 4). Essi assumeranno pertanto valori compresi tra -100 (quando tutte le risposte sono concentrate nella modalità 4="per nulla" e, quindi nel caso di valutazione massimamente negativa) a +100 (quando tutte le risposte sono concentrate nella modalità 1="molto", e, quindi, nel caso di valutazione massima positiva).

Al generico gruppo di laureati i , in conseguenza della distribuzione delle risposte da essi fornite dai laureati all'item h_i , corrisponde la distribuzione $N_i^h(1) N_i^h(2) N_i^h(3) N_i^h(4)$ delle frequenze associate ai quattro gradi della scala. Si indichino:

- x_i^h la percentuale di giudizi positivi per l'item h espressi dai laureati appartenenti al gruppo i :

$$x_i^h = \frac{(N_i^h(1) + N_i^h(2)) \times 100}{\sum_{j=1}^4 N_i^h(j)}$$

- y_i^h la percentuale di giudizi molto positivi per l'item h calcolati sul totale dei giudizi positivi espressi dai laureati appartenenti al gruppo i :

$$y_i^h = \frac{N_i^h(1) \times 100}{(N_i^h(1) + N_i^h(2))}$$

- x_i^h la percentuale di giudizi negativi sempre per l'item h espressi dai laureati appartenenti al gruppo i :

$$\bullet \quad x_i^h = \frac{(N_i^h(3) + N_i^h(4)) \times 100}{\sum_{j=1}^4 N_i^h(j)} = 100 - x_i^h$$

- y_i^h la percentuale di valutazioni molto negative sul totale delle valutazioni negative:

$$y_i^h = \frac{N_i^h(3) \times 100}{(N_i^h(3) + N_i^h(4))}$$

Con riferimento all'item h , nel versante delle valutazioni positive, l'unità i è rappresentata dal punto $P_i^h(x_i^h, y_i^h)$ appartenente al quadrato di lato 100 situato nel primo quadrante (spazio delle valutazioni positive); nel versante delle valutazioni negative, invece, la stessa unità i è rappresentata dal punto $Q_i^h(x_i^h, y_i^h)$ appartenente al quadrato di lato 100 situato nel terzo quadrante (spazio delle valutazioni negative). Il posizionamento dei singoli gruppi d'analisi all'interno delle due aree fornisce una visualizzazione immediata del posizionamento di ogni gruppo. Come la figura sottostante evidenzia immediatamente, al posizionamento del punto $P_i^h(x_i^h, y_i^h)$ sul vertice superiore destro corrisponde un gruppo i di laureati che hanno risposto tutti "molto" all'item h e quindi che hanno espresso la valutazione massima.

Analogamente, al posizionamento del punto $P_i^h(x_i^h, y_i^h)$ nell'origine degli assi corrisponderà un gruppo i di laureati che hanno risposto solo "poco" e/o "per nulla" e quindi il minimo giudizio.

Per qualunque altra situazione, stabilire se al gruppo 1 è associato un giudizio più positivo che al gruppo 2 (e quindi un valore numerico dell'indicatore per il gruppo 1 maggiore di quello del gruppo 2) dipende dal giudizio di valore dello "sperimentatore/decision maker", dipende cioè dal grado di importanza che egli vuole attribuire alla quota di giudizi molto positivi (e, analogamente, molto negativi). Sulla base di queste considerazioni, si è costruito, con riferimento al quadrante delle valutazioni positive, l'indice CI_i^{h+} , avente la seguente espressione:

$$CI_i^{h+} = \frac{100(x_i^h + ky_i^h)}{\max(CI_i^{h+})}$$

$$\max(CI_i^{h+}) = 100(1 + k).$$

Del tutto analogamente si procede per il quadrante negativo definendo l'indice:

$$CI_i^{h-} = \frac{-100(x_i^h + ky_i^h)}{\max(CI_i^{h-})}$$

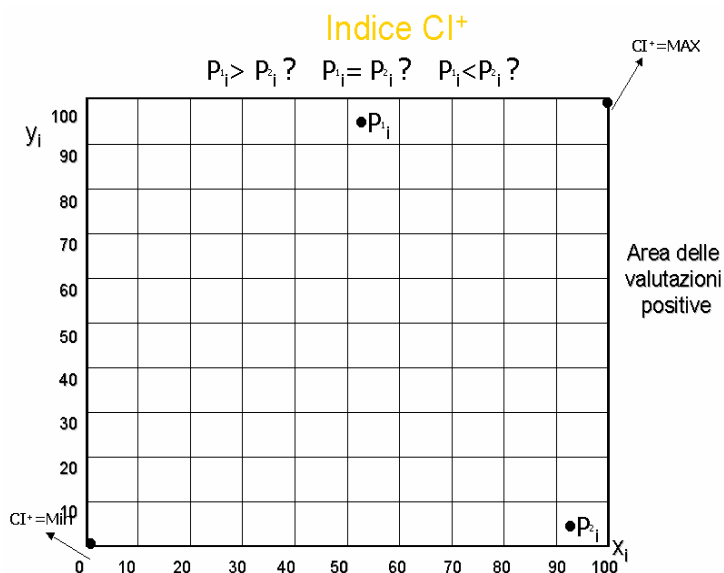
dove $\max(CI_i^{h-}) = 100(1+k)$

Si arriva quindi all'indice:

$$CI_i^h = CI_i^{h+} + CI_i^{h-}$$

con $-100 \leq CI_i^h \leq 100$.

$0 \leq k \leq 1$ è il parametro, scelto dalla "sperimentatore", che esprime appunto il livello di importanza da attribuire ai giudizi "molto positivi".



La scelta di $k=0$, equivale a ridurre la scala a quattro gradi a una scala dicotomica e, quindi, a scegliere di non attribuire alcun peso alle risposte "molto" e "per nulla". In questa applicazione si è scelto $k=0,5$, che, con riferimento al quadrante delle valutazioni positive, implica il considerare equivalenti ad esempio il livello di coerenza di un primo gruppo di laureati che hanno tutti risposto "abbastanza" con quello di un secondo gruppo in cui solo il 50% ha risposto "molto" e nessuno "abbastanza". Il valore dell'indice dei due gruppi, ovviamente, differirà in dipendenza di come, nel secondo gruppo, ha risposto il restante 50% (che ha fornito risposte con modalità negative).

È opportuno richiamare l'attenzione sul fatto che per confrontare correttamente i diversi gruppi sulla base dei valori assunti dall'indice CI è necessario che per ciascuno di essi la numerosità di rispondenti sia molto simile o, quantomeno, che non ci siano gruppi con un numero di rispondenti molto piccolo.

In questo contesto, poiché il calcolo degli indici è stato condotto sulle distribuzioni di risposte riportate all'universo, il problema su accennato non assume particolare rilevanza.

Nel volume i risultati degli indici di sintesi sono stati rappresentati graficamente suddividendo, per ogni aspetto del lavoro considerato, l'area dei giudizi positivi e l'area dei giudizi negativi.

NOTA

Nel rapporto molte tabelle presentano le distribuzioni di frequenza di un carattere condizionate alle modalità di un altro.

Qualora la dimensione della sottopopolazione intervistata e individuata dal condizionamento sia inferiore a 30, i valori delle percentuali sono stati scritti in grigio per evidenziare la cautela necessaria nell'interpretazione di stime soggette ad un'elevata variabilità campionaria.

In alcune tabelle del presente volume, inoltre, i totali, sia nei valori assoluti sia nei valori percentuali, potrebbero non corrispondere alla somma dei singoli valori pubblicati o, nel secondo caso, potrebbero non essere pari a 100. Ciò non costituisce un errore ma è semplicemente dovuto alle necessarie approssimazioni a valori interi dei risultati del coefficiente di espansione calcolato per ottenere la stima del totale della popolazione indagata.