

# Numerus

di Donato Speroni

## L'operazione chiarezza dell'Istat smonta il teatrino mediatico sui dati del lavoro

7 OTTOBRE 2015 | di Donato Speroni

La disoccupazione in Italia in agosto è davvero scesa di un punto rispetto a un anno fa, come risulta dai dati grezzi diffusi con l'ultimo comunicato dell'Istat sulle Forze di lavoro? In realtà non ne siamo certi, ci dice lo stesso Istituto di statistica: potrebbe essere diminuita di quasi due punti o rimasta stazionaria. E l'occupazione nello stesso periodo è aumentata di circa 350mila unità, come si deduce dalle stime (grezza e destagionalizzata) pubblicate nel comunicato? Forse: potrebbe essere cresciuta di oltre 600mila o di meno di 100mila. Questi margini di incertezza, detti "intervalli di confidenza", sono ben noti agli statistici perché sono connaturati a tutte le indagini campionarie. La novità è che ora l'Istat, per volontà del suo presidente Giorgio Alleva, nel quadro di un progressivo miglioramento dell'informazione statistica, rende questi intervalli facilmente calcolabili, cominciando dall'indagine sulle Forze di lavoro. Questa "operazione chiarezza" non toglie importanza alle stime diffuse dall'Istituto di statistica, ma smonta il teatrino di quei giornalisti che sono abituati a costruire titoloni su variazioni minime, ben all'interno dell'errore campionario.

Il presidente dell'Istat **Giorgio Alleva** continua il suo impegno per il **miglioramento dell'informazione statistica sul lavoro**, che avevamo già descritto in un [precedente post](#). Il [comunicato sull'occupazione](#) in agosto, diffuso il 30 settembre, ha avuto grande risonanza perché il tasso di disoccupazione, sia nel dato grezzo che in quello corretto con la destagionalizzazione, è calato a livelli che non si vedevano dal 2013. Ma è anche importante sottolineare che **per la prima volta l'Istat ha diffuso informazioni specifiche sul cosiddetto "errore campionario", cioè sul livello di precisione dei dati che rilascia.**

Il comunicato sul lavoro in agosto fornisce per la prima volta indicatori mensili sugli **occupati dipendenti e indipendenti**. Inoltre, rende disponibili gli **intervalli di confidenza dei principali indicatori non destagionalizzati**. Infatti la [Nota metodologica](#) pubblicata con il comunicato contiene un nuovo paragrafo intitolato: "Gli intervalli di confidenza". Riportiamo i passaggi più significativi con un'avvertenza: *chi non ama le spiegazioni troppo tecniche può anche saltare la citazione e leggere invece più sotto le conseguenze di questa innovazione.*

Scriva l'Istat:

Al fine di valutare l'accuratezza delle stime prodotte da un'indagine campionaria è necessario tenere conto dell'errore campionario che deriva dall'aver osservato la variabile di interesse solo su una parte (campione) della popolazione. Tale errore può essere espresso in termini di errore assoluto (standard error) o di errore relativo (cioè l'errore assoluto diviso per la stima, che prende il nome di coefficiente di variazione, CV). In questo paragrafo, per ciascuna delle principali variabili di interesse, sono riportate la stima puntuale e l'errore relativo ad essa associato.

A partire da questi è possibile costruire l'intervallo di confidenza che con un prefissato livello di fiducia, contiene al suo interno il valore vero, ma ignoto, del parametro oggetto di stima.

L'intervallo di confidenza è calcolato aggiungendo e sottraendo alla stima puntuale il suo errore campionario assoluto, moltiplicato per un coefficiente che dipende dal livello di fiducia; considerando il tradizionale livello di fiducia del 95%, il coefficiente corrispondente è pari a 1,96.

Nel prospetto A si riportano gli errori relativi (CV) delle stime non destagionalizzate dei principali indicatori riferiti al mese di agosto 2015.

PROSPETTO A. ERRORI RELATIVI DELLE STIME NON DESTAGIONALIZZATE DEI PRINCIPALI INDICATORI  
Agosto 2015

	Stima puntuale	Errore relativo (CV)
Occupati (migliaia di unità)	22.615	0,003153
Disoccupati (migliaia di unità)	2.490	0,023791
Inattivi 15-64 anni (migliaia di unità)	14.397	0,004952
Tasso di occupazione 15-64 anni (valore percentuale)	56,75	0,003413
Tasso di disoccupazione (valore percentuale)	9,92	0,023502
Tasso di inattività 15-64 anni (valore percentuale)	36,89	0,004952

(...)

Attraverso semplici calcoli, è possibile ricavare gli intervalli di confidenza con livello di fiducia pari al 95% ( $\alpha=0,05$ ). Tali intervalli comprendono pertanto i parametri ignoti della popolazione con probabilità pari a 0,95. Nel prospetto seguente sono illustrati i calcoli per la costruzione dell'intervallo di confidenza della stima degli occupati e del tasso di disoccupazione.

PROSPETTO B. CALCOLO ESEMPLIFICATIVO DELL'INTERVALLO DI CONFIDENZA

	Occupati (migliaia di unità)	Tasso di disoccupazione (%)
Stima puntuale:	22.615	9,92
Errore relativo (CV)	0,003153	0,023502
Stima intervallare		
Semi ampiezza dell'intervallo:	$(22.615 \times 0,003153) \times 1,96 = 140$	$(9,92 \times 0,023502) \times 1,96 = 0,46$

Limite inferiore dell'intervallo di confidenza:	$22.615 - 140 = 22.475$	$9,92 - 0,46 = 9,46$
Limite superiore dell'intervallo di confidenza:	$22.615 + 140 = 22.755$	$9,92 + 0,46 = 10,38$

In parole più semplici, il dato sul tasso di disoccupazione in agosto diffuso dall'Istat (9,92%) in realtà ha una probabilità del 95% per cento di collocarsi in una "forchetta" tra 9,46 e 10,38%. Se invece volessimo una probabilità del 99% la "forchetta" sarebbe ben più ampia.

Queste incertezze sono dovute al cosiddetto "errore campionario", cioè al fatto che una **indagine per campione**, come è appunto la rilevazione Forze di lavoro, per quanto ben fatta, è **soltanto una stima della realtà** e non può esserne una rappresentazione esattissima, come quella che si avrebbe, in teoria, interrogando l'universo degli interessati.

Ho applicato i parametri sui margini di errore diffusi [in un tabella](#) dall'Istat insieme al comunicato e ne ho ricavato questi dati.

LIMITE DI CONFIDENZA NELLA STIMA NON DESTAGIONALIZZATA DEI PRINCIPALI INDICATORI 15 - 64						
Mese	Occupati			Tasso di occupazione		
	Stima (in migliaia)	Limite superiore dell'intervallo di confidenza (LS)	Limite inferiore dell'intervallo di confidenza (LI)	Stima (in migliaia)	LS	LI

ago-2014	22.231	22.376	22.086	55,6	56,0	55,2
set-2014	22.437	22.574	22.299	56,1	56,4	55,7
ott-2014	22.526	22.662	22.389	56,3	56,7	56,0
nov-2014	22.232	22.369	22.095	55,6	56,0	55,3
dic-2014	22.330	22.474	22.185	55,8	56,2	55,5
gen-2015	22.151	22.281	22.020	55,4	55,8	55,1
feb-2015	22.126	22.268	21.984	55,4	55,8	55,0
mar-2015	22.200	22.338	22.062	55,6	56,0	55,3
apr-2015	22.467	22.599	22.334	56,2	56,5	55,8
mag-2015	22.455	22.587	22.323	56,2	56,6	55,8
giu-2015	22.577	22.716	22.437	56,5	56,9	56,2
lug-2015	22.668	22.808	22.528	56,7	57,1	56,3
ago-2015	22.615	22.755	22.475	56,7	57,1	56,3

	Disoccupati			Tasso di disoccupazione		
	Stima (in migliaia)	LS	LI	Stima (in migliaia)	LS	LI
ago-2014	2.705	2.809	2.602	10,9	11,3	10,4
set-2014	3.211	3.331	3.092	12,5	13,0	12,1
ott-2014	3.499	3.604	3.394	13,5	13,8	13,1
nov-2014	3.702	3.824	3.581	14,3	14,7	13,8
dic-2014	3.037	3.144	2.931	12,0	12,4	11,6
gen-2015	3.243	3.348	3.138	12,8	13,2	12,4
feb-2015	3.316	3.424	3.209	13,0	13,4	12,6
mar-2015	3.361	3.471	3.251	13,2	13,6	12,7
apr-2015	3.101	3.205	2.997	12,1	12,5	11,7
mag-2015	3.137	3.246	3.028	12,3	12,7	11,8
giu-2015	3.067	3.183	2.951	12,0	12,4	11,5
lug-2015	2.702	2.818	2.585	10,7	11,1	10,2
ago-2015	2.490	2.606	2.374	9,9	10,4	9,5

*Elaborazione Numerus su dati Istat*

Che cosa se ne deduce? Per esempio, che secondo le stime diffuse dall'Istat, il tasso di disoccupazione dall'agosto 2014 all'agosto 2015 ha avuto una variazione tendenziale pari a -1 (da 10,9 a 9,9%). Però dall'esame degli intervalli di confidenza possiamo dire che la variazione può anche essere stata di -1,8 (differenza tra il limite inferiore della forchetta di agosto 2015 e il limite superiore di un anno prima) oppure la disoccupazione può essere rimasta stazionaria (confronto tra il superiore 2015 e l'inferiore 2014). E la variazione nel numero degli occupati è stata davvero di 384mila unità, come ci dice il dato grezzo? In realtà sappiamo solo che c'è un 95% di probabilità che si sia collocata tra 99mila e 669mila unità.

**È importante tutto questo? A mio avviso sì: l'“operazione chiarezza” fatta dall'Istituto di statistica mi sembra meritoria e spero che venga estesa anche alle altre indagini campionarie. L'esistenza degli “intervalli di confidenza” (cioè delle “forchette” in cui si collocano i dati) non cancella il valore delle stime diffuse nei comunicati e che si collocano al centro di questi intervalli: se per esempio il numero degli occupati è costantemente in crescita rispetto allo stesso mese dell'anno precedente, come risulta dalla [tabella](#) diffusa dall'Istat insieme al comunicato, non c'è dubbio che la situazione in Italia sta migliorando. Perdono valore, però, i confronti tra i dati, quando si tratta di variazioni di decimali ben all'interno del margine di errore. I media tendono a esaltare variazioni spesso irrilevanti, magari dello 0,1%. In realtà non sappiamo se queste variazioni sono reali e semplicemente frutto dell'errore statistico.**

sappiamo se queste variazioni sono reali o semplicemente frutto dell'errore statistico.

**Per ora, l'operazione chiarezza nei dati sull'occupazione riguarda solo i dati grezzi e non quelli destagionalizzati. Spiega infatti l'Istituto di via Balbo:**

Nella pagina web del comunicato stampa è disponibile il file excel che riporta la tabella completa degli errori relativi riferiti alle stime mensili non destagionalizzate dei principali indicatori, calcolati a partire da gennaio 2004.

I principali Istituti di statistica non pubblicano errori campionari riferiti a stime destagionalizzate. In alcuni casi sono pubblicati gli errori campionari delle stime non destagionalizzate ritenendo che questi siano del tutto simili a quelli riferiti alle corrispondenti stime destagionalizzate. L'Istat sta conducendo studi al fine di verificare se tale approccio sia applicabile anche agli indicatori diffusi dall'Istituto.

Sarebbe anche interessante conoscere l'errore campionario relativo ad alcune disaggregazioni, per esempio quella relativa alla disoccupazione giovanile tra i 15 e i 24 anni. Le tabelle dell'Istat ci dicono che su 100 giovani di quell'età, un po' più di 15 lavorano, meno di 11 cercano lavoro e tutti gli altri (circa 74) studiano o comunque non sono interessati in questo momento a far parte del mercato del lavoro. Ma il tasso di disoccupazione si calcola in percentuale solo sulla cosiddetta "popolazione attiva" (occupati più persone in cerca) e quindi se ne ricavano quei tassi roboanti, oltre il 40% di disoccupazione giovanile, a cui alcuni giornalisti tendono ad attribuire quasi più importanza del dato generale su occupati e disoccupati tra i 15 e i 64 anni. Commette sistematicamente questo errore **Enrico Mentana**, direttore del Tg La7; ma non è certo l'unico. In realtà si tratta di **numeri senza importanza** (quelli che contano davvero per descrivere la situazione dei giovani sono quelli diffusi trimestralmente dall'Istat nella sua banca dati per le fasce 18 – 29 e 25 – 34) e **sarebbe anche interessante conoscerne l'intervallo di confidenza, che per la esiguità del campione è certamente più ampio del dato generale 15-64, rendendo questa informazione molto incerta.**

**Un'ultima osservazione.** Ho scritto che la pubblicizzazione degli intervalli di confidenza è uno dei passi compiuti da Alleva per migliorare l'informazione statistica sul lavoro. Il [comunicato sui dati trimestrali](#) diffuso il 1° settembre ne ha annunciati altri:

Nell'ambito di un programma di diffusione sempre più integrata delle informazioni statistiche, a partire dal 15 settembre l'informazione trimestrale su domanda e offerta di lavoro sarà diffusa congiuntamente nel nuovo comunicato dal titolo "Il mercato del lavoro". I contenuti del comunicato "Indicatori del lavoro nelle imprese", in calendario lo stesso giorno, verranno inglobati in tale nuova diffusione.

Il 15 settembre, infatti, l'Istat ha diffuso un [nuovo comunicato trimestrale sul mercato del lavoro](#), con questa premessa:

A partire da oggi l'Istat rilascia un nuovo comunicato trimestrale sul mercato del lavoro. Le informazioni più recenti sulla domanda di lavoro da parte delle imprese e quelle sulla offerta di lavoro degli individui (queste ultime già pubblicate lo scorso 1° settembre) vengono presentate congiuntamente allo scopo di fornire una lettura più completa delle caratteristiche dell'occupazione e delle dinamiche in atto. I contenuti del comunicato "Indicatori del lavoro nelle imprese", in calendario sempre oggi, sono stati inglobati in questa nuova diffusione. A partire dal prossimo dicembre, tuttavia, le linee di produzione corrispondenti saranno allineate e il comunicato integrato trimestrale sostituirà sia il comunicato sugli occupati/disoccupati sia quello sugli indicatori del lavoro nelle imprese.

Insomma, i dirigenti dell'Istat hanno fatto un grosso passo avanti per rendere i dati più completi e comprensibili nel loro effettivo significato. **Manca ora il passaggio successivo, al quale Alleva lavora da tempo: integrare questi dati con quelli delle altre fonti ufficiali: Ministero del Lavoro, Inps e Inail.**

Tag: Alleva, Disoccupazione giovanile, Errore campionario, Istat, Lavoro

**CONTRIBUTI** > 0

#### PARTECIPA ALLA DISCUSSIONE



Scrivi qui il tuo commento

> INVIA