

STATISTICA ECONOMICA

STATISTICA PER L'ECONOMIA

a.a. 2009-2010

Facoltà di Economia, Università Roma Tre

Archivio Statistico delle Imprese Attive (ASIA)

L'archivio è costituito dalle unità economiche che esercitano arti e professioni nelle attività industriali, commerciali e dei servizi alle imprese e alle famiglie.

Fornisce informazioni identificative (denominazione e indirizzo) e di struttura (attività economica, addetti dipendenti e indipendenti, forma giuridica, data di inizio e fine attività, fatturato) di tali unità.

Oltre che costituire la base informativa per le analisi sull'evoluzione della struttura delle imprese italiane e sulla loro demografia, l'archivio Asia rappresenta la popolazione di riferimento delle indagini sulle imprese condotte dall'Istat.

Nasce nel 1996 in base al Regolamento del Consiglio Europeo relativo al coordinamento comunitario dello sviluppo dei registri d'impresa utilizzati a fini statistici.

Viene aggiornato ogni anno integrando informazioni desumibili da fonti amministrative, gestite da enti pubblici o da società private (ad esempio: Agenzia delle Entrate, Camere di Commercio Industria Artigianato ed Agricoltura, INPS) e da fonti statistiche.

L'impresa è definita come "la più piccola combinazione di unità giuridiche, che costituisce un'unità organizzativa per la produzione di beni e servizi che fruisce d'una certa autonomia decisionale. In particolare per quanto attiene alla destinazione delle sue risorse correnti. Un'impresa esercita una o più attività in uno o più luoghi. Un'impresa può corrispondere a una sola unità giuridica".

Secondo questa definizione sono considerati impresa anche i lavoratori autonomi e i liberi professionisti.

Ai fini della produzione dell'informazione statistica, le imprese sono classificate per attività economica.

L'attività economica è la combinazione di risorse, quali attrezzature, manodopera, tecniche di fabbricazione, reti di informazione, o di prodotti, che porta alla creazione di specifici beni o servizi.

Le imprese sono distinte in base all'attività economica esclusiva o principale, secondo il criterio della prevalenza. Quando, nell'ambito di una stessa unità, sono esercitate più attività, la prevalenza è individuata sulla base del valore aggiunto ovvero, in mancanza di tale dato, sulla base del fatturato, del numero medio annuo di addetti, delle spese per il personale, delle retribuzioni lorde.

La classificazione delle attività economiche, in vigore dal primo gennaio 2008, è l'**Ateco 2007**, versione nazionale della classificazione (NACE Rev. 2) definita in ambito europeo. E' adottata anche dall'Agenzia delle Entrate e dalle Camere di Commercio.

I dati si riferiscono alle **imprese attive**, cioè alle unità che hanno svolto una effettiva attività produttiva per almeno sei mesi nell'anno di riferimento.

La *dimensione* di ciascuna impresa è misurata in termini di *persone occupate o addetti* (lavoratori dipendenti ed indipendenti), che esercitano un'attività a tempo pieno o a tempo parziale nell'impresa.

Nell'archivio Asia tutte le imprese posseggono almeno un *addetto indipendente*, perché non si considera possibile lo svolgimento di un'attività economica senza l'utilizzo di una quantità minima di lavoro.

Il numero di persone occupate è calcolato in termini di media annua.

Ad esempio, un'impresa che sia stata attiva per sei mesi con un solo addetto risulta avere 0.5 addetti. Mentre un'impresa attiva nei dodici mesi e che abbia avuto due addetti per l'intero anno ed un terzo addetto per quattro mesi risulta avere 2.3 addetti.

Indicatori di dimensione delle unità economiche

Facciamo riferimento alle unità produttive operanti nell'intero sistema economico oppure in un generico settore di attività economica o in una determinata regione. Consideriamo la distribuzione di tali unità produttive rispetto ad un carattere d'interesse (ad esempio il fatturato o il numero di addetti) suddiviso in n classi.

Con riferimento a ciascuna classe i ($i = 1, \dots, n$), indichiamo con

N_i il numero di unità economiche nella classe i ;

A_i l'ammontare del carattere posseduto dalle N_i unità appartenenti alla classe i ,

e per il totale, indichiamo con

N il numero complessivo di unità economiche,

A l'ammontare complessivo del carattere rilevato.

Ovviamente

$$N = \sum_{i=1}^n N_i \qquad A = \sum_{i=1}^n A_i$$

In assenza di altre informazioni, facciamo l'ipotesi di distribuzione uniforme del carattere tra le unità appartenenti ad una stessa classe, cioè che ciascuna unità della classe i -esima abbia dimensione $\frac{A_i}{N_i}$, pari alla dimensione media nella classe.

La distribuzione delle imprese per numero di addetti (ASIA 2006) è molto dispersa e fortemente asimmetrica.

Dal momento che il numero degli addetti di un'impresa è calcolato come media annua la classe dimensionale "1" comprende le unità con addetti da 0.5 a 1.49, la classe "2 - 9" comprende quelle con addetti da 1.50 a 9.49 e così via.

CLASSI DI ADDETTI	Asia 2007		
	Imprese	Addetti	N. Medio addetti
1	2,593,079	2,591,717	1.0
2-9	1,654,102	5,571,620	3.4
10-19	148,840	1,954,006	13.1
20-49	58,064	1,732,869	29.8
50-249	22,758	2,211,289	97.2
250 e più	3,630	3,524,529	970.9
Totale	4,480,473	17,586,031	3.9

La dimensione media

è calcolata come media aritmetica del numero di addetti per impresa:

$$\mu = \frac{A}{N} = \frac{17586031}{4480473} = 3.9$$

A causa della marcata asimmetria della distribuzione non soddisfa in modo adeguato l'esigenza di sintetizzare la dimensione del fenomeno.

In realtà nel nostro caso risente eccessivamente della presenza di impianti di piccole dimensioni.

Si noti che la moda è la classe "1" addetto e che in tale classe cade anche la mediana della distribuzione.

Una forte asimmetria a destra e una marcata dispersione sono caratteristiche di solito presenti nella distribuzione di un carattere che misuri la dimensione delle unità economiche (numero di addetti, valore aggiunto, superficie per le aziende agricole).

Dal momento che, per tali ragioni gli usuali indicatori di posizione non forniscono una sintesi soddisfacente della dimensione delle imprese, sono stati proposti indicatori alternativi che tengano conto in un certo senso della dispersione del fenomeno attraverso gli ammontari del carattere in ciascuna delle classi dimensionali.

La dimensione caratteristica

viene misurata tramite la **media entropica** definita da

$$ME = \exp \left[\frac{1}{A} \sum_i A_i \log \left(\frac{A_i}{N_i} \right) \right]$$

è una media geometrica delle dimensioni $\left(\frac{A_i}{N_i}\right)$ delle unità appartenenti alle n classi, ponderata con pesi pari ai corrispondenti ammontari del carattere nelle classi A_i , infatti

$$ME = \sqrt[A]{\left(\frac{A_1}{N_1}\right)^{A_1} \dots \left(\frac{A_n}{N_n}\right)^{A_n}}$$

Hanno peso maggiore le unità con maggiore dimensione. Il valore di ME è sempre maggiore o uguale al valore della media aritmetica semplice. *Nel nostro caso* $ME = 19.3$

Dimensione prevalente

si individua facendo ricorso ad un criterio che consiste nell'individuare non la classe dimensionale più diffusa (sarebbe la moda), ma la classe di ampiezza comprendente la quota più elevata di addetti.

E' un indicatore qualitativo, basato sulla definizione di una conveniente scala ordinale di gradi di ampiezza delle unità economiche, organizzata in modo che la definizione sia univoca, cioè che si possa rientrare in una sola classe di ampiezza prevalente.

La dimensione prevalente risulta:

molto piccola se più del 50% degli addetti è occupato in imprese con meno di 10 addetti;

piccola se più del 50% degli addetti è occupato in imprese con meno di 20 addetti e meno del 50% è occupato in imprese con meno di 10 addetti;
medio-piccola se più del 50% degli addetti è occupato in imprese con meno di 50 addetti e meno del 50% è occupato in imprese con meno di 20 addetti;
media se più del 50% degli addetti è occupato in imprese da 50 a 249 addetti;
medio-grande se più del 50% degli addetti è occupato in imprese da 250 a 499 addetti;
grande se più del 50% degli addetti è occupato in imprese con oltre 500 addetti;
senza ampiezza prevalente quando non sussistono le condizioni precedenti.

Nel nostro caso: il numero di addetti in imprese con meno di 10 addetti è $2591717 + 5571620 = 8163337$, pari al 46.4% del totale, mentre il numero di addetti in imprese con meno di 20 addetti è $8163337 + 1954006 = 10117343$ pari al 57.5% del totale. Quindi la *dimensione prevalente* è *piccola*.

In pratica la dimensione prevalente viene individuata sulla base delle q_i (gli ammontari relativi cumulati) utilizzate nella concentrazione.

La misura non è espressa in termini numerici. *Problema:* può succedere che nel tempo non si percepiscano mutamenti dimensionali di entità insufficiente a far variare la classe di ampiezza prevalente, mutamenti ai quali potrebbero anche corrispondere significative modificazioni strutturali.

Indicatori di concentrazione

Un aspetto particolare della variabilità, specifico dei caratteri quantitativi, positivi e trasferibili è la concentrazione. La concentrazione studia il grado di disuguaglianza tra le unità della popolazione nella distribuzione del totale del carattere rilevato.

E' di solito interessante valutare se l'ammontare complessivo è equamente distribuito tra tutte le unità oppure è "concentrato" in corrispondenza di una o poche di esse.

Ad esempio diremo che il reddito di un paese è tanto più concentrato quanto più il reddito complessivo è posseduto da una frazione modesta delle unità statistiche (cioè quanti più poveri vi sono in quel paese).

La concentrazione cresce con il crescere della frazione di unità statistiche che possiedono il carattere in misura inferiore alla media e quindi con il diminuire del numero di unità statistiche che possiedono il carattere in misura superiore alla media.

La concentrazione varia tra due estremi teorici:

- *concentrazione minima* o *assenza di concentrazione*: se tutte le unità possiedono lo stesso ammontare del carattere (coincide con il caso di variabilità minima);
- *concentrazione massima*: se una sola unità possiede l'ammontare complessivo del carattere e quindi tutte le altre non ne possiedono affatto (coincide con il caso di variabilità massima).

I casi reali fanno riferimento a situazioni intermedie, in cui vi sono unità statistiche che possiedono il carattere in misura inferiore e altre in misura superiore alla media.

Un indicatore di concentrazione assume il valore minimo (di solito 0) nel caso di equidistribuzione, cioè di minima concentrazione, ed il valore massimo (di solito 1) nel caso di massima concentrazione.

La misura di concentrazione più nota è data dal **rapporto di concentrazione di Gini**

$$R = 1 - \sum_{i=0}^{n-1} (q_i + q_{i+1})(p_{i+1} - p_i)$$

in cui $p_0 = q_0 = 0$ e (per ogni classe, cioè $i = 1, \dots, n$)

$p_i = \sum_{j=1}^i N_j/N$ è la frazione delle unità economiche che possiedono il carattere in misura inferiore od uguale all'estremo superiore della classe i ,

$q_i = \sum_{j=1}^i A_j/A$ è la frazione dell'ammontare complessivo del carattere posseduto da tali unità.

Esempio (Asia 2007):

classi di addetti	imprese N_i	addetti A_i	$p_i = \sum_{j=1}^i N_j/N$	$q_i = \sum_{j=1}^i A_j/A$
1	2593079	2591717	0.579	0.147
2-9	1654102	5571620	0.948	0.464
10-19	148840	1954006	0.981	0.575
20-49	58064	1732869	0.994	0.674
50-249	22758	2211289	0.999	0.800
250 e più	3630	3524529	1.000	1.000
Totale	4480473	17586031		

Il rapporto di concentrazione di Gini vale

$$R = 1 - \sum_{i=0}^{n-1} (q_i + q_{i+1})(p_{i+1} - p_i) = 0.715$$

Il rapporto di concentrazione di Gini non dipende in modo diretto dalla numerosità totale del collettivo di unità e non è detto che il suo valore cambi se varia n .

Ciò non rappresenta un problema se lo studio è rivolto alla distribuzione dei redditi, in cui possono sussistere unità la cui quota di carattere spettante diventi progressivamente minore fino ad annullarsi.

Dal punto di vista economico, la *concentrazione* viene definita come un processo di diffusione del controllo (accentramento) di tutto o di una quota rilevante dell'attività di un settore o di un intero comparto di attività da parte di un numero ristretto di imprese.

Processo che implica la progressiva riduzione del numero di imprese operanti sul mercato, poiché la concentrazione tende ad aumentare con la costituzione di imprese di grandi dimensioni, attraverso lo sviluppo e la fusione di imprese esistenti.

Un indicatore semplice di *concentrazione economica* è dato dalla frazione della variabile considerata (ad esempio numero di occupati o valore aggiunto) posseduto dalle k imprese più grandi.

Problemi: la scelta di k è arbitraria e l'indicatore non tiene conto del numero totale di imprese.

Un indicatore di concentrazione *economica* (riferita alle imprese), *finanziaria* (riferita ai gruppi, cioè ad imprese collegate da partecipazioni azionarie che operano in mercati diversi), *tecnica* (riferita agli stabilimenti, le unità locali) o *di mercato* (riferita alla distribuzione delle vendite di beni e servizi rispetto al numero di venditori) è necessario che tenga conto direttamente, oltre che del modo con cui il carattere è distribuito tra le unità, anche della numerosità del collettivo.

Ad un indicatore di concentrazione *economica*, oltre a godere delle usuali proprietà, cioè ad esempio

umentare di valore in presenza di: trasferimenti di ammontari da un'unità ad un'altra più grande, un aumento della quota di carattere rilevato spettante alle unità di maggiori dimensioni, processi casuali di crescita delle unità;

diminuire di valore quando la quota di ammontare spettante ad un'unità diventa progressivamente minore;

è richiesto di

umentare di valore se avvengono fusioni tra unità,

diminuire di valore se entrano nel sistema nuove unità al di sotto di una certa dimensione.

Indice di concentrazione di Theil:

$$T = \sum_{i=1}^n \frac{A_i}{A} \log \left[N_i \frac{A}{A_i} \right]$$

Varia tra 0 (nel caso di massima disuguaglianza) e $\log(N)$ (nel caso di equidistribuzione).

Infatti se siamo nel caso di massima disuguaglianza vi è una sola unità che possiede l'ammontare complessivo del carattere, cioè $n = 1$, $N_1 = 1$ e $A_1 = A$.

Mentre se siamo nel caso di equidistribuzione tutte le N unità sono in una sola classe a cui corrisponde l'ammontare complessivo del carattere, cioè $n = 1$, $N_1 = N$ e $A_1 = A$.

L'indice di Theil è più propriamente una misura dell'*entropia* (dell'eterogeneità) della distribuzione, cioè di quanto siamo lontani dalla situazione in cui tutte le unità presentano la stessa modalità del carattere.

Mette in evidenza che quanto maggiore è il livello di entropia, tanto più è incerto il controllo di un settore da parte di un'impresa.

L'indice relativo

$$\frac{\log(N) - T}{\log(N)} = 1 - \frac{T}{\log(N)}$$

varia tra 0 (nel caso di equidistribuzione) e 1 (nel caso di massima disuguaglianza).

Nel nostro caso

$$T = 13.7$$

($\log(N) = 15.3$), mentre l'indice relativo vale 0.1

Indice di concentrazione di Herfindal:

$$H = \sum_{i=1}^n N_i \left[\frac{A_i}{A} \frac{1}{N_i} \right]^2$$

varia tra $1/N$ (nel caso di equidistribuzione) e 1 (nel caso di massima disuguaglianza).

Possiamo ottenere il corrispondente indice relativo, che varia tra 0 (nel caso di equidistribuzione) e 1 (nel caso di massima disuguaglianza) come

$$\frac{H - 1/N}{1 - 1/N} = \frac{NH - 1}{N - 1}.$$

Nel nostro caso $H = 0.00001$, con $1/N = 0.0000002$, e l'indice relativo vale 0.00001.