

**STRUMENTI PER L'ANALISI  
DEI FENOMENI DEMOGRAFICI**

**SOMMARIO:** 1. Introduzione. - 2. Il diagramma di Lexis. - 3. Coorti e generazioni. Analisi per generazioni e analisi per contemporanei. - 4. Gli strumenti elementari dell'analisi demografica: tassi e probabilità.

## 1. INTRODUZIONE

Scopo dell'analisi demografica è la trasformazione delle informazioni statistiche grezze in dati elaborati al fine di renderli significativi. Tale trasformazione si attua attraverso le operazioni di sistemazione dei dati in tabelle, nella loro opportuna sintesi attraverso indici, etc.

In questo capitolo esamineremo i concetti e gli strumenti elementari a disposizione della demografia per analizzare i fenomeni che concorrono a determinare la dinamica demografica.

## 2. IL DIAGRAMMA DI LEXIS

Il **diagramma di Lexis** è una rappresentazione grafica di eventi demografici (nascita, matrimonio, divorzio, ..., morte) di un individuo, dovuta allo statistico tedesco Lexis (1837 - 1914). Il diagramma, ideato nel 1875, pone in corrispondenza le date di osservazione di questi eventi e le età (o le durate) a queste date; consente, dunque, di classificare le informazioni attinenti tali eventi in funzione del flusso del tempo indicato secondo due modalità contemporanee: la data e l'età.

Il diagramma di Lexis si rappresenta nel piano attraverso un sistema di assi cartesiani ortogonali in cui in ascissa si riportano i tempi (teoricamente milioni di anni, dall'origine dell'uomo) e in ordinata le età (fino al massimo  $\omega$  che, convenzionalmente, è pari a più di 100 anni).

Per rappresentare l'insieme degli eventi concernenti la vita di un individuo, il diagramma si serve della **linea di vita**, ossia di un segmento parallelo alla bisettrice del quadrante formato dal verso positivo degli assi. Essa può essere definita, quindi, come il *luogo geometrico di tutti gli eventi sopravvenuti nella vita di un individuo*; consente, conoscendo l'una delle coordinate temporali di un punto — evento (data o età) — di determinare l'altra.

Una linea di vita può indicare semplicemente i due eventi fondamentali attinenti la vita umana: la nascita e la morte.

Si consideri la figura seguente:

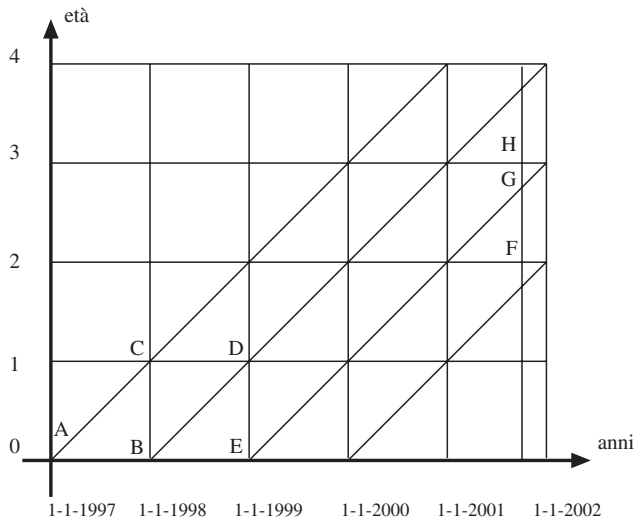


Figura 3.1

in essa le linee oblique e tra loro parallele tracciate rappresentano le linee di vita di individui diversi. La prima, che parte dal punto  $A$ , è quella di un individuo nato il primo gennaio 1997, la seconda, che parte dal punto  $B$  è quella di un individuo nato il primo gennaio 1998. Le due linee delimitano un fascio di linee di vita relative ad individui nati tra il primo gennaio 1997 e il primo gennaio 1998, la cosiddetta **generazione** di nati nel 1997. Pertanto, il segmento  $AB$  definisce l'insieme delle linee di vita degli individui nati nel 1997. Analogamente, le linee di vita che partono dal punto  $B$  al punto  $E$  delimitano un fascio di linee di vita relative ad individui nati tra il primo gennaio 1998 e il primo gennaio 1999, per cui il segmento  $BE$  definisce l'insieme delle linee di vita degli individui nati nel 1998, etc.

A questo punto è possibile dare un significato ai vari segmenti e figure geometriche che si evincono dalla figura:

- il segmento  $AB$ , come detto, definisce il fascio delle linee di vita degli individui nati nel 1997;
- il segmento  $BE$  definisce il fascio delle linee di vita degli individui nati nel 1998;
- il triangolo  $ABC$  definisce le linee di vita degli individui che, nati nel 1997, non sopravvivono al 31 dicembre 1997, dunque le linee di vita che non intersecano il segmento  $BC$ ;
- il segmento  $BC = AB - ABC$  definisce l'insieme degli individui nati nel 1997 e che sopravvivono al 31 dicembre 1997;
- il triangolo  $BCD$  definisce l'insieme degli individui nati nel 1997 e che muoiono nel corso del 1998;
- il parallelogramma  $ACDB$  definisce, dunque, l'insieme degli individui nati nel 1997 e che non sopravvivono al 31 dicembre 1998;
- il segmento  $CD = AB - ABCD$  definisce l'insieme degli individui nati nel 1997 che sopravvivono e che festeggiano il primo compleanno nel corso del 1998 ma in momenti diversi;

- il segmento verticale che si individua tra il primo gennaio 2001 e il primo gennaio 2002 indica il censimento tenuto nell'ottobre 2001 e rappresenta tutti gli individui che a quella data vivono contemporaneamente; il segmento *FH* su tale linea identifica gli individui censiti di età compresa tra 2 e 3 anni, in particolare:
- il segmento *FG* definisce gli individui nati nel 1999;
- il segmento *GH* definisce gli individui nati nel 1998.

Pertanto, dopo aver illustrato i segmenti e le superfici della figura si può affermare che:

- tutti i **segmenti paralleli all'asse delle ordinate** definiscono gli individui che vivono nello stesso momento e sono detti **linee dei contemporanei**;
- tutti i **segmenti paralleli all'asse delle ascisse** definiscono gli individui che compiono lo stesso numero di anni ma in momenti diversi nel corso dello stesso anno solare e sono detti **linee dei coetanei**;
- tutte le **superfici** (triangoli, quadrati o parallelogrammi) definiscono gli insiemi di individui che non sopravvivono ad una certa data, ossia le linee di vita che si interrompono.

Nella spiegazione della figura ci si è limitati ad annoverare solo i due eventi nascita e morte concernenti la vita di individui. In realtà, ciascun punto di una linea di vita di un individuo rappresenta un evento sopravvenuto durante il corso della sua esistenza, ed ha per coordinate le coppie tempo - età che lo contrassegna.

### 3. COORTI E GENERAZIONI. ANALISI PER GENERAZIONI E ANALISI PER CONTEMPORANEI

L'analisi demografica si interessa di eventi concernenti la vita di un insieme di individui, non di uno solo.

Si è già visto che nella figura 3.1 il segmento *AB* definisce la fascia di linee di vita rappresentative di eventi sopravvenuti nella vita di individui nati nel corso del 1997 in una data popolazione e che assumono la denominazione di **generazione** del 1997.

Si definisce **coorte** un insieme di individui caratterizzati da un **evento — origine** comune nel corso di uno stesso anno. La generazione è, dunque, un caso particolare di coorte in cui l'evento — origine comune è la nascita.

Nell'analisi demografica due nozioni appaiono di frequente: quella di *intensità* e quella di *cadenza*.

L'**intensità** di un fenomeno in una coorte misura la frequenza con cui si presenta il fenomeno ed è data dal numero totale di eventi corrispondenti vissuti da una data coorte rapportato alla consistenza iniziale della stessa; essa rappresenta, dunque, il numero medio di eventi vissuti da ciascun individuo della coorte. Nei capitoli dedicati all'analisi dei fenomeni demografici ricaveremo indici atti a misurare l'intensità di ciascuno di essi.

L'intensità può essere misurata sia all'atto dell'estinzione di una coorte (**intensità finale**), sia in un momento qualsiasi (**intensità attuale**).

L'intensità finale del fenomeno mortalità è ovviamente pari a 1.

La **cadenza** o **calendario** di un fenomeno è la distribuzione per età del fenomeno.

Generalmente si assume come indicatore della cadenza di un fenomeno l'età media alla quale esso si presenta. Pertanto, la cadenza della fecondità è data dall'età media alla nascita di un figlio.

Alcuni cambiamenti dell'indice di fecondità possono riflettere i cambiamenti, da un anno all'altro, nelle scelte della donna di avere figli: prolungamento degli studi, aspirazione a fare carriera, etc.

Date queste definizioni, è possibile, a questo punto, affermare che l'analisi demografica, avendo per oggetto un evento che contraddistingue un insieme di individui, può essere svolta seguendo due modalità differenti:

- per generazioni o per coorti;
- per contemporanei.

La figura seguente mostra le due tipologie di analisi che andiamo adesso ad illustrare:

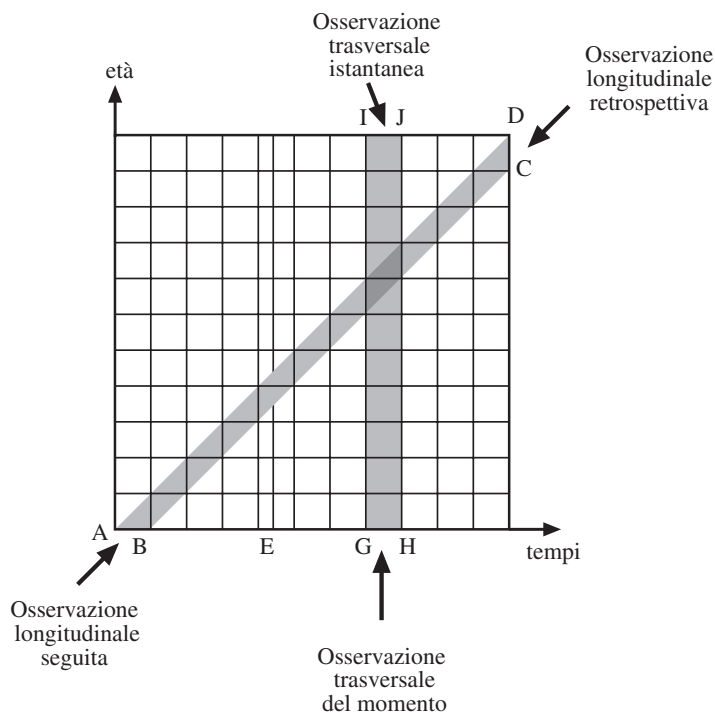


Figura 3.2

### A) Analisi per generazioni

L'**analisi per generazioni**, o **per coorti**, o **analisi longitudinale**, consiste nell'osservare, appunto, una coorte per un dato periodo di tempo. Ovviamente, alle coorti appartengono individui con conoscenze e tradizioni differenti. Solitamente si assume come tempo di osservazione dall'origine della generazione alla sua scomparsa (nella figura 3.2 dal segmento *AB* al segmento *CD*) se si tratta di analizzare il fenomeno della mortalità, oppure dalla nascita all'anno di età se il fenomeno oggetto di analisi è la mortalità infantile, oppure un'età compresa tra i 15 e i 49 anni in una generazione di donne per lo studio della fecondità, etc.

È con questa tipologia di analisi che sono generalmente studiati i fenomeni demografici, in quanto essa pone in rilievo come una medesima generazione sia esposta agli stessi condizionamenti ambientali.

A questo punto è necessario individuare le due modalità di osservazioni diacroniche dei fenomeni demografici: l'*osservazione seguita* e l'*osservazione retrospettiva*.

#### a) *Osservazione seguita*

L'**osservazione seguita** consiste nel registrare gli eventi nel momento in cui essi si producono. Per una nuova coorte si raccolgono le informazioni che la riguardano dalla sua origine.

L'osservazione condotta in questi termini presenta un'eccellente qualità dei dati raccolti ed è utilizzata fundamentalmente per l'analisi delle determinanti dei percorsi scolastici.

Tale osservazione, tuttavia:

- comporta costi notevoli ed è gravosa da effettuarsi, per cui è eseguita su campioni, i cosiddetti *panels*. Si tratta, quindi, di indagini caratterizzate da campioni, permanenti o continui, costituiti dalle stesse unità statistiche che sono intervistate, salvo sostituzioni richieste da esigenze tecniche, in successivi periodi di tempo;
- è soggetta, talvolta, ad una dispersione di una parte della coorte (per morte o emigrazione) e, di conseguenza, sussistono difficoltà di sostituzione con individui dalle caratteristiche omogenee.

#### b) *Osservazione retrospettiva*

Con l'**osservazione retrospettiva** si interrogano, in un dato momento, i superstiti di una coorte per ricostruirne la storia.

Il metodo in questione consente di raccogliere solo le informazioni sulla biografia che interessano ai fini dell'analisi.

Tuttavia, esso presenta dei limiti:

- non è utilizzabile per lo studio di fenomeni come la mortalità non essendo, ovviamente, possibile interrogare gli interessati;
- il fenomeno studiato può essere positivamente correlato con il fenomeno perturbatore eliminato; il caso tipico è la fecondità con la mortalità, entrambe assumono valori relativamente alti per date classi sociali;
- la qualità dei dati raccolti dipende dall'affidabilità della memoria degli intervistati.

### **B) Analisi per contemporanei**

L'**analisi per contemporanei** o **analisi trasversale** consiste nell'osservare simultaneamente (da qui la denominazione di **analisi sincronica**) ad un dato momento, o nel corso di un dato periodo, le coorti di una data popolazione relativamente ad uno stesso fenomeno.

Essa può essere effettuata relativamente ad un dato:

- istante;
- periodo.

a) *Osservazione ad un dato istante*

Tale osservazione trae origine da un'indagine, generalmente il censimento della popolazione, che consente, ad un certo **istante**, di conoscere lo stato di una popolazione (nella figura 3.2 è il segmento verticale *EF*). I movimenti o i cambiamenti di stato si possono evincere solo dal confronto di osservazioni successive.

b) *Osservazione su un dato periodo*

Tale osservazione consente di registrare, nel corso di un dato periodo (nella figura 3.2 la colonna *GHIJ*), il sopraggiungere di eventi relativi ad un dato fenomeno demografico.

L'osservazione condotta in questi termini differisce dall'analisi longitudinale in quanto, mentre quest'ultima studia il fenomeno in funzione dell'età di una medesima coorte di cui si segue il processo di invecchiamento, la prima raccoglie informazioni su diverse coorti che, durante il periodo di osservazione, hanno differente età con tradizioni evidentemente differenti.

#### 4. GLI STRUMENTI ELEMENTARI DELL'ANALISI DEMOGRAFICA: TASSI E PROBABILITÀ

Per lo studio della distribuzione nel tempo di un dato evento, l'analisi demografica si serve di due indici elementari: tassi e probabilità.

##### A) Tassi

Si è già visto che, nell'analisi demografica, il termine tasso rappresenta il rapporto tra il numero di eventi osservati nel corso di un dato periodo e la media della consistenza numerica della popolazione nel corso del periodo.

Sia  $E_{t, t+1}$  il numero di eventi osservati tra il tempo  $t$  e il tempo  $t + 1$  e siano  $P_t$  e  $P_{t+1}$  la popolazione al tempo  $t$  e al tempo  $t + 1$ , rispettivamente, si ha che un generico tasso  $r$  è dato da:

$$r = \frac{E_{t, t+1}}{\frac{P_t + P_{t+1}}{2}}$$

##### B) Probabilità

Nell'analisi demografica il concetto di **probabilità** è utilizzato per studiare fenomeni non rinnovabili o per natura, come la mortalità, o per il loro grado, come il primo matrimonio, il primogenito etc.

Secondo la definizione classica, la probabilità di un evento è data dal rapporto tra numero di casi favorevoli al verificarsi dell'evento e numero di casi possibili.

Come vedremo meglio in seguito, per calcolare la probabilità di morte di un individuo di età  $x$  si rapporta la differenza tra il numero di persone viventi di età  $x$  e il numero di persone di età  $x + 1$  (ossia il numero di decessi avvenuti in quell'intervallo) al numero di persone di età  $x$ .

Il complemento a 1 del rischio di subire l'evento è la probabilità di sottrarsi allo stesso, ossia la probabilità di sopravvivere, di restare celibe etc.

### ANALISI DELLA STRUTTURA DELLA POPOLAZIONE

**SOMMARIO:** 1. Introduzione. - 2. Analisi della struttura per età. - 3. Analisi della struttura per sesso. - 4. Analisi della struttura per stato civile. - 5. I tassi specifici. Intensità della mortalità. - 6. Eliminazione degli effetti della struttura per età. - 7. La transizione demografica.

#### 1. INTRODUZIONE

L'analisi della distribuzione di una popolazione in funzione di una o più variabili, quali età, sesso, stato civile, livello d'istruzione, professione etc., ossia in funzione delle sue **caratteristiche strutturali**, assume un rilievo fondamentale in demografia in quanto, attraverso strumenti matematico — statistici in grado di rappresentare e sintetizzare la distribuzione di una popolazione in funzione di tali variabili, permette di rendere immediate eventuali affinità o differenze tra popolazioni a confronto, inoltre consente di porre le basi per previsioni demografiche.

In questo capitolo analizzeremo la distribuzione di una popolazione in funzione dell'età, del sesso e dello stato civile.

#### 2. ANALISI DELLA STRUTTURA PER ETÀ

Nell'analisi della struttura per età di una popolazione è di fondamentale interesse disporre di grafici e indici dai quali si possa evincere in quale misura la composizione per età di una popolazione possa influenzare l'evoluzione dei tassi di natalità e di mortalità, lo stato di avanzamento del processo di invecchiamento demografico che a sua volta dipende dall'andamento attuale o passato delle nascite e dei decessi.

##### A) Piramide delle età

Per rappresentare la distribuzione della popolazione di un dato paese secondo l'età ed il sesso e con riferimento ad un dato arco temporale si ricorre ad una particolare tipologia di grafico: la **piramide delle età**.

Si tratta di un doppio istogramma che presenta su un asse verticale le classi d'età, solitamente ad intervalli di cinque anni, e su un asse orizzontale il numero complessivo di appartenenti a ciascuna classe di età, in modo che ciascuna di esse sia rappresentata da superfici rettangolari aventi basi uguali o proporzionali al numero di individui ed altezze uguali all'ampiezza comune di classe. La rappresentazione grafica si ottiene dalla sovrapposizione di questi rettangoli ed esprime, appunto, le proporzioni esistenti fra il numero di persone a diverse classi di età e la ripartizione dei sessi per ciascuna di queste classi.

Generalmente, i maschi sono posti alla sinistra dell'asse verticale, le femmine alla destra e, in mancanza di ulteriori indicazioni, deve intendersi impiegata tale convenzione.

La rappresentazione grafica consente di cogliere con evidenza visiva la struttura per età e per sesso e il suo andamento. Attraverso la piramide dell'età si riescono ad evidenziare aspetti quali la variabilità dei fenomeni investigati, eventuali valori anomali (a seguito di eventi bellici, o boom demografici).

La mente umana deduce dal grafico tali aspetti pressoché istantaneamente, mentre deve impiegare assai più tempo per leggere tutte le cifre esposte in una tabella e collegarle tra loro.

Per operare confronti delle strutture demografiche di popolazioni appartenenti a territori diversi, o a periodi diversi, si pone sull'asse orizzontale la consistenza numerica degli individui per classi di età in percentuale di quella totale della popolazione.

La piramide delle età può assumere andamento diverso a seconda della tipologia demografica e può essere:

- a **triangolo**, o ad **accento circonflesso**, che è la forma tipica di una popolazione in cui la mortalità riduce progressivamente la consistenza numerica della popolazione con l'età, con un numero elevato nelle età giovanili (graficamente ciò si traduce in un'ampia base della piramide) e man mano sempre più scarsa al crescere dell'età, fino a giungere al vertice del triangolo, cui corrispondono i pochi individui di età avanzata sopravvissuti. Questa forma è caratteristica di un andamento demografico con alta natalità e alta mortalità ad ogni età;
- a **campana** con base media e profilo arrotondato, tipica di una popolazione stazionaria in cui la mortalità è diminuita a tutte le età e il tasso di natalità ha subito una flessione;
- a **salvadanaio** con una piramide a base ampia ma che si restringe nella parte media a fronte di un corpo finale di rilevante consistenza, indica una ripresa della natalità in un paese caratterizzato nel recente passato da una decadenza demografica;
- ad **albero** quando la piramide è stretta alla base. Tale forma rappresenta la situazione dei paesi demograficamente senili dove la natalità e la mortalità decrescono continuamente.

Le figure seguenti rappresentano la piramide dell'età della popolazione residente (in percentuale) in Italia per sesso e per classi di età, rispettivamente, al 10 giugno 1911 e al 1° gennaio 2001:

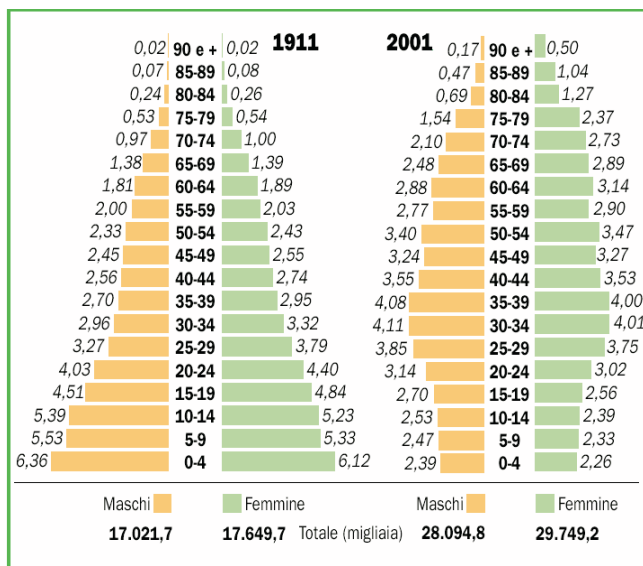


Figura 4.1 - Fonte: ISTAT

Dalla piramide relativa al 2001 si desume una sostanziale uguaglianza, per le diverse fasce d'età, tra le rispettive percentuali, maschili e femminili. Dalla fascia 0-4 anni fino alla fascia 35-39 anni, la percentuale dei maschi è più alta di quella delle femmine, successivamente, per la fascia di età 40-44 anni vi è scarso scostamento tra le percentuali e, infine, dalla fascia 45 - 49 anni la percentuale delle femmine supera quella dei maschi.

Si noti che la curva si accosta all'asse delle ordinate in corrispondenza della classe d'età 55-59 anni, ciò ad evidenziare il declino delle nascite durante il periodo bellico, ossia delle generazioni 1941-1945.

Le percentuali più alte si riscontrano, sia per i maschi sia per le femmine, in corrispondenza della fascia di età tra i 30 e i 39 anni, ciò a testimoniare l'elevato numero di nascite durante il boom economico degli anni dal 1960 al 1970.

Ancora più interessante è il confronto tra le due piramidi. Ciò che risulta evidente dai grafici è che la vita media era più bassa nel 1911 rispetto al 2001, in quanto le percentuali relative alle fasce d'età nella parte superiore della piramide del 1911 sono più basse rispetto alle corrispondenti del 2001. Inoltre, la composizione percentuale dei giovani fino a 24 anni era più alta nel 1911 che nel 2001.

I movimenti demografici dovuti a trasformazioni delle condizioni generali di natalità e mortalità sono lenti e modificano il profilo della piramide solo nel lungo periodo. Tuttavia, il grafico porta i segni di eventi eccezionali che, durante i decenni precedenti l'osservazione, hanno fortemente influenzato il numero delle nascite, delle migrazioni e dei decessi; tra questi, senza ombra di dubbio, hanno esercitato maggiori effetti, la seconda guerra mondiale e il baby-boom degli anni dal 1960 al 1970.

## B) Indici della struttura per età

Come si è già anticipato, l'analisi demografica dispone di indici, ossia di misure sintetiche in grado di disperdere al minimo le informazioni sui dati originari e che assumono espressioni diverse a seconda delle caratteristiche che si misurano.

Di seguito ci occuperemo dei principali indici della struttura per età di una popolazione.

### a) *Età media*

L'**età media** costituisce un primo indice per un'analisi della struttura demografica. Essa è, in generale, la media aritmetica delle età relative agli individui che costituiscono la popolazione. Generalmente, la distribuzione delle età del collettivo è per classi dunque, l'età media si esprime come la media delle età ponderata con la consistenza numerica di ciascuna classe.

La sua espressione analitica è la seguente:

$$\bar{x} = \frac{\sum \left( x + \frac{1}{2} \alpha \right) P_{x,x+\alpha}}{\sum P_{x,x+\alpha}}$$

in cui:

- $\alpha$  è l'ampiezza di ciascuna classe;
- $P_{x,x+\alpha}$  indica la frequenza assoluta di ciascuna classe.

In una popolazione in cui i movimenti migratori sono trascurabili, per cui l'eventuale processo di senilizzazione o ringiovanimento della stessa dipendono esclusivamente dalla mortalità o natalità, l'indice assume rilevanza fondamentale.

Non bisogna confondere i concetti di *età media* e *vita media*. La prima dipende dalla struttura per età della popolazione, la seconda, invece, dipende dalla legge di mortalità. Di essa ci occuperemo diffusamente in seguito.

b) *Età mediana*

L'**età mediana** è quell'indice che bipartisce la distribuzione ordinata in senso crescente delle età della popolazione. Per cui, al primo gruppo apparterranno le osservazioni inferiori o uguali all'età mediana, al secondo gruppo le osservazioni superiori o uguali alla stessa.

Sia  $N$  la numerosità della popolazione, se si suppone di disporre della distribuzione delle età non raggruppate in classi allora, se il numero  $N$  delle osservazioni è dispari, l'età mediana sarà il valore che occupa la posizione centrale nella distribuzione, ossia  $C = \frac{N+1}{2}$ ; quando, invece,  $N$  è pari, l'età mediana coincide con la semisomma delle intensità individuate dai due posti centrali  $C_1 = \frac{N}{2}$  e  $C_2 = \frac{N}{2} + 1$ .

Generalmente, si dispone della distribuzione della popolazione per classi di età per cui, per specificare l'età mediana è necessario supporre che la popolazione si distribuisca linearmente in ogni classe e si applica un procedimento di interpolazione lineare. Si calcolano, dapprima, le frequenze cumulate; la classe mediana è quella in cui ricadono la metà delle frequenze cumulate. Quindi, si può applicare l'espressione analitica dell'età mediana, che è:

$$x_{Me} = L_1 + \left( \frac{C - C_x}{C_{x,x+\alpha} - C_x} \right) \alpha$$

in cui:

- $L_1$  è l'estremo inferiore della classe mediana;
- $C$  è il posto centrale;
- $C_x$  è l'accumulo delle frequenze di tutte le classi inferiori alla mediana;
- $C_{x,x+\alpha}$  è la frequenza cumulata che corrisponde alla classe mediana;
- $\alpha$  è l'ampiezza della classe mediana.

Nell'espressione entro parentesi il denominatore del rapporto rappresenta quella che in statistica è denominata frequenza della classe mediana; essa si individua facilmente scorrendo lungo la colonna della popolazione nella distribuzione per classi d'età.

Procedimento analogo si applica per determinare l'età del primo e del terzo quartile, altri due indici mutuati dall'analisi statistica e che consentono, dopo aver ripartito la distribuzione della popolazione per età (o classi d'età), di misurare l'incidenza delle classi d'età giovani — se si tratta del primo quartile — o anziane — se si tratta del terzo quartile — sul totale della popolazione.