

Lezione 2: Variabili statistiche

Corso di Statistica
Facoltà di Economia
Università della Basilicata

Prof. Massimo Aria

- aria@unina.it

Introduzione

I dati raccolti attraverso un'indagine sulle unità statistiche di un particolare collettivo (sia esso una popolazione o un campione) rispetto a uno o più fenomeni rappresentano un'**informazione statistica**.

Questa informazione è codificata attraverso il concetto di **carattere**, che può essere inteso come il fenomeno oggetto di studio, rilevato sulle unità statistiche della popolazione di riferimento e codificato secondo le esigenze dell'analisi statistica.

Caratteri statistici

Un carattere prende il nome di:

- **variabile** quando assume come modalità dei numeri reali
- **mutabile** quando invece le modalità rappresentano delle qualità (dette *attributi*)

Nel linguaggio moderno, spesso il termine variabile viene usato in maniera indistinta sia per le variabili in senso stretto sia per le mutabili.

In questo caso alla parola variabile si fa seguire l'aggettivo *quantitativa* o *quantitativa* per indicare la corretta natura del carattere considerato.

Esempi di caratteri statistici

Variabili:



età	peso	altezza
19	86	180
20	78	174
21	84	192
22	63	158
34	71	169
...
...
...

Modalità

Mutabili:



Colore Capelli	Professione	Grado istruzione
Rosso	Operaio	elementare
Bruno	Impiegato	laurea
Biondo	Dirigente	diploma
Bruno	Disoccupato	media
Rosso	Operaio	laurea
...
...
...

Esempi di caratteri statistici

Scale di misura

Una **scala di misura** rappresenta lo strumento logico attraverso cui avviene la misurazione di un carattere statistico.

Nel caso di una mutabile si parla di **scale di misura per categorie** mentre per una variabile si fa riferimento a **scale di misura metriche** (o dimensionali).

Tipologie di scale di misura

Le diverse **tipologie di scale di misura** sono:

- per **variabili qualitative** (o mutabili)
 - *Nominale*
 - *Ordinale*
- per **variabili quantitative**
 - *Intervalli*
 - *Rapporti*

Scala Nominale

Un **carattere su scala nominale** (o *carattere sconnesso*) è una mutabile le cui modalità sono degli attributi che si caratterizzano per l'assenza di un ordine preconstituito.

Ne deriva che non è possibile ordinare le unità statistiche sulla base delle risposte ma solo stabilire se due di esse possiedono o meno lo stesso attributo.

Un particolare carattere di questo tipo è la **variabile dicotomica**, cioè quella mutabile che assume solo due possibili modalità (vero-falso, si-no, Maschio-Femmina, ecc.)

mutabili con scala nominale:

- *simmetria* ossia se $x_i = x_j; x_j = x_k$; allora $x_i = x_k$
- *transitività* ossia se $x_i = x_j$, anche $x_j = x_i$

Proprietà dei caratteri su scala nominale

Esempi di carattere su scala nominale

- Stato civile

attributi:

- Celibe
- Nubile
- Sposato
- ecc

- Settore di attività di un'azienda

attributi:

- Tessile
- Meccanico
- Commercio
- Ecc

- Nazionalità di un turista

attributi:

- Italiana
- Francese
- Tedesca
- Ecc.

Scala Ordinale

Un **carattere su scala ordinale** (o *carattere ordinale*) è una variabile le cui modalità sono degli attributi logicamente sequenziali, in ordine crescente o decrescente.

Ne deriva che è possibile ordinare le unità statistiche sulla base delle risposte anche se, essendo le risposte di natura qualitativa, non è possibile misurarne la distanza.

mutabili con scala ordinale:

- *simmetria, transitività*
- *ordine* $x_i < x_j$

Proprietà dei caratteri su scala ordinale

Esempi di carattere su scala ordinale

- Titolo di studio

attributi:

- Nessun titolo
- Licenza media inferiore
- Licenza media superiore
- Laurea triennale
- Laurea magistrale

- Indice di gradimento

attributi:

- Per niente
- Poco
- Abbastanza
- Molto

- Gerarchia militare

attributi:

- Soldato semplice
- Caporale
- Sergente
- Maresciallo
- Ecc.

Scala ad intervallo

Un **carattere su scala ad intervallo** (o *carattere intervallare*) è una variabile che assume valori numerici che consentono confronti solo per differenza tra le modalità che le unità assumono.

Ne deriva che è possibile ordinare le unità statistiche sulla base delle risposte ed inoltre misurare la differenza esistente tra i valori assunti da ognuna di esse.

Le scale ad intervallo sono quelle che assumono un zero arbitrario che rappresenta una convenzione e non l'assoluta assenza del fenomeno.

variabili con scala ad intervallo:

- *simmetria, transitività, ordine*

- *differenza* $x_i - x_j$

Proprietà dei caratteri intervallari

Esempio di carattere su scala ad intervallo

- Temperatura in gradi

La misurazione può avvenire con la scala Celsius oppure con quella Fahrenheit. Lo zero della prima scala è solo una convenzione che equivale nella seconda a 32 gradi. Ciò perché la gradazione $F = (9/5)C + 32$.

In questo caso si possono confrontare le differenze tra temperature ma non i rapporti:

Milano 10°C equivalenti a 50°F

Palermo 20°C equivalenti a 68°F

Non è corretto affermare che a Palermo la temperatura è doppia di Milano in quanto ciò vale solo per la scala Celsius e non per quella Fahrenheit dove invece la temperatura sarebbe superiore solo del 36%.

Scala di rapporto

Un **carattere su scala di rapporto** (o *carattere proporzionale*) è una variabile che assume valori numerici che consentono confronti sia per differenza che per rapporto tra le modalità che le unità assumono.

Ne deriva che oltre ad ordinare e confrontare le differenze, è possibile rapportare in termini proporzionali i valori assunti dalle diverse unità.

Le scale di rapporto sono quelle che assumono un zero assoluto che indica l'assoluta assenza del fenomeno.

variabili con scala di rapporto:

- *simmetria, transitività, ordine, differenza*
- *rapporto x_i/x_j*

Proprietà dei caratteri proporzionali

Esempio di carattere su scala di rapporto

- **Peso**

La misurazione può avvenire con diverse unità di misura (kg, g, quintali, tonnellate) ma comunque se tra due individui il primo pesa il doppio del secondo ciò sarà vero per ognuna delle unità di misura utilizzate.

Tizio peso 80kg equivalente a 0,8 quintali

Caio peso 60kg equivalente a 0,6 quintali

Il peso di tizio sarà superiore a quello di Caio del 33% a prescindere della scala utilizzata!!!

-Altri esempi sono:

- Altezza
- Età
- Numero di figli
- Numero addetti di un'azienda
- Ecc.

Tipi di variabili quantitative

Sulla base dei valori assunti dalle modalità, le variabili quantitative (ad intervallo o di rapporto) posso distinguersi in:

- **Variabili discrete**

possono assumere un numero discreto (cioè finito) di modalità che possono essere sempre poste in corrispondenza con l'insieme dei numeri naturali.

esempio: numero figli, voto di laurea, punteggio nel lancio di un dado, ecc.

- **Variabili continue**

possono assumere un qualunque valore contenuto in un intervallo reale predefinito.

esempio: età, peso, altezza, temperatura, ecc.

Nella prossima lezione

Nella prossima lezione si affronteranno i seguenti argomenti:

- Distribuzioni unitarie
- Distribuzioni di frequenza
- Rappresentazioni grafiche