

Metodi quantitativi per le scienze sociali

Servizio Sociale e Scienze Criminologiche (L-39)

andrea.priulla@unikore.it

Statistica - Concetti Principali

Il codice della lezione

La lezione

1 Informazioni

2 Introduzione alla metodologia statistica

- Cosa è la Statistica?
- Statistica Descrittiva e Statistica Inferenziale
- Popolazioni e Campioni

3 Le variabili statistiche

- Introduzione
- Le variabili e la loro misurazione

Informazioni

Il corso di **statistica**:

- 6 CFU (36 ore)
 - 1 Introduzione alla statistica
 - 2 Statistiche descrittive
 - 3 Cenni di probabilità
 - 4 Cenni di regressione
- contatti:
 - andrea.priulla@unikore.it
- libro di testo:
 - *Metodologie per le scienze economiche e sociali*. Borra, Di Ciaccio (2014).

Introduzione alla metodologia statistica

La lezione

1 Informazioni

2 Introduzione alla metodologia statistica

- Cosa è la Statistica?
- Statistica Descrittiva e Statistica Inferenziale
- Popolazioni e Campioni

3 Le variabili statistiche

- Introduzione
- Le variabili e la loro misurazione

Perché studiare la statistica?

Negli ultimi decenni si è osservata una crescita nell'impiego delle **metodologie statistiche** in **molti settori**:

- nelle **scienze sociali** per raccogliere ed elaborare le informazioni relative a un determinato fenomeno sociale
 - quali fattori influenzano i risultati scolastici degli studenti?
 - quali elementi influenzano le opinioni politiche degli individui?
 - cosa influisce sulla scelta di ritirarsi dall'attività lavorativa?
- nel **settore privato**:
 - analisi di mercato, marketing, data scientists, ingegneri del dato, ...
 - social media manager, esperti di comunicazione, ...
- nel **settore pubblico**:
 - pubbliche amministrazioni
 - aziende sanitarie
 - sistema scolastico

Perché studiare la statistica?

Questo incremento è dovuto a **numerosi fattori**:

- la **disponibilità** immediata di informazioni **quantitative** direttamente utilizzabili
 - open data, internet, social networks, dati satellitari, ...
- lo sviluppo di **software** e **computer** sempre più potenti e accessibili
 - **gratuiti** e liberamente utilizzabili (es. R, Gretl, Python, ...)
 - **più semplici** da utilizzare (es. tutorial gratuiti, blogs, ChatGPT, ...)
 - servizi di **cloud computing/storage**
- in molti settori la **conoscenza dei metodi statistici** è diventata una **competenza fondamentale**
 - occorre saper **applicare** i **metodi statistici**
 - occorre saper **leggere** e **interpretare** report contenenti informazioni statistiche

Perché studiare la statistica?

Nella **vita di tutti i giorni** affrontiamo un flusso continuo di informazioni quantitative:

- **messaggi pubblicitari**: il 70% dei consumatori sceglie ...
- **notizie**: evoluzione della pandemia e del numero di contagi
- **sondaggi**: il 50% degli italiani pensa che ...
- **social network**: lo youtuber con 5 milioni di followers...

Perché studiare la statistica?

Nella **vita di tutti i giorni** affrontiamo un flusso continuo di informazioni quantitative:

- **messaggi pubblicitari**: il 70% dei consumatori sceglie ...
- **notizie**: evoluzione della pandemia e del numero di contagi
- **sondaggi**: il 50% degli italiani pensa che ...
- **social network**: lo youtuber con 5 milioni di followers...

In tutti questi casi la **statistica** ci aiuta a:

- **comprendere** e **interpretare** i fenomeni che ci circondano
- effettuare **scelte informate** e **consapevoli**

I dati

La **raccolta delle informazioni** rappresenta un punto cruciale dell'analisi statistica, essa permette di ottenere i **dati** :

- **dati**: l'**insieme** di tutte le **osservazioni** sulle caratteristiche di interesse in uno studio

I dati

La **raccolta delle informazioni** rappresenta un punto cruciale dell'analisi statistica, essa permette di ottenere i **dati** :

- **dati**: l'**insieme** di tutte le **osservazioni** sulle caratteristiche di interesse in uno studio

Ad esempio, per ciascun individuo (**unità statistica**) di uno studio i **dati** potrebbero avere la seguente struttura:

Caratteristiche	Osservazioni
Legalizzazione marijuana	Non favorevole
Partito politico	Repubblicano
Anni di istruzione	14 anni
Frequenza funzioni religiose	Una volta a settimana
Reddito annuo	Tra i 40 e i 60 mila \$
Stato civile	Coniugata
Sesso	Femminile

I dati

I **dati** raccolti devono poi essere esaminati **nel modo opportuno** per comprendere come queste caratteristiche sono in **relazione** tra loro:

- le persone che frequentano con maggiore frequenza le funzioni religiose tendono a essere maggiormente conservatrici?
- le persone con più anni di istruzione tendono a avere un reddito annuo maggiore?

Per ottenere i **dati** possono essere utilizzati diversi **metodi**:

- indagini, esperimenti, osservazioni dirette dei comportamenti

Spesso si analizzano dati raccolti per **altre finalità**:

- rapporti delle forze di polizia, archivi ospedalieri, censimenti ...

Le raccolte di **dati** in archivi sono dette **database** e spesso sono disponibili su **internet** (es. <http://dati.istat.it/>)

Indagine Multiscopo sulle famiglie: aspetti della vita quotidiana

L'indagine **multiscopo sulle famiglie: aspetti della vita quotidiana** dell'ISTAT:

- si inserisce nel sistema delle indagini multiscopo sulle famiglie
 - **diversi temi:** scuola, lavoro, vita familiare e di relazione, abitazione e zona in cui si vive, tempo libero, partecipazione politica e sociale, salute, stili di vita, . . .
- è a cadenza annuale ed è eseguita su un campione di circa 25.000 famiglie distribuite in 800 comuni
- consente di avere informazioni sulle abitudini dei cittadini e il loro livello di soddisfazione rispetto al funzionamento dei servizi
- informazioni sono disponibili online sul sito ISTAT:
<https://www.istat.it/it/archivio/91926>

La Statistica?

Cosa è la statistica?

- la **Statistica** è l'insieme delle **metodologie** finalizzate alla **raccolta** e all'**analisi** dei **dati**
- L'obiettivo primario della Statistica è **spiegare la variabilità** osservata nei fenomeni studiati.

La Statistica?

Cosa è la statistica?

- la **Statistica** è l'insieme delle **metodologie** finalizzate alla **raccolta** e all'**analisi** dei **dati**
- L'obiettivo primario della Statistica è **spiegare la variabilità** osservata nei fenomeni studiati.

Per far ciò è necessario **osservare**, ovvero **misurare**:

- le diverse caratteristiche o stati (**fenomeni qualitativi**)
- le diverse intensità (**fenomeni quantitativi**)

con le quali il fenomeno, o i fenomeni studiati si manifestano sulle (**unità statistiche**) in esame.

La Statistica?

Più precisamente, la **statistica** è l'insieme dei metodi per:

1 - Progettare: pianificare come devono essere raccolti i dati

La Statistica?

Più precisamente, la **statistica** è l'insieme dei metodi per:

- 1 - **Progettare**: pianificare come devono essere raccolti i dati
- 2 - **Descrivere**: organizzazione dei dati in **distribuzioni di frequenza** e **sintesi** tramite appositi indici numerici o rappresentazione grafiche

La Statistica?

Più precisamente, la **statistica** è l'insieme dei metodi per:

- 1 - **Progettare**: pianificare come devono essere raccolti i dati
- 2 - **Descrivere**: organizzazione dei dati in **distribuzioni di frequenza** e **sintesi** tramite appositi indici numerici o rappresentazione grafiche
- 3 - **Inferire**: formulare previsioni basate sui dati raccolti
 - ...
- n - **Diffusione dei risultati**

Cosa è la Statistica?

Il termine **Progettare** è riferito alla pianificazione di come dovranno essere ottenuti i dati:

- come saranno **selezionati** gli individui da intervistare?
- come dovrà essere **strutturato** il questionario da somministrare?
- quali **caratteristiche** ci interessano?
- quali **strumenti** possiamo utilizzare per raccogliere le informazioni?

Cosa è la Statistica?

Il termine **Descrivere** si riferisce ai modi in cui si possono sintetizzare i dati per comprendere le informazioni che contengono:

- consideriamo un'analisi sul **numero di figli desiderati** basata sull'indagine multiscopo
- i **dati grezzi** sono, solitamente, una lista completa delle osservazioni registrate per ogni unità rilevata
 - una lista di circa 25.000 osservazioni (una per famiglia)
- Come possiamo presentare i dati **efficacemente**?
 - riportare il numero medio di figli desiderati o la risposta più frequente
 - grafici, tabelle e sintesi numeriche → **statistiche descrittive**

Cosa è la Statistica?

Il termine **Descrivere** si riferisce ai modi in cui si possono sintetizzare i dati per comprendere le informazioni che contengono:

- consideriamo un'analisi sul **numero di figli desiderati** basata sull'indagine multiscopo
- i **dati grezzi** sono, solitamente, una lista completa delle osservazioni registrate per ogni unità rilevata
 - una lista di circa 25.000 osservazioni (una per famiglia)
- Come possiamo presentare i dati **efficacemente**?
 - riportare il numero medio di figli desiderati o la risposta più frequente
 - grafici, tabelle e sintesi numeriche → **statistiche descrittive**

L'analisi descrittiva deve **sintetizzare** informazioni complesse in un **formato facilmente comprensibile**:

- ... prendiamo ad esempio la campagna di Napoleone in Russia

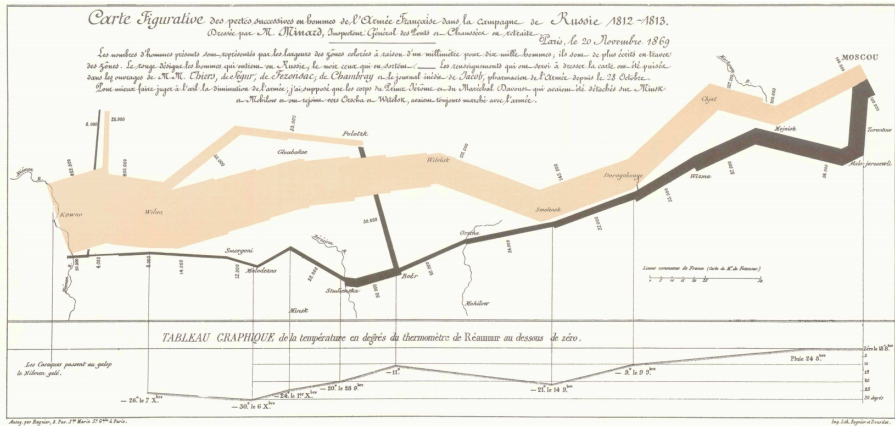


Figure: Mappa Figurativa delle perdite successive in uomini dell'Armata Francese nella Campagna di Russia 1812-1813, M. Charles Joseph Minard (1861)

Dalla legenda: Il numero degli uomini presenti è rappresentato dalla larghezza delle zone colorate in ragione di un millimetro per 10.000 uomini; le cifre sono anche riportate in perpendicolare. Il rosso rappresenta gli uomini entranti in Russia, il nero quelli che ne escono. I dati impiegati per disegnare la mappa sono stati ricavati . . .

Cosa è la Statistica?

Il termine **Inferire** si riferisce alle previsioni fatte attraverso i dati:

- immaginiamo che il 12.4% dei rispondenti all'indagine multiscopo ha riferito di desiderare un figlio
- possiamo utilizzare questa informazione per prevedere la percentuale di adulti italiani che desiderava avere un figlio al momento in cui è stata realizzata l'indagine?

Cosa è la Statistica?

Il termine **Inferire** si riferisce alle previsioni fatte attraverso i dati:

- immaginiamo che il 12.4% dei rispondenti all'indagine multiscopo ha riferito di desiderare un figlio
- possiamo utilizzare questa informazione per prevedere la percentuale di adulti italiani che desiderava avere un figlio al momento in cui è stata realizzata l'indagine?

Le previsioni fatte utilizzando i dati sono chiamate **inferenze statistiche**

- **diversi esempi:** numero di contagi da COVID-19, riscaldamento globale, rendimenti finanziari, esami universitari ...

La lezione

1 Informazioni

2 Introduzione alla metodologia statistica

- Cosa è la Statistica?
- **Statistica Descrittiva e Statistica Inferenziale**
- Popolazioni e Campioni

3 Le variabili statistiche

- Introduzione
- Le variabili e la loro misurazione

Statistica Descrittiva e Statistica Inferenziale

I metodi dell'**analisi statistica** dei dati vengono classificati in due grandi categorie:

- **Statistica Descrittiva** → l'analisi viene definita **descrittiva** quando il suo obiettivo principale è quello di **descrivere e sintetizzare** i dati
- **Statistica Inferenziale** → l'analisi viene definita **inferenziale** quando il suo obiettivo principale è quello di fare **previsioni** sulla base dei dati raccolti

Statistica Descrittiva e Statistica Inferenziale

I metodi dell'**analisi statistica** dei dati vengono classificati in due grandi categorie:

- **Statistica Descrittiva** → l'analisi viene definita **descrittiva** quando il suo obiettivo principale è quello di **descrivere e sintetizzare** i dati
- **Statistica Inferenziale** → l'analisi viene definita **inferenziale** quando il suo obiettivo principale è quello di fare **previsioni** sulla base dei dati raccolti

... Ma perché ci serve fare **previsioni**?

- due concetti importanti: **popolazione** e **campione**

La lezione

1 Informazioni

2 Introduzione alla metodologia statistica

- Cosa è la Statistica?
- Statistica Descrittiva e Statistica Inferenziale
- Popolazioni e Campioni

3 Le variabili statistiche

- Introduzione
- Le variabili e la loro misurazione

Popolazioni e Campioni

Si definisce come **unità statistica** l'unità elementare su cui vengono osservati alcune informazioni:

- persone fisiche, famiglie, scuole, aziende, città, ...

Popolazioni e Campioni

Si definisce come **unità statistica** l'unità elementare su cui vengono osservati alcune informazioni:

- persone fisiche, famiglie, scuole, aziende, città, ...

La **popolazione** è costituita dal **totale** delle unità statistiche di interesse di uno studio

- esempio: **tutte** le scuole italiane, **tutte** le persone fisiche residenti in Italia, **tutte** le ...

Popolazioni e Campioni

Si definisce come **unità statistica** l'unità elementare su cui vengono osservati alcune informazioni:

- persone fisiche, famiglie, scuole, aziende, città, ...

La **popolazione** è costituita dal **totale** delle unità statistiche di interesse di uno studio

- esempio: **tutte** le scuole italiane, **tutte** le persone fisiche residenti in Italia, **tutte** le ...

Il **campione** è un **sottoinsieme** della popolazione di riferimento dello studio

Popolazioni e Campioni

L'obiettivo principale di qualunque ricerca è quello di **acquisire conoscenze** sulla **popolazione**:

- **problema:** spesso non è possibile ottenere le informazioni su tutti i soggetti di interesse
 - costi, tempo, difficoltà di accesso alle informazioni, ...

Popolazioni e Campioni

L'obiettivo principale di qualunque ricerca è quello di **acquisire conoscenze** sulla **popolazione**:

- **problema**: spesso non è possibile ottenere le informazioni su tutti i soggetti di interesse
 - costi, tempo, difficoltà di accesso alle informazioni, ...
- **soluzione (parziale)**: usiamo i dati del campione per fare **inferenza** sulle caratteristiche della popolazione
 - problemi legati alla selezione del campione e alla generalizzabilità dei risultati della ricerca

Popolazioni e Campioni

L'obiettivo principale di qualunque ricerca è quello di **acquisire conoscenze sulla popolazione**:

- **problema**: spesso non è possibile ottenere le informazioni su tutti i soggetti di interesse
 - costi, tempo, difficoltà di accesso alle informazioni, ...
- **soluzione (parziale)**: usiamo i dati del campione per fare **inferenza** sulle caratteristiche della popolazione
 - problemi legati alla selezione del campione e alla generalizzabilità dei risultati della ricerca

Quindi:

- la **statistica descrittiva** sintetizza le informazioni raccolte in formati semplici limitando la perdita di informazioni e le distorsioni
- la **statistica inferenziale** fa delle previsioni su una popolazione sulla base delle informazioni raccolte su un campione selezionato da essa

Parametri e Statistiche

Quando si utilizzano le tecniche **statistiche inferenziali** è fondamentale comprendere la distinzione tra **parametro** e **statistica**:

Parametri e Statistiche

Quando si utilizzano le tecniche **statistiche inferenziali** è fondamentale comprendere la distinzione tra **parametro** e **statistica**:

- **parametro**: sintesi numerica di dati di **popolazione**
 - il numero medio di figli desiderati da **tutta la popolazione italiana**
 - numero medio → **sintesi numerica**
- **statistica**: sintesi numerica di dati di **campionari**
 - il numero medio di figli desiderati dai soggetti che fanno parte **del campione**
 - numero medio → **sintesi numerica**

Parametri e Statistiche

Quando si utilizzano le tecniche **statistiche inferenziali** è fondamentale comprendere la distinzione tra **parametro** e **statistica**:

- **parametro**: sintesi numerica di dati di **popolazione**
 - il numero medio di figli desiderati da **tutta la popolazione** italiana
 - numero medio → **sintesi numerica**
- **statistica**: sintesi numerica di dati di **campionari**
 - il numero medio di figli desiderati dai soggetti che fanno parte **del campione**
 - numero medio → **sintesi numerica**

Il nostro **principale obiettivo** è quello di conoscere il valore assunto dai **parametri della popolazione**

- il campione e le statistiche sono **utili** in quanto ci permettono di fare **inferenza** sul **parametro incognito** di interesse

Esempio: Hai fiducia nella gente?

Nell'indagine Multiscopo Aspetti della Vita Quotidiana è stata formulata la seguente **domanda**:

- lei pensa che **ci si possa fidare** della maggior parte della gente oppure **bisogna stare molto attenti**?

I dati dell'indagine:

- **popolazione**: tutte le famiglie italiane
- **campione**: 20.275 intervistati (uno per famiglia)
- **sintesi numerica**: il 66% degli intervistati del campione ha risposto **bisogna stare molto attenti**

Cosa ci interessa:

- le opinioni dell'intera **popolazione** italiana
- la statistica calcolata nel campione è la nostra fonte di informazioni sul **parametro incognito** della popolazione di interesse

Esempio: Hai fiducia nella gente?

I metodi di **statistica inferenziale** ci permettono di formulare la previsione:

- la **percentuale di adulti** che non ha fiducia nella gente è **compresa** fra il 64% e il 68%
 - la **statistica** ha un “margine di errore” del 2%
- **statistica**: percentuale di adulti nel **campione** che non si fida della gente
- **parametro**: percentuale di adulti nella **popolazione** che non si fida della gente
- **percentuale di adulti** → **statistica descrittiva**

Se, invece, si hanno i dati su **tutta la popolazione** si possono calcolare i **valori esatti** dei parametri senza utilizzare **metodi inferenziali**:

- es. censimenti, dati amministrativi, ...

Le variabili statistiche

La lezione

1 Informazioni

2 Introduzione alla metodologia statistica

- Cosa è la Statistica?
- Statistica Descrittiva e Statistica Inferenziale
- Popolazioni e Campioni

3 Le variabili statistiche

- Introduzione
- Le variabili e la loro misurazione

Introduzione

Riprendiamo due concetti:

- **Popolazione**: l'insieme di **tutti i soggetti** di interesse di uno studio
- **Campione**: **sottoinsieme** della popolazione di riferimento

Utilizziamo il **campionamento** per effettuare un'indagine su una **popolazione** utilizzando **una parte** di essa. Perché?

- contenere i costi e i tempi di realizzazione della ricerca
- avere la possibilità di replicare la ricerca nel breve periodo

Quindi, in ogni indagine, diventa fondamentale:

- **selezionare** i soggetti di studio → **Campionamento**
- **misurare** le **caratteristiche** dei soggetti → **Variabili e misurazione**

La lezione

1 Informazioni

2 Introduzione alla metodologia statistica

- Cosa è la Statistica?
- Statistica Descrittiva e Statistica Inferenziale
- Popolazioni e Campioni

3 Le variabili statistiche

- Introduzione
- Le variabili e la loro misurazione

Le variabili e la loro misurazione

Variabile: una **caratteristica** che assume **valori diversi** tra i **soggetti** di un campione o una popolazione.

Modalità: i diversi valori assunti dalla variabile in corrispondenza di ogni unità statistica.

- **sex:** maschio o femmina
 - due valori, non posso dire che uno sia maggiore dell'altro
- **numero di auto:** 0, 1, 2, ...
 - numeri interi, non posso avere 0,2 auto
- **reddito:** 'infiniti' valori
 - posso osservare sia redditi inferiori a 0 (perdite) che numeri decimali

Le variabili e la loro misurazione

Le variabili precedenti differiscono tra loro per la **scala di misura**:

- ogni variabile assume diversi valori ed è importante capire come **misurare** e **leggere** i valori osservati

Le modalità di una variabile devono essere:

- **Esaustive**: devono rappresentare **tutti i modi** di manifestarsi del carattere
- **Non sovrapposte**: a ogni unità si può associare **una e una sola** modalità

Variabili quantitative e categoriali

Le variabili possono essere suddivise in due sottogruppi principali:

- **Variabili quantitative:** i valori osservati sono **numeri**
 - es. età, reddito, numero di auto, ...
 - età = 10 anni
 - reddito = 14.450 euro
 - numero di auto = 2
- **Variabili categoriali:** i valori osservati sono rappresentati da **categorie**, o **parole**
 - es. sesso, titolo di studio, religione, frutto preferito, ...
 - sesso = femmina
 - titolo di studio = laureata
 - religione = cattolica
 - frutto preferito = ciliegie

Variabili quantitative e categoriali

Le variabili **quantitative** sono di due tipi:

- **Discrete**: i valori sono **numeri** distinti (di solito interi)
 - es. numero di auto: 0, 1, 2, 3, ...
- **Continue**: può assumere come valori ogni possibile numero reale incluso in un continuum '**infinito**'
 - es. età: 1 anno, 2 mesi, 3 giorni, ...
- la **differenza** tra due modalità è **informativa**

Variabili quantitative e categoriali

Le variabili **quantitative** sono di due tipi:

- **Discrete**: i valori sono **numeri** distinti (di solito interi)
 - es. numero di auto: 0, 1, 2, 3, ...
- **Continue**: può assumere come valori ogni possibile numero reale incluso in un continuum '**infinito**'
 - es. età: 1 anno, 2 mesi, 3 giorni, ...
- la **differenza** tra due modalità è **informativa**

Le variabili **categoriali** sono di due tipi:

- **Ordinali**: i valori possono essere ordinati
 - es. titolo di studio: laurea è maggiore di diploma ...
- **Nominali**: i valori non possono essere ordinati
 - es. colori: rosso non è maggiore di blu ...

Variabili discrete e continue

Consideriamo la variabile **discreta** numero di auto per la famiglia:

- **possiamo** usare solo numeri **interi** (1, 2, 3, 4, ...)
- **non possiamo** rilevare osservazioni il cui valore sia un numero decimale (1.45 o 2.3)
- la **differenza** tra due valori è **informativa**:
 - se una famiglia ha 3 auto e un'altra ne ha 1 possiamo dire che la differenza è di 2 auto

Variabili discrete e continue

Consideriamo la variabile **discreta** numero di auto per la famiglia:

- **possiamo** usare solo numeri **interi** (1, 2, 3, 4, ...)
- **non possiamo** rilevare osservazioni il cui valore sia un numero decimale (1.45 o 2.3)
- la **differenza** tra due valori è **informativa**:
 - se una famiglia ha 3 auto e un'altra ne ha 1 possiamo dire che la differenza è di 2 auto

Consideriamo la variabile **continua** età in anni:

- **può** assumere valori che vanno da 0 a più di 100
- esempio: consideriamo due studenti che hanno 20 anni
 - in realtà hanno un'età diversa se consideriamo i mesi, i giorni, i secondi, ecc
- la **differenza** tra due valori è **informativa**:
 - se tra due studenti c'è una differenza di 73 giorni allora ci sarà una differenza di $73/365 = 0,2$ anni

Variabili discrete e continue

Le variabili quantitative si possono anche distinguere in:

- **Scala a intervalli:** per queste variabili non esiste uno **zero assoluto**.
 - Es.: la temperatura misurata in gradi centigradi ($^{\circ}\text{C}$).
 - Lo zero utilizzato è uno zero **convenzionale**, non ha quindi senso dire che 40°C sono il doppio di 20°C .

Variabili discrete e continue

Le variabili quantitative si possono anche distinguere in:

- **Scala a intervalli:** per queste variabili non esiste uno **zero assoluto**.
 - Es.: la temperatura misurata in gradi centigradi ($^{\circ}\text{C}$).
 - Lo zero utilizzato è uno zero **convenzionale**, non ha quindi senso dire che 40°C sono il doppio di 20°C .
- **Scala a rapporti:** per queste variabili esiste uno **zero assoluto**.
 - Es.: Peso, reddito, età..
 - Si può dire che Tizio guadagni **il doppio** di Caio.

Variabili discrete e continue

Le variabili quantitative si possono anche distinguere in:

- **Carattere trasferibile:** quando un'unità statistica **può cedere parte del carattere posseduto** ad un'altra unità.
 - Es.: il reddito è un carattere trasferibile.
- **Carattere non trasferibile:** quando un'unità statistica **non può cedere parte del carattere posseduto** ad un'altra unità.
 - (*Purtroppo*) l'età, il peso, ...

Variabili discrete e continue

La differenza tra variabili **continue** e **discrete** è solo **apparente**:

- nella realtà tutte le variabili **continue** vengono trattate come **discrete**
 - esempio: età misurata in anni, distanza in km, ...
- stabilendo a quale unità o cifra decimale fermarci, compiamo un'operazione di **approssimazione**
 - da infiniti valori passiamo a un numero **finito** di valori

Variabili discrete e continue

La differenza tra variabili **continue** e **discrete** è solo **apparente**:

- nella realtà tutte le variabili **continue** vengono trattate come **discrete**
 - esempio: età misurata in anni, distanza in km, ...
- stabilendo a quale unità o cifra decimale fermarci, compiamo un'operazione di **approssimazione**
 - da infiniti valori passiamo a un numero **finito** di valori

In sintesi:

- le variabili sono **quantitative** o **categoriali**
- le **categoriali** sono sempre **discrete**
- le **categoriali** si dividono in **nominali** (categorie non ordinate) e **ordinali** (categorie ordinate)
- le **quantitative** possono essere **continue** o **discrete**

Esempio

Nome	Età	Sesso	Tit. studio	Attività	Peso (kg)	Punteggio
Mario	32	M	Laurea	Occupato/a	72	65
Alessia	39	F	Laurea	Occupato/a	55	53
Giovanni	46	M	Diploma	Disoccupato/a	79	53
Francesco	28	M	Diploma	Studente	63	78
Claudia	51	F	Diploma	Casalingo/a	64	30
Flavio	60	M	Diploma	Casalingo/a	62	21
Fabio	35	M	Dottorato	Occupato/a	71	85
Elisa	38	F	Diploma	Occupato/a	64	45
Clara	41	F	Diploma	Casalingo/a	65	28

Esempio

Nome	Età	Sesso	Tit. studio	Attività	Peso (kg)	Punteggio
Mario	32	M	Laurea	Occupato/a	72	65
Alessia	39	F	Laurea	Occupato/a	55	53
Giovanni	46	M	Diploma	Disoccupato/a	79	53
Francesco	28	M	Diploma	Studente	63	78
Claudia	51	F	Diploma	Casalingo/a	64	30
Flavio	60	M	Diploma	Casalingo/a	62	21
Fabio	35	M	Dottorato	Occupato/a	71	85
Elisa	38	F	Diploma	Occupato/a	64	45
Clara	41	F	Diploma	Casalingo/a	65	28

- **Variabili nominali:** Nome, Sesso, Attività
- **Variabili ordinali:** Titolo di studio
- **Variabili quantitative:** Età, Peso, Punteggio.