Probabilità e Statistica

Alex Narder

September 26, 2022

1 Contenuti del corso

2 Introduzione

Per estrarre informazioni dai dati dobbiamo processarli adeguatamente;
La statistica ci fornisce quegli stumenti per estrarre le informazioni dai dati, uno statistico sa:

- combinare informazioni di diverso tipo
- valutare l'affidabilità
- sintetizzare e presentare molti dati
- costruire modelli
- calcolare previsioni e formulare ipotesi di decisione

La statistica non è l'unico strumento usato per analizzare i dati, ma è quello più adatto in presenza di incertezze.
Anche l'informatica svolge un ruolo fondamentale nel salvataggio e nella gestione dei dati.

3 Termini della statistica

- La **popolazione di riferimento**, è un insieme di elementi chiamati **unità statistiche**.
- I **dati** sono il risultato di rilevazioni o misurazioni.
- La **statistica** ci permette di estrarre le informazioni dai dati, generando nuove conoscenze o ipotesi di decisione.
- Ogni caratteristica rilevata sulle unità statistiche si chiama **variabile** e i dati corrispondenti a ogni variabile sono le **realizzazioni**.
- Se le variabili non sono rilevate su tutte le statistiche, il sottoinsieme della popolazione oggetto della rilevazione è chiamato il **campione**.

4 Tipi di variabili

Una variabile è **qualitativa** o **categorica** quando i suoi possibili valori o modalità prendono la forma di aggettivi o di altre espressioni verbali.
Le variabili qualitative possono essere:

- Sconnesse
- Ordinali

Le variabili categoriche possono essere:

- Discrete o intere
- Continue o reali

5 Calcolo della Probabilità

Perché l'inferenza porti a risultati sensati, bisogna che sia noto il legame fra popolazione e campione.

• Il calcolo delle probabilità fornisce i modelli matematici utili per descrivere la relazione fra campione e popolazione.
• Il calcolo delle probabilità è lo strumento necessario per l'inferenza. Permette di quantificare gli errori che commettiamo nel passaggio dal particolare (campione) al generale (popolazione).

6 Esempi dipprobabilità

- Lancio di 1 dado (normale), qual è la probabilità che esca 6? 1/6
- Stessa domanda ma il dado ha 20 facce. 1/20

Qual è la probabilità che il risultato sia 2 o 6, sempre in riferimento della prima domanda? La probabilità che esca è 2/6.

- Pensiamo di avere delle antenne in linea, se due antenne di fila non funzionano allora il sistema non funziona. Se una antenna è rotta ma quella successiva no il sistema funziona lo stesso perchè il segnale passa alla successiva.
Sapendo che n antenne sono difettose, qual è la probabilità che il sistema funzioni?
$m <= n$

Se n = 4 e m = 2:
1 → se l'antenna funziona
0 → se l'antenna non funziona

n rappresenta il numero delle antenne, mentre m è il nunero delle antenne rotte, quindi nel momento in cui ho un gruppo di antenne 1100 sò che le prime 2 funzionano, mentre le ultime 2 non funzionano e sono consecutive, quindi il sistema non funziona. Nel secondo caso invece le antenne non funzionanti non sono consecitive, 0110 quindi è un sistema che funziona. Vado a prendere solo i casi in cui so che solo 2 non funzionano, perciò 1111, 0000 oppure 1110 ecc ecc non vado a considerarli. →

1100 non funziona
0110 funziona
1010 funziona
0011 non funziona
0101 funziona
1001 non funziona

probabilità richiesta: 3/6 = 1/2

totale: 6, quindi sono 3 funzionanti

7 Contare le probabilità

Principio fondamentale del calcolo combinatorio:
Se una scelta può essere fatta in m1 modi diversi e un'altra scelta può essere fatta in m2 modi diversi, allora esistono in totale

m1 * m2 possibilità di scelta

esempio:
10 cavalieri e 12 dame partecipano a un ballo. Ci sono $10 \times 12 = 120$ possibili coppie danzanti.

Principio fondamentale generalizzato:
Se ciascuna delle r scelte successive può essere fatta in m_i modi, allora esistono rispettivamente in totale:

$$\prod_{i=1}^{r} m_i = m_1 * ... * m_r$$

possibilità di scelta esempio:
Una commissione parlamentare deve essere composta da un membro del partito A, che conta 10 rappresentanti, da un membro del partito B, che conta 15 rappresentanti, e da un membro del partito C, che conta 2 rappresentanti. Ci sono in totale $10 \times 15 \times 2 = 300$ possibili commissioni parlamentari.

8 Disposizioni

Consideriamo un insieme di n elementi. Una **disposizione** di r di essi è una scelta ordinata di r elementi tra quegli n.

• Si distinguono le disposizioni con **ripetizione** da quelle **semplici** (o **senza ripetizione**), a seconda che uno stesso elemento possa essere scelto più di una volta:

• Le disposizioni con **ripetizione** di n elementi presi r alla volta sono in numero di

$$\prod_{i=1}^{r} n = n^r$$

esempio:
le parole lunghe due lettere che si possono comporre con le lettere I, L, A sono 32 = 9: II, IL, IA, LI, LL, LA, AI, AL, AA.

• Le **disposizioni semplici o senza ripetizione** di n elementi presi r alla volta sono in numero di

$$n * (n-1) \; ... \; *(n-r+1)$$

esempio:
le parole di due lettere diverse che si possono comporre con le lettere I, L, A sono $3 \times 2 = 6$: IL, IA, LI, LA, AI, AL.

9 Campionamento da un'urna

Il **campionamento casuale da un'urna** è un'estrazione di palle da un'urna. Può essere fatto **con o senza reintroduzione**.

- Per **casuale** si intende dire che prima di ogni estrazione l'urna viene mescolata appropriatamente per essere riportata a una condizione di irriconoscibilità delle palle. Un'operazione del genere viene fatta, ad esempio, per le estrazioni del lotto.

- La **reintroduzione** fa invece riferimento al fatto di rimettere nell'urna ciascuna palla subito dopo averla estratta e averne registrate le caratteristiche di interesse, per esempio il suo numero o il suo colore.

Quindi:
- Se un'urna contiene n palle distinguibili e r palle vengono estratte con reintroduzione, ci sono n^r possibilità di estrazione.
- Se un'urna contiene n palle distinguibili e r palle vengono estratte senza reintroduzione, le estrazioni possibili sono $n * (n-1) \; ... \; *(n-r+1)$.

10 Permutazioni

Le disposizioni semplici di n elementi presi n alla volta si chiamano anche **permutazioni** perché rappresentano tutti i modi in cui n elementi possono essere ordinati in fila. Esse sono in numero di

$$n * (n-1) ... * 2 * 1 =: n!$$

Il simbolo speciale n! che rappresenta questa quantità si legge n **fattoriale**.

Esempio 1:
Le permutazioni delle lettere I, L, A sono $3 \times 2 \times 1 = 3! = 6$:
ILA, IAL, LIA, LAI, AIL, ALI.

Esempio 2:
Supponiamo di fare due file, i 6 maschietti a destra e le 4 femminucce a sinistra. Ci sono 6! possibili file di maschietti e 4! file di femminucce possibili. Quindi, dal principio fondamentale del calcolo combinatorio, in tutto ci sono $6! \times 4! = 17280$ possibili file.

11 Combinazioni

In generale; Un sottoinsieme di numerosità r scelto da un insieme con n elementi si chiama **combinazione** di n elementi presi r alla volta.
Il numero di combinazioni di n elementi r alla volta è

$$\frac{n * (n-1) ... (n-r+1)}{r!} =: \binom{n}{r}$$

esempio:
La professoressa di matematica interroga ogni lunedì 10 studenti da una classe di 25. Esistono per lei $\binom{25}{10} = 3.268.760$ possibilità di scelta.

→ Il nome coefficente binomiale deriva dalla seguente espressione:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

detta formula del **Binomio di Newton**.

Esempio:

$$(a+b)^2 = \binom{2}{0} a^2 b^0 + \binom{2}{1} ab + \binom{2}{2} a^0 b^2 = a^2 + 2ab + b^2$$