Statistica I

Unità E: barche turistiche & crescita di batteri

Tommaso Rigon

Università Milano-Bicocca

Anno Accademico 2020-2021

Descrizione del problema

- Una barca turistica offre un servizio di navigazione nel Lago di Como.
- Il proprietario propone una visita della durata massima di 3 ore, partendo da Como, il cui itinerario può essere deciso dal cliente.
- Quando la barca procede **controvento**, si muove a **6.3 km/h**. Viceversa, a **favore di vento** la stessa barca procede a **14.7 km/h**.
- Un turista quindi chiede: "*Sarebbe possibile visitare la villa di George Clooney a Laglio*?". Laglio è un paese vicino a Como, distante circa 14.5km.
- **Domande del turista**. Qual è la velocità media? È possibile percorrere questo viaggio in meno di tre ore?

La soluzione "diretta"

- Anzitutto, è bene ricordare che

$$(\text{velocità}) = \frac{(\text{distanza})}{(\text{tempo})}.$$

- Supponendo che il vento non cambi direzione, la barca viaggerà per 14.5km controvento e per altri 14.5km con il vento a favore.

- **Favore di vento**. La barca viaggerà per una distanza di 14.5km ad una velocità di 14.7km/h. Invertendo la formula precedente, si ottiene

$$(\text{tempo a favore di vento}) = \frac{(\text{distanza a favore di vento})}{(\text{velocità a favore di vento})} = \frac{14.5 \text{ km}}{14.7 \text{ km/h}} \approx 0.986 \text{ h}.$$

- **Controvento**. La barca viaggerà per una distanza di 14.5km ad una velocità di 6.3km/h. Quindi

$$(\text{tempo controvento}) = \frac{(\text{distanza controvento})}{(\text{velocità controvento})} = \frac{14.5 \text{ km}}{6.3 \text{ km/h}} \approx 2.302 \text{ h}.$$

La soluzione "diretta"

- Quindi, possiamo calcolare il **tempo totale** necessario per completare l'itinerario:

$$(\text{tempo totale}) = 0.986\,\text{h} + 2.302\,\text{h} = 3.288\,\text{h}.$$

  Di conseguenza, non è possibile visitare la villa di George Clooney in meno di 3 ore.

- Inoltre, la **velocità media** complessiva si ottiene tramite

$$(\text{velocità media totale}) = \frac{(\text{distanza totale})}{(\text{tempo totale})} = 29\,\text{km}/3.288\,\text{h} = 8.82\,\text{km}/\text{h}.$$

La soluzione statistica sbagliata

- La seguente soluzione è **sbagliata** da un punto di vista fisico.
- Si potrebbe (erroneamente) ragionare come segue. Dal momento che le distanze sono uguali, si potrebbe considerare la media aritmetica delle velocità, ovvero:

$$\bar{x} = \frac{14.7\,\text{km/h} + 6.3\,\text{km/h}}{2} = 10.5\,\text{km/h}.$$

- Questo valore tuttavia **non** coincide con la vera velocità media, calcolata nella slide precedente.
- Usando le formule di prima, si potrebbe quindi giungere al valore errate 2.762 km/h per il tempo totale.
- **Nota**. La media aritmetica delle velocità **non** restituisce la velocità media complessiva né il tempo totale corretto.

La soluzione statistica corretta

- La soluzione corretta prevede invece l'uso della **media armonica**.
- **Nota**. La media armonica delle velocità coincide con la velocità media.
- In questo esempio infatti si ottiene le velocità media

$$\mathbb{A} = 2\left(\frac{1}{14.7\ \text{km/h}} + \frac{1}{6.3\ \text{km/h}}\right)^{-1} = 8.82\ \text{km/h},$$

che coincide con quella precedentemente calcolata e da cui si può rapidamente ottenere la distanza totale.

- Come mai questo accade? Si noti che possiamo riscrivere

$$\mathbb{A} = 2 \times 14.5\ \text{km}\left(\frac{14.5\ \text{km}}{14.7\ \text{km/h}} + \frac{14.5\ \text{km}}{6.3\ \text{km/h}}\right)^{-1},$$

che sono gli stessi conti effettuati nella soluzione "diretta".

La media armonica ponderata

- La media armonica può essere usata per calcolare la velocità media anche nel caso in cui le distanze non siano uguali.

- Un oggetto percorre le distanze $d_1, \dots, d_n$ con velocità $v_1, \dots, v_n$.

- La velocità complessiva del viaggio si ottiene tramite una **media armonica ponderata**, ovvero
$$\mathbb{A}_w = \frac{d_1 + \cdots + d_n}{d_1/v_1 + \cdots + d_n/v_n} = \frac{\sum_{i=1}^n d_i}{\sum_{i=1}^n d_i/v_i},$$
dove i pesi sono in questo caso le distanze $d_1, \dots, d_n$.

- **Esercizio**. Si supponga che barca considerata finora proceda per 10km a favore di vento e per 19km controvento.

- Qual è la velocità media? Quanto tempo occorre per l'intero viaggio?

- Si ottenga lo stesso risultato tramite la soluzione "diretta".

L'approccio di Chisini

- La media armonica ponderata si può anche ottenere seguendo il **criterio di Chisini**, che rende trasparente la sua scelta.

- La media cercata $\mathbb{A}_w$ è quel valore che, sostituito ai dati, lascia **invariato** il **tempo complessivo**.

- In altri termini, la funzione aggregatrice $g(\cdot)$ in questo caso è pari a

$$g(v_1, \dots, v_n) = \frac{d_1}{v_1} + \cdots + \frac{d_n}{v_n} = (\text{tempo complessivo}).$$

- La media secondo Chisini pertanto si ottiene risolvendo la seguente equazione:

$$\frac{d_1}{\mathbb{A}_w} + \cdots + \frac{d_n}{\mathbb{A}_w} = \frac{d_1}{v_1} + \cdots + \frac{d_n}{v_n},$$

  la cui soluzione coincide in effetti con la media armonica ponderata della slide precedente.

Descrizione del problema

- È stato osservato il processo di crescita di una popolazione di batteri, rilevando il numero di individui $x_0, \dots, x_4$, ad intervalli regolari di un'ora ciascuno, ai tempi $t = 0, \dots, 4$.

x_0	x_1	x_2	x_3	x_4
10	170	3290	42230	173510

- Si vuole conoscere il **tasso di crescita medio** della popolazione di batteri, ovvero la crescita percentuale media tra due istanti successivi.

- In simboli, chiamiamo il **tasso di crescita** per l'istante t-esimo

$$r_t = \frac{x_t - x_{t-1}}{x_{t-1}} \qquad \text{e pertanto} \qquad x_t = (1 + r_t)x_{t-1}, \qquad t = 1, 2, 3, 4.$$

- Abbiamo quindi descritto l'**evoluzione temporale** dei conteggi x_t. Per esempio, il valore finale x_4 si ottiene moltiplicando il valore iniziale x_0 come segue

$$x_4 = x_0(1 + r_1)(1 + r_2)(1 + r_3)(1 + r_4).$$

Tassi di crescita

- I tassi di crescita sono pari a

r_1	r_2	r_3	r_4
16	18.35	11.84	3.11

- In questo caso, la **media geometrica** sembra essere un indicatore sintetico appropriato, perché ha una chiara interpretazione evolutiva.

- Infatti, il **tasso di crescita medio** del periodo considerato è quel valore $\bar{r}$ tale che

$$(1+\bar{r})^4 = (1+r_1)(1+r_2)(1+r_3)(1+r_4)$$

oppure, equivalentemente, quel valore tale per cui

$$x_4 = x_0(1+\bar{r})^4.$$

- Questo corrisponde alla media geometrica degli $(1+r_t)$, ovvero

$$\mathbb{G} = (1+\bar{r}) = \sqrt[4]{16 \times 18.35 \times 11.84 \times 3.11} = 11.477,$$

Il tasso di crescita medio è quindi pari a $\bar{r} = 10.477$.

L'approccio di Chisini

- Anche in questo caso, la scelta della media $\bar{r}$ è naturale seguendo il **criterio di Chisini**.

- La media cercata $\bar{r}$ è quel valore che, sostituito ai dati, lascia **invariato** il **numero di batteri complessivo**.

- In altri termini, la funzione aggregatrice $g(\cdot)$ in questo caso è pari a

$$g(r_1, \dots, r_n) = x_0(1 + r_1)(1 + r_2)\cdots(1 + r_n) = (\text{numero complessivo di batteri } x_n),$$

seguendo lo **schema evolutivo** che abbiamo descritto.

- La media secondo Chisini pertanto si ottiene risolvendo la seguente equazione:

$$x_0(1 + \bar{r})(1 + \bar{r})\cdots(1 + \bar{r}) = x_0(1 + r_1)(1 + r_2)\cdots(1 + r_n),$$

la cui soluzione coincide in effetti con quella calcolata nella slide precedente, ovvero

$$\bar{r} = \sqrt[n]{\prod_{t=1}^{n}(1 + r_t)} - 1.$$