Modelli descrittivi, statistica e simulazione

Master per Smart Logistics specialist

Roberto Cordone
(roberto.cordone@unimi.it)

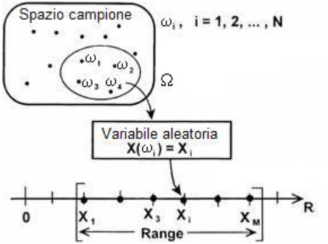
Variabile aleatoria

Variabile aleatoria X è una funzione che associa a ogni esito $\omega \in \Omega$ un valore reale $X(\omega) \in \mathbb{R}$

Esempi

- al lancio di un dado, il numero che compare sulla faccia superiore;
- al lancio di una moneta, 0 per la testa e 1 per la croce (o viceversa);
- all'arrivo di un treno, il numero di minuti di ritardo;
- all'arrivo di un ordine, l'ora esatta della chiamata;
- ...

Lo spazio campione Ω si trasforma in un sottoinsieme di numeri reali detto supporto o range $X(\Omega) = \{t \in \mathbb{R} : X(\omega_i) = t \text{ per ogni } \omega_i \in \Omega\}$



Vantaggi

Lavorare con i numeri reali è molto più comodo
che lavorare su elementi di un insieme astratto

- si può usare l'analisi matematica (somme, derivate, integrali)
- si possono usare tabelle
- si possono usare strumenti informatici come i fogli elettronici (Excel)

Ora vedremo come

Il passo fondamentale sarà spostare la definizione di probabilità

- dagli eventi $A \subseteq \Omega$
- agli insiemi di numeri reali $X(A) \subseteq \mathbb{R}$

Variabili aleatorie discrete e continue

In base al tipo di supporto, una variabile aleatoria si dice

- discreta quando $X(\Omega)$ è un insieme
  - finito
  - infinito numerabile
- continua quando $X(\Omega)$ è un insieme continuo

Probabilità per variabili discrete

Per le variabili discrete è facile passare da eventi a valori reali

- ci concentriamo sugli eventi in cui la variabile assume un dato valore, cioè gli eventi $(X = t)$ definiti come $\{\omega \in \Omega : X = t\}$
- per l'assioma della somma, la probabilità di un tale evento è

$$P(X = t) = \sum_{\omega \in \Omega : X(\omega) = t} p(\{\omega\})$$

- ora definiamo funzione di probabilità (discreta) della variabile X la probabilità dell'evento che la variabile X assuma ogni valore $t \in \mathbb{R}$

$$p_X(t) = \begin{cases} P(X = t) & \text{per } t \in X(\Omega) \\ 0 & \text{per } t \notin X(\Omega) \end{cases}$$

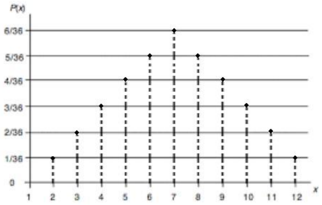
Funzione di probabilità discreta

La funzione di probabilità eredita le proprietà della probabilità di eventi

1. è positiva: $p_X(t) \geq 0$ per ogni $t \in \mathbb{R}$
2. ha somma 1: $\sum_{t \in X(\Omega)} p_X(t) \geq 0$

È sempre nulla, salvo in un insieme numerabile di punti dove fa dei "salti"

Esempio: sia X il risultato totale del lancio di due dadi



Probabilità per variabili continue

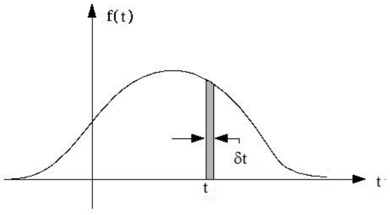
Per variabili continue non ha senso parlare di probabilità di un solo valore

- la probabilità che X assuma un singolo valore è sempre nulla
- si valuta la probabilità che il valore di X cada in un dato intervallo

Invece che una funzione di probabilità, avremo una

- funzione di densità di probabilità, che è il limite del rapporto fra probabilità e ampiezza di un intervallo di valori

$$f_X(t) = \lim_{\delta \to 0} \frac{P(t < X \leq t + \delta)}{\delta}$$



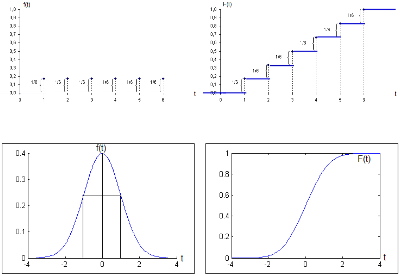
Se interessano gli eventi per cui la variabile X cade in un intervallo

$$t_1 < X \leq t_2$$

lo strumento adatto è la funzione di ripartizione:

$$F_X(t) = P(X \leq t)$$

- esiste per variabili aleatorie sia discrete sia continue
- se si considera Ω come popolazione statistica e X come carattere, F_X è la funzione di distribuzione cumulata del carattere

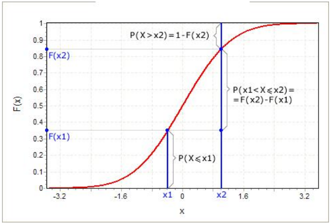


Proprietà

La funzione di ripartizione $F_X(t)$ di qualsiasi variabile aleatoria X

1. parte da 0: $\lim_{t \to -\infty} F_X(t) = 0$
2. arriva a 1: $\lim_{t \to +\infty} F_X(t) = 1$
3. cresce sempre: $F_X(t_1) \leq F_X(t_2)$ per ogni $t_1 \leq t_2$
4. se interessa la probabilità che X cada in un intervallo $(t_1, t_2]$, basta fare la differenza dei valori di F_X nei due estremi

$$P(t_1 < X \leq t_2) = F_X(t_2) - F_X(t_1)$$



Statistica descrittiva e probabilità

La statistica descrittiva e la teoria della probabilità sono parenti stretti

Si tratta di vedere

- lo spazio campione Ω come una popolazione statistica
- ciascun esito come un'unità statistica
- una variabile aleatoria come un carattere quantitativo (discreto o continuo)
- le frequenze relative come probabilità (interpretazione frequentista)

Da questa corrispondenza deriva che

- la funzione di ripartizione corrisponde alla distribuzione cumulata
- gli indici di posizione, variabilità, simmetria e forma si possono ridefinire pari pari per i valori della variabile aleatoria

Indici di sintesi

Media di una variabile aleatoria è la
media statistica della popolazione costituita dai valori della variabile, pesati con le loro probabilità o densità di probabilità

$$E(X) = \begin{cases} \sum_{i=1}^{m} t_i \cdot p_X(t_i) & \text{(caso discreto)} \\ \int_{-\infty}^{+\infty} t \cdot f_X(t)\, dt & \text{(caso continuo)} \end{cases}$$

Varianza di una variabile aleatoria è la
varianza statistica della popolazione costituita dai valori della variabile, pesati con le loro probabilità o densità di probabilità

$$V(X) = \begin{cases} \sum_{i=1}^{m} (t_i - E(X))^2 \cdot p_X(t_i) & \text{(caso discreto)} \\ \int_{-\infty}^{+\infty} (t - E(X))^2 \cdot f_X(t)\, dt & \text{(caso continuo)} \end{cases}$$

A che serve tutto questo?

Distribuzioni notevoli

Tutto questo serve perché

- dalla funzione di ripartizione si ricavano le probabilità degli eventi descrivibili come intervalli di valori
- dagli indici di sintesi si ricavano stime di probabilità (ad es., con la disuguaglianza di Čebišev)
- per opportune famiglie di funzioni di ripartizione, dagli indici di sintesi si ricostruisce la funzione, e quindi le probabilità

Quindi, sapendo che un fenomeno segue un dato tipo di distribuzione, pochi indici numerici bastano per fare previsioni probabilistiche

Vediamo alcune distribuzioni notevoli, e quando si usano

Distribuzione uniforme (discreta)

Si usa quando X è uno di m valori equiprobabili $(a_1, \ldots, a_m)$

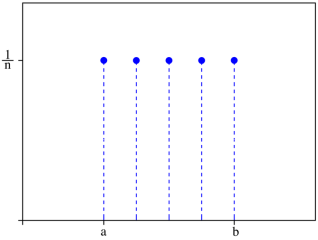
Ha funzione di probabilità discreta

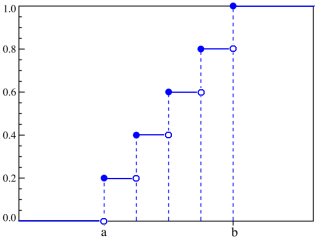
$$p_X(a_i) = \frac{1}{m} \text{ per ogni } i \in \{1, \ldots, m\}$$

e quindi funzione di ripartizione $F_X(t) = 0$ per $t < a_1$,

$$F_X(t) = \frac{i}{m} \text{ per } t \in \left[a_1 + \frac{i-1}{m-1}(a_m - a_1), a_1 + \frac{i}{m-1}(a_m - a_1)\right)$$

per $i \in \{1, \ldots, m-1\}$ e $F_X(t) = 1$ per $t \geq a_m$





Distribuzione binomiale

Si usa se X è il numero di esiti positivi su n ripetizioni indipendenti di un esperimento con probabilità p di esito positivo e $1 - p$ di esito negativo

$$X \sim \text{Bin}\,(n, p)$$

Esempi: qual è la probabilità che

- escano t teste lanciando n monete? *(ad es., 2 teste su 5 lanci)*
- esca al massimo t volte un numero ≥ 5 lanciando n volte un dado?
- almeno t pezzi su n siano difettosi?
- un numero di clienti compreso fra t_1 e t_2 paghi con carta di credito?

ipotizzando che i singoli lanci, pezzi, clienti siano indipendenti fra loro

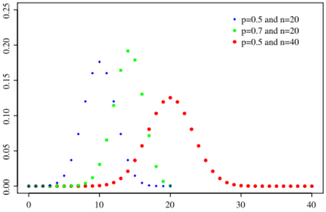
Pensate di calcolarle con i diagrammi di Venn e le regole della probabilità!

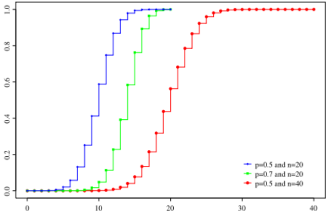
Distribuzione binomiale

$$p_X(t) = P(X = t) = \binom{n}{t} p^t (1-p)^{n-t} \quad \text{per ogni } t = 0, \ldots, n$$

da cui derivano gli indici

- $E[X] = np$
- $V[X] = np(1-p)$





Non occorre studiare le intersezioni e unioni fra eventi per fare il calcolo!

Distribuzione poissoniana

Si usa se X è il numero di occorrenze indipendenti tra loro di un evento durante un intervallo di tempo, sapendo che il valor medio è λ

$$X \sim \mathcal{P}(\lambda)$$

Esempi: qual è la probabilità che

- arrivino t telefonate a un call center durante la mattina?
- succedano al massimo t incidenti in una città al sabato sera?
- almeno t clienti si presentino alla cassa fra le 9 e le 10?

C'è una somiglianza con la distribuzione binomiale: infatti, equivale a un caso limite della distribuzione binomiale con $n \to +\infty$, $p \to 0$ e $\lambda = np$

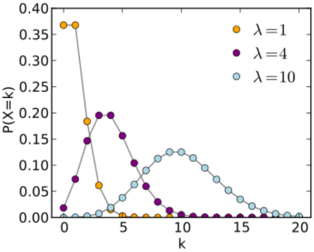
È come eseguire infiniti esperimenti di durata brevissima (in modo che un evento possa accadere al massimo una volta per ogni intervallo)

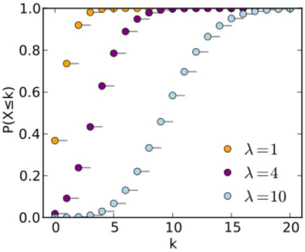
Distribuzione poissoniana

$$p_X(t) = P(X = t) = e^{-\lambda}\frac{\lambda^t}{t!}$$

da cui derivano gli indici

- $E[X] = \lambda$
- $V[X] = \lambda$





Distribuzione uniforme (continua)

Si usa quando X cade in modo equiprobabile in un intervallo $[a, b]$

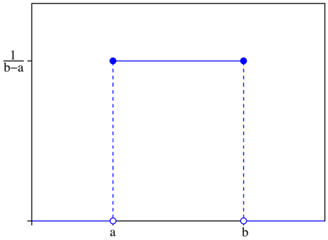
Ha funzione di densità

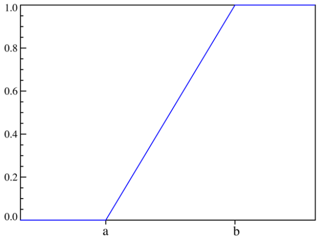
$$p_X(t) = \frac{1}{b-a} \text{ per } t \in [a, b], 0 \text{ altrove}$$

e quindi funzione di ripartizione $F_X(t) = 0$ per $t < a$,

$$F_X(t) = \frac{t-a}{b-a} \text{ per } t \in [a, b]$$

e $F_X(t) = 1$ per $t \geq b$





Distribuzione esponenziale

$$X \sim \mathcal{E}(\lambda)$$

Si usa se X è la durata di un fenomeno "privo di memoria", cioè per il quale la probabilità condizionata coincide con quella a priori

$$P(X > t + \delta | X > t) = P(X > \delta) \qquad \text{per ogni } t, \delta \geq 0$$

La probabilità che l'autobus abbia un ulteriore ritardo δ è fissa, qualunque sia il ritardo già accumulato

Esempi: qual è la probabilità che

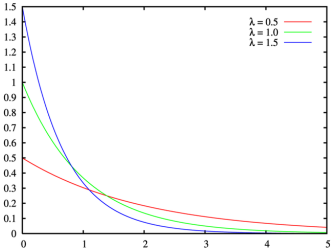
- la prossima telefonata arrivi fra t minuti?
- il prossimo cliente arrivi alla cassa fra al massimo t minuti?
- il tempo necessario a sbrigare il cliente sia di almeno t secondi?
- il prossimo guasto si verifichi fra t_1 e t_2 giorni?

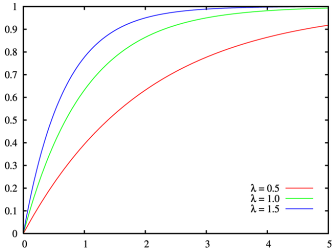
Per gli autobus probabilmente non è valida

$$f_X(t) = \lambda e^{-\lambda t} \qquad F_X(t) = 1 - e^{-\lambda t}$$

da cui derivano gli indici

- $E[X] = \dfrac{1}{\lambda}$
- $V[X] = \dfrac{1}{\lambda^2}$





Distribuzione normale

Si usa se X è una misura che tende a concentrarsi intorno a un valor medio significativo con disturbi secondari simmetrici

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

Esempi: qual è la probabilità che

- il peso di una confezione di pasta sia t grammi?
- il tempo di carico di un camion sia al massimo pari a t minuti?
- il giudizio numerico dei clienti sulla qualità del servizio sia almeno t?
- la durata dei pneumatici dei camion sia compresa fra t_1 e t_2 giorni?

Ipotizziamo valori continui

Distribuzione normale

$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

da cui derivano gli indici

- $E[X] = \mu$
- $V[X] = \sigma^2$