Teoria dei Segnali – Richiami di teoria della probabilità

Valentino Liberali

Dipartimento di Fisica
Università degli Studi di Milano
valentino.liberali@unimi.it

UNIVERSITÀ DEGLI STUDI DI MILANO

Teoria dei Segnali – Richiami di teoria della probabilità – 13 dicembre 2010

Contenuto

Probabilità (1/4)

- Da un punto di vista puramente matematico, il risultato di un'operazione può essere calcolato con una precisione infinita.
- Invece, quando abbiamo a che fare con grandezze fisiche, nessuna operazione fornisce un risultato "esatto".

Esempio: la misura di una stessa grandezza fisica può dare risultati diversi, a causa delle inaccuratezze del sistema di misura e delle fluttuazioni dei valori di alcuni parametri.

Probabilità (2/4)

La teoria della probabilità serve per trattare i fenomeni "di massa", cioè quelli che coinvolgono un numero molto grande di persone, particelle, eventi.
Benché spesso non sia possibile fare affermazioni **certe** su una singola persona, una singola particella, o un singolo evento, esistono tecniche matematiche che permettono di calcolare i **valori medi** (e altri indici statistici) **con un elevato grado di attendibilità**.

Probabilità (3/4)

- A livello macroscopico, la teoria della probabilità è necessaria ogni volta che non si può avere una conoscenza completa di un fenomeno per motivi pratici.

Esempio: voglio sapere quante persone oggi in Lombardia hanno usato l'automobile, quante hanno usato il treno e quante la bicicletta.
Siccome non posso contare tutti i pendolari ad uno ad uno, devo limitare l'indagine ad un sottoinsieme rappresentativo della popolazione (il "*campione statistico*").
Aumentando la numerosità del campione statistico, il risultato dell'indagine *tende* al valore "esatto" in percentuale.

Probabilità (4/4)

- A livello microscopico, esistono limiti fondamentali che impediscono la conoscenza completa di alcuni fenomeni.

Esempio: in base al principio di indeterminazione di Heisenberg, è impossibile conoscere contemporaneamente e con esattattezza sia la posizione sia la quantità di moto di una particella.
Di conseguenza, non posso prevedere la traiettoria futura del singolo elettrone (ma posso prevedere la velocità media di un grande numero di elettroni).

Definizioni di probabilità

Esistono più definizioni del concetto matematico di **probabilità**:

- Definizione assiomatica (Kolmogorov), basata sulla teoria della misura
- Definizione basata sul concetto di frequenza relativa (von Mises)
- Definizione classica "a priori", come il rapporto tra il numero dei casi favorevoli e il numero dei casi equiprobabili

Definizione assiomatica (1/2)

Un **esperimento casuale** $\mathcal{E}$ può dare diversi risultati.
$\mathcal{S}$ è l'insieme di tutti i risultati possibili ("spazio degli eventi").
Un **evento** $\mathcal{A}$ è un sottoinsieme (proprio o improprio) di $\mathcal{S}$ ($\mathcal{S}$ è l'evento certo).
$\mathcal{A} + \mathcal{B}$ è l'evento che si ha quando si presenta l'evento $\mathcal{A}$ o l'evento $\mathcal{B}$ o entrambi (unione dei due sottoinsiemi).

La **probabilità dell'evento** $\mathcal{A}$ è un **numero** $\Pr(\mathcal{A})$ che soddisfa questi postulati:

- la probabilità è positiva o nulla: $\Pr(\mathcal{A}) \geq 0$
- la probabilità dell'evento certo è uguale a uno: $\Pr(\mathcal{S}) = 1$
- se due eventi $\mathcal{A}$ e $\mathcal{B}$ si escludono a vicenda, allora $\Pr(\mathcal{A} + \mathcal{B}) = \Pr(\mathcal{A}) + \Pr(\mathcal{B})$

Definizione come frequenza relativa

Un esperimento casuale $\mathcal{E}$ è ripetuto N volte.
Se l'evento $\mathcal{A}$ si presenta $N_{\mathcal{A}}$ volte, allora la sua probabilità $\Pr(\mathcal{A})$ è il limite della frequenza relativa $N_{\mathcal{A}}/N$, facendo tendere ad infinito il numero di prove N:

$$\Pr(\mathcal{A}) = \lim_{N\to\infty} \frac{N_{\mathcal{A}}}{N}$$

Nota: questa è comunque una definizione assiomatica; qui l'assioma consiste nell'esistenza del limite per $N \to \infty$

Definizione classica

Si considerano "a priori" tutti i risultati possibili di un esperimento casuale $\mathcal{E}$.
Se N è il numero totale dei risultati possibili, e l'evento $\mathcal{A}$ si presenta in un numero $N_{\mathcal{A}}$ di risultati, allora la probabilità $\Pr(\mathcal{A})$ è:

$$\Pr(\mathcal{A}) = \frac{N_{\mathcal{A}}}{N}$$

a condizione che tutti i risultati siano ugualmente verosimili.

Nota: questa definizione è una **tautologia**, perchè definisce la probabilità attraverso un'altra probabilità ("ugalmente verosimili" significa "equiprobabili")!

Esempio 1

Un episodio di vita vissuta (27 settembre 2006):

Dobbiamo recarci al seguente indirizzo:
41 Posidonos, Glyfada, Attica (Grecia).

La piantina di Glyfada, ottenuta da MapQuest (`www.mapquest.com`), riporta la via Posidonos, e ci rechiamo nella via indicata sulla piantina.

Qual è la probabilità di essere arrivati nel posto giusto?

Esempio 1 (continuazione)

41 Posidonos, Glyfada, Attica (Grecia)
Qual è la probabilità p di essere arrivati nel posto giusto usando la piantina di MapQuest?

Per risolvere questo problema, occorre **sbarazzarsi del pregiudizio che in una qualunque città ci sia una sola via con un dato nome**. Se esistono N vie con lo stesso nome, e i nomi delle vie vengono stampati sulle piantine di MapQuest con uguale probabilità, allora

$$p = \frac{1}{N}$$

A Glyfada esistono DUE vie Posidonos (come ho potuto verificare *sperimentalmente*), e quindi la probabilità è $p = \frac{1}{2}$.

Esempio 2

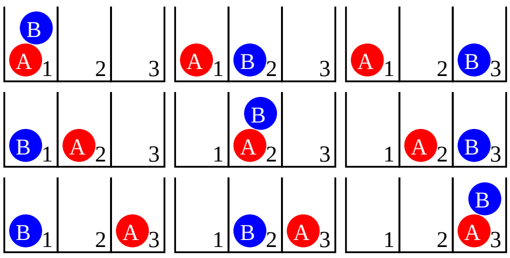
Disponiamo "a caso" n palline in m scatole ($n < m$). Qual è la probabilità p di avere le palline in n scatole prescelte, una per ogni scatola?

Il problema, formulato in questo modo, non ha una soluzione univoca!
Infatti, per trovare la probabilità p, siamo costretti a fare delle ipotesi aggiuntive sulle regole per disporre "a caso" le n palline nelle m scatole. Si ottengono questi risultati:

- statistica di **Maxwell-Boltzmann**: $p = \frac{n!}{m^n}$
- statistica di **Bose-Einstein**: $p = \frac{n!(m-1)!}{(n+m-1)!}$
- statistica di **Fermi-Dirac**: $p = \frac{n!(m-n)!}{m!}$

Esempio 2: prima soluzione

Le palline sono distinguibili l'una dall'altra e ogni scatola può contenere un numero qualsiasi di palline.



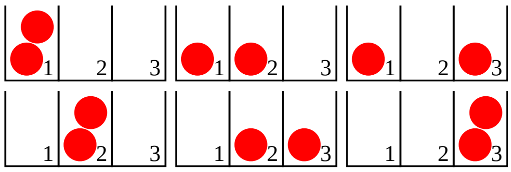
Statistica di Maxwell-Boltzmann

Il numero delle possibili disposizioni è $N = m^n$.
Il numero delle disposizioni favorevoli è uguale al numero delle permutazioni delle n palline: $N_A = n!$.
La probabilità p è:

$$p = \frac{n!}{m^n}$$

Esempio 2: seconda soluzione

Le palline sono indistinguibili l'una dall'altra e ogni scatola può contenere un numero qualsiasi di palline.



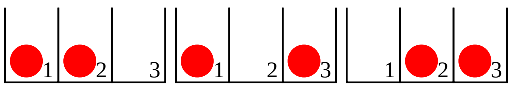
Statistica di Bose-Einstein

Il numero delle possibili disposizioni è $N = \frac{(n+m-1)!}{n!(m-1)!}$.
Il numero delle disposizioni favorevoli è $N_A = 1$.
La probabilità p è:

$$p = \frac{n!(m-1)!}{(n+m-1)!}$$

Esempio 2: terza soluzione

Le palline sono indistinguibili l'una dall'altra e ogni scatola può contenere una sola pallina.



Statistica di Fermi-Dirac

Il numero delle possibili disposizioni è $N = \frac{m!}{n!(m-n)!}$.
Il numero delle disposizioni favorevoli è $N_A = 1$.
La probabilità p è:

$$p = \frac{n!(m-n)!}{m!}$$

Considerazioni sugli esempi

Dai due esempi visti, si vede che il ragionamento "a priori" può trarre in errore!
Le discipline scientifiche devono basarsi su *ipotesi* da cui derivino risultati *in accordo con i dati sperimentali*.

Nel caso dell'esempio 2 (n palline in m scatole):

- le molecole di un gas all'interno di un recipiente seguono la statistica di **Maxwell-Boltzmann**;
- le particelle elementari con numero quantico di spin intero (protoni, neutroni, fotoni) seguono la statistica di **Bose-Einstein**;
- le particelle elementari con numero quantico di spin semi-intero (elettroni) seguono la statistica di **Fermi-Dirac**.

Definizione assiomatica (2/2)

Secondo la definizione assiomatica, la **probabilità dell'evento $\mathcal{A}$** è un **numero** $\Pr(\mathcal{A})$ che soddisfa questi assiomi:

1. la probabilità è positiva o nulla: $\Pr(\mathcal{A}) \geq 0$
2. la probabilità dell'evento certo $\mathcal{S}$ è uguale a uno: $\Pr(\mathcal{S}) = 1$
3. se due eventi $\mathcal{A}$ e $\mathcal{B}$ si escludono a vicenda, allora $\Pr(\mathcal{A} + \mathcal{B}) = \Pr(\mathcal{A}) + \Pr(\mathcal{B})$

Occorre aggiungere un altro assioma per l'unione di un numero infinito di eventi:

4. se per ogni coppia $\mathcal{A}_i$ e $\mathcal{A}_j$, con $i, j = 1, \ldots, n, \ldots$ e $i \neq j$, gli eventi $\mathcal{A}_i$ e $\mathcal{A}_j$ si escludono a vicenda, allora $\Pr(\mathcal{A}_1 + \ldots + \mathcal{A}_n + \ldots) = \Pr(\mathcal{A}_1) + \ldots + \Pr(\mathcal{A}_n) + \ldots$

Probabilità condizionata

Dati due eventi $\mathcal{A}$ e $\mathcal{B}$, l'evento $\mathcal{AB}$ è l'evento che si presenta quando si presentano contemporaneamente i due eventi $\mathcal{A}$ e $\mathcal{B}$.
Se due eventi $\mathcal{A}$ e $\mathcal{B}$ si escludono a vicenda, allora $\Pr(\mathcal{AB}) = 0$.

Se $\mathcal{M}$ è un evento con probabilità $\Pr(\mathcal{M}) \neq 0$, la **probabilità dell'evento $\mathcal{A}$ condizionata dall'evento $\mathcal{M}$** è:

$$\Pr(\mathcal{A}|\mathcal{M}) = \frac{\Pr(\mathcal{AM})}{\Pr(\mathcal{M})}$$

Esempio 3

Una scatola contiene tre carte, delle quali la prima ha entrambe le facce rosse, la seconda ha una faccia rossa e l'altra bianca, e la terza ha entrambe le facce bianche.
A caso, una delle tre carte viene estratta dalla scatola e posata sul tavolo.
La faccia visibile della carta estratta di colore rosso.
Qual è la probabilità che anche la faccia non visibile sia di colore rosso?

Esempio 3: soluzione

Usiamo la formula della probabilità condizionata, tenendo presente che le carte sono 3, mentre le facce sono 6: 3 rosse e 3 bianche.
$\mathcal{A}$ è l'evento "faccia inferiore rossa", mentre $\mathcal{M}$ è l'evento "faccia superiore rossa"; $\mathcal{AM}$ è l'evento "entrambe le facce rosse". La probabilità di estrarre la faccia superiore rossa è $\Pr(\mathcal{M}) = \frac{3}{6} = \frac{1}{2}$, mentre la probabilità di estrarre la carta con entrambe le facce rosse è $\Pr(\mathcal{AM}) = \frac{1}{3}$.
Quindi, se la prima faccia è rossa, la probabilità che lo sia anche la seconda è:

$$\Pr(\mathcal{A}|\mathcal{M}) = \frac{\Pr(\mathcal{AM})}{\Pr(\mathcal{M})} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} \qquad \text{e non } \frac{1}{2} \text{ !!!}$$

Eventi indipendenti

Due eventi $\mathcal{A}$ e $\mathcal{B}$ sono indipendenti se

$$\Pr(\mathcal{AB}) = \Pr(\mathcal{A})\Pr(\mathcal{B})$$

Se gli eventi $\mathcal{A}$ e $\mathcal{B}$ sono indipendenti, allora:

$$\Pr(\mathcal{A}|\mathcal{B}) = \Pr(\mathcal{A})$$

$$\Pr(\mathcal{B}|\mathcal{A}) = \Pr(\mathcal{B})$$

Variabili aleatorie (1/2)

Per un esperimento casuale $\mathcal{E}$, associamo ad ogni possibile risultato ζ un numero $X(\zeta)$.
X è una funzione che ha come dominio l'insieme dei possibili risultati dell'esperimento casuale $\mathcal{E}$, e come codominio un certo insieme di numeri.
Questa funzione X è detta **variabile aleatoria**.
Una variabile aleatoria può essere **discreta** oppure **continua**.

Fissato un certo numero x_1, la notazione $\{X \leq x_1\}$ rappresenta l'insieme di tutti i risultati ζ per cui $X(\zeta) \leq x_1$.
Analogamente, $\{X = x_1\}$ rappresenta l'insieme dei risultati ζ per cui $X(\zeta) = x_1$, e così via.
Esempio: se lancio un dado a 6 facce, $\{X < 4\}$ è l'insieme dei risultati {⚀, ⚁, ⚂}.

Variabili aleatorie (2/2)

Una variabile aleatoria deve soddisfare queste proprietà:

1. L'insieme $\{X \leq x\}$ è un evento per ogni $x \in \mathbb{R}$, cioè è definita la probabilità $\Pr\{X \leq x\}$
2. La probabilità degli eventi $\{X = +\infty\}$ e $\{X = -\infty\}$ è nulla:

$$\Pr\{X = +\infty\} = \Pr\{X = -\infty\} = 0$$

Funzione cumulativa di distribuzione

La **funzione cumulativa di distribuzione** $F_X(x)$ è la probabilità che la variabile aleatoria $X(\zeta)$ abbia un valore minore o uguale a x:

$$F_X(x) = \Pr\{X \leq x\}$$

Proprietà:

- $0 \leq F_X(x) \leq 1$
- $F_X(-\infty) = 0$
- $F_X(+\infty) = 1$
- monotonicità: $F_X(x_1) \leq F_X(x_2) \quad \text{se } x_1 < x_2$

Funzione densità di probabilità (1/6)

La **funzione densità di probabilità** $f_X(x)$ è la derivata della funzione cumulativa di distribuzione; di solito si indica con l'abbreviazione **pdf** (= probability density function).

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Proprietà:

- $f_X(x) \geq 0 \quad \text{per } \forall x$. Poiché $F_X(x)$ è monotona non decrescente, la sua derivata è non negativa.
- $\int_{-\infty}^{+\infty} f_X(x) dx = 1$. Infatti: $\int_{-\infty}^{+\infty} f_X \, dx = F_X(+\infty) - F_X(-\infty) = 1 - 0 = 1$.
- $F_X(x) = \int_{-\infty}^{x} f_X(x) dx$

Funzione densità di probabilità (2/6)

La variabile casuale X può essere continua o discreta.

- Se X è continua, allora (di solito) $f_X(x)$ è una funzione **liscia**, cioè $f_X(x)$ è derivabile infinite volte, e di conseguenza anche $F_X(x)$ è una funzione liscia.
- Se X è discreta, allora $F_X(x)$ presenta dei gradini, e $f_X(x)$ è una somma di funzioni delta di Dirac.

Funzione densità di probabilità (3/6)

Esempio: L'esperimento casuale consiste nel lancio di un dado a sei facce. Lo spazio degli eventi è:

$$\mathcal{S} = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$$

La variabile aleatoria X è il punteggio indicato dal dado: $X \in \{1, 2, 3, 4, 5, 6\}$. Se il dado non è truccato, ciascuno dei risultati possibili si presenta con probabilità $\frac{1}{6}$.
Poiché nessun risultato è minore di 1, si ha:

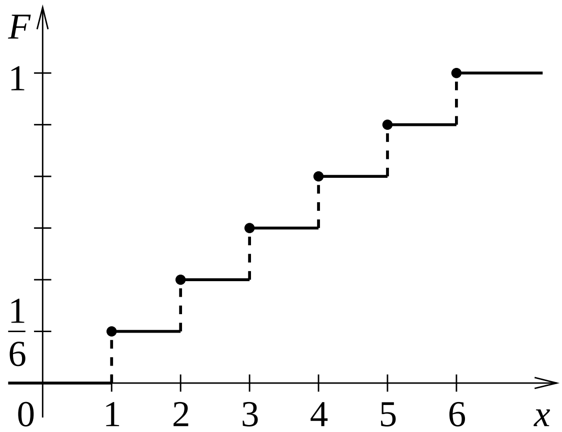
$$F_X(x) = 0 \quad \text{per } \forall x < 1$$

e poiché nessun risultato è maggiore di 6, si ha:

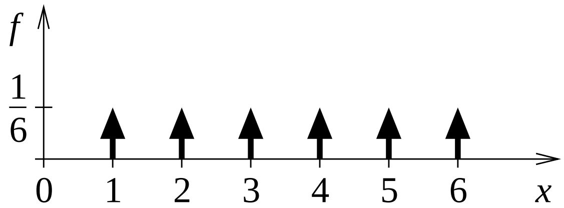
$$F_X(x) = 1 \quad \text{per } \forall x \geq 6$$

Funzione densità di probabilità (4/6)

$F_X(x) = \frac{1}{6}\sum_{i=1}^{6} u(x-i)$



$f_X(x) = \frac{1}{6}\sum_{i=1}^{6} \delta(x-i)$



Funzione densità di probabilità (5/6)

La probabilità che la variabile casuale $X(\zeta)$ assuma valori compresi nell'intervallo $[x_1, x_2]$ è:

$$\begin{aligned}
\Pr\{x_1 \le X(\zeta) \le x_2\} &= \Pr\{X(\zeta) \le x_2\} - \Pr\{X(\zeta) < x_1\} = \\
&= F_X(x_2) - F_X(x_1 - \varepsilon) = \\
&= \int_{-\infty}^{x_2} f_X(x)dx - \int_{-\infty}^{x_1-\varepsilon} f_X(x)dx = \\
&= \int_{x_1-\varepsilon}^{x_2} f_X(x)dx = \\
&= \int_{x_1}^{x_2} f_X(x)dx
\end{aligned}$$

(l'ultimo passaggio è lecito solo se $f_X(x)$ è liscia in x_1)

Funzione densità di probabilità (6/6)

Se $f_X(x)$ è una funzione liscia, allora $f_X(x)dx$ rappresenta la probabilità che la variabile aleatoria $X(\zeta)$ abbia un valore compreso tra x e $x + dx$:

$$f_X(x)dx = \Pr\{x \leq X(\zeta) \leq x + dx\}$$