

# Modelli probabilistici di fenomeni aleatori

Da Wikiversità, l'università aperta.

La **teoria della probabilità** si occupa dello studio dei fenomeni aleatori. Quando non siamo in grado di dare una caratterizzazione esatta del fenomeno e dobbiamo dare una descrizione globale del fenomeno stesso, usiamo la probabilità.

## Indice

- 1 Probabilità e statistica
  - 1.1 Esperimenti aleatori
  - 1.2 Teoria della probabilità e statistica
  - 1.3 Teoria dell'informazione
- 2 Statistica
  - 2.1 Insieme delle parti
  - 2.2 Classe di insiemi
  - 2.3 Definizioni di probabilità
    - 2.3.1 Probabilità secondo la frequenza relativa
    - 2.3.2 Probabilità secondo il modello probabilistico
- 3 Note

## Probabilità e statistica

### Esperimenti aleatori

Sono aleatori tutti gli esperimenti per i quali è difficile o impossibile prevedere in modo esatto il risultato, ma presentano una qualche forma di regolarità. Il comportamento dei fenomeni aleatori può essere descritto solo attraverso grandezze globali e/o medie.

*Non ci interessa solo il caso in cui sia impossibile, ma anche sia molto difficile, così tanto da rendere la descrizione irrealizzabile.*

Pensando alla definizione di probabilità, i valori medi possono essere i momenti e le regolarità del primo o second'ordine. Tanto per dare un esempio, è difficile predire esattamente il risultato di ogni lancio del dado, ma se il dado non è truccato posso dire che ogni faccia ha la stessa probabilità di uscire. Posso prevedere il valor medio del risultato e la statistica collegata.

### Teoria della probabilità e statistica

La teoria della probabilità si occupa della costruzione di modelli probabilistici (matematici) che descrivano i fenomeni aleatori. La statistica, invece, si occupa di verificare l'aderenza di un modello rispetto ai dati sperimentali.

Nella parte di teoria della probabilità possiamo dire qual è la funzione di densità di probabilità del dado. La statistica si preoccupa di dire se, dato un dado, questo aderisce al modello della teoria della probabilità o se questo è truccato. Gli ambiti in cui viene utilizzata la teoria della probabilità sono molti, per esempio:

- teoria delle code;
- instradamento ottimo dei pacchetti;
- analisi fatta a livello statistico;
- meccanica statistica relativa ai gas (posso descrivere la pressione, che è un valore medio, e non la posizione di ogni molecola);
- elaborazione e trasmissione dell'informazione.

## Teoria dell'informazione

La teoria dell'informazione studia i problemi legati all'elaborazione e alla trasmissione dell'informazione utilizzando un approccio probabilistico.

### Esempio:

Se dico:

1. oggi il treno per Milano delle 17.25 sarà in ritardo di 10 minuti
2. oggi il treno per Milano delle 17.25 sarà puntuale

*Pensando a come funzionano le ferrovie in Italia, qual è l'informazione più importante? La seconda, perché si verifica raramente: un evento raro è molto più informativo. La misura di informazione è basata sulla probabilità. L'informazione associata ad un evento è inversamente proporzionale alla sua probabilità di occorrenza.*

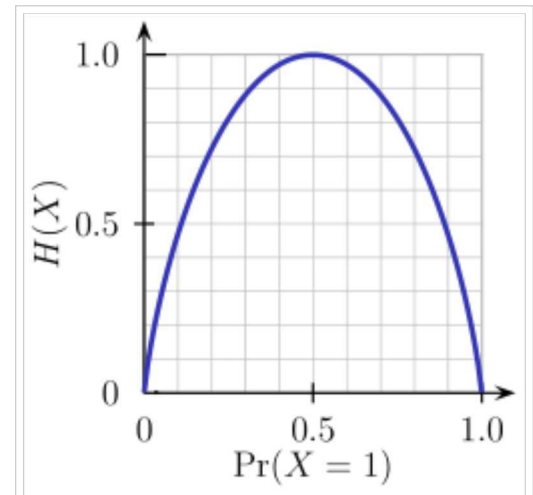


Fig.1 - Entropia di una sorgente binaria

L'informazione si misura con la definizione<sup>[1]</sup>

$$i(m_k) = \log \frac{1}{p(m_k)}$$

Dall'informazione  $i(m)$  si passa alla definizione di entropia:

$$H(M) = \sum_{k=1}^m p(m_k) \cdot i(m_k) = \sum_{k=1}^m p(m_k) \cdot \log \frac{1}{p(m_k)} = E[i] \text{ con } m_k \in M$$

L'entropia non è altro che l'informazione media di una sorgente.

### Esempio:

*Supponiamo di avere una sorgente  $X$  che emette simboli 0 e 1. La sorgente emette simboli in modo equiprobabile, quindi*

- $p(0) = 1/2$
- $p(1) = 1/2$

*Di solito si hanno delle stringhe di bit 1101110011. Un'operazione importante è la codifica della sorgente; quello che si vuole fare è trovare un codice per rappresentare questa stringa, con o senza perdita, con un numero minore di bit. Con la probabilità data non è possibile comprimere la stringa, perché i simboli sono equiprobabili. Se i simboli sono indipendenti, il fatto che sia uscito un 1 o uno 0 non influenza il risultato del prossimo simbolo.*

*In questo caso, posso rappresentare con meno bit i simboli della sorgente, posso comprimere.*

*Sappiate che la costruzione di modelli semplificati può cambiare nettamente le prestazioni di un canale o di un sistema di telecomunicazioni.*

## Statistica

### Definizione: Fenomeno aleatorio

Un fenomeno aleatorio è un esperimento i cui possibili risultati appartengono ad un insieme ben definito e dove l'esito non è prevedibile (o predicibile) a priori. È importante che l'insieme dei possibili risultati sia ben definito, deve essere noto.

### Definizione: Spazio degli esiti

Lo spazio degli esiti, o spazio campione  $\Omega$  associato ad un esperimento aleatorio, è l'insieme di tutti i possibili risultati di un esperimento. Può essere finito o infinito, sia numerabile che non numerabile.

### Definizione: Evento

Dato uno spazio campione  $\Omega$ , si dice evento un qualsiasi sottoinsieme  $A$  di  $\Omega$ ,  $A \subseteq \Omega$ .

### Esempio:

Si consideri il lancio di un dado a 6 facce. Si ha:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

### Definizione: Spazio degli eventi

Dato uno spazio campione  $\Omega$ , si definisce spazio degli eventi  $F$  l'insieme non vuoto che contiene tutti gli elementi di interesse (determinabili su  $\Omega$ ) che soddisfano le seguenti proprietà:

1.  $\Omega, \emptyset \in F$
2.  $A, B \in F \Rightarrow A \cap B \in F$  (dai teoremi di De Morgan)
3.  $\forall A \subseteq F \Rightarrow \overline{A} \subseteq F$

Un spazio  $F$  è una  $\sigma$ -algebra se vale anche:

$$\blacksquare \forall A_1, \dots, A_k \text{ dove } k \in [1, \infty] \Rightarrow \bigcup_{i=1}^k A_i \in F$$

cioè, se si ha chiusura rispetto all'unione numerabile. Noi useremo esclusivamente  $\sigma$ -algebre.

Ci saranno, in generale, più di uno spazio  $F$  degli eventi. Il più banale deve contenere l'unione ed il complemento degli eventi.

$$F_1 = \{\emptyset, \Omega\}$$

### Esempio: Esempio di utilizzo della teoria della probabilità: il cut detection

In un filmato, si assume che frame vicini siano simili tra loro. Qual è l'interframe, la distanza tra due frame? Quando l'interframe è troppo elevato, posso dichiarare che c'è stata una transizione del filmato. Posso usare soglie fisse o non fisse (le soglie adattative). Grazie al modello probabilistico, si può introdurre la soglia adattativa, cioè che si adatta in base al modello di probabilità che stimo sui dati.

## Insieme delle parti

*Definizione: Insieme delle parti*

Si dice insieme delle parti lo spazio degli eventi  $F$  che contiene tutti i possibili eventi di  $\Omega$ , cioè tutti i possibili sottoinsiemi che posso costruire con gli elementi di  $\Omega$ .

$$F_2 = P(\Omega) = 2^\Omega$$

*Esempio: I dadi*

L'insieme delle parti è

$$F_2 = \{\emptyset, \{1\}, \{2\}, \{1, 3, 5\}, \{2, 4, 6\}, \{1, 2, 3\}, \dots, \Omega\}$$

## Classe di insiemi

*Definizione: Classi di insiemi*

Dato un insieme  $X$  si dice classe  $C$  una collezione di sottoinsiemi di  $X$ . La classe di tutti i possibili sottoinsiemi di  $X$  si chiama insieme (o collezione, o classe) delle parti.

*Definizione: Partizione di un insieme*

Una partizione è la classe di sottoinsiemi  $\{X_1, X_2, \dots, X_n\}$  tali che

- $X_i \cdot X_j = \{\emptyset\} \quad \forall i \neq j$
- $\sum_{i=1}^n X_i = X$

*Esempio: Il dado*

Con  $\Omega = \{1, 2, 3, 4, 5, 6\}$ , una partizione può essere

$$P = \{(1,3), (2,4), (5,6)\}$$

*Definizione: Cardinalità*

La cardinalità di un insieme è il numero di elementi che esso contiene. Se la cardinalità di  $\Omega$  è  $N$ , allora la cardinalità dell'insieme delle parti  $F$  è

$$|F| = 2^N$$

## Definizioni di probabilità

### Probabilità secondo la frequenza relativa

Una delle possibili definizioni di probabilità è quella che usa la frequenza relativa. Si dice che la probabilità  $P(A)$  di un evento  $A$  è data da

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

dove  $n$  è il numero di volte che si ripete l'esperimento, mentre  $n_A$  è il numero di volte che si verifica l'evento  $A$ .

### Probabilità secondo il modello probabilistico

Un modello probabilistico di un fenomeno aleatorio è lo spazio di probabilità identificato da tre elementi  $(\Omega, F, P)$ , dove:

- $\Omega$  è lo spazio degli esiti;
- $F$  è lo spazio degli eventi;
- $P$  è la probabilità.

#### Definizione: Probabilità $P$

Assegnato uno spazio campione  $\Omega$  ed una  $\sigma$ -algebra  $F$  di eventi di  $\Omega$ , si definisce probabilità una funzione  $P$  definita su  $F$  a valori in  $\mathbb{R}$  (non negativi), tale che

1.  $P(A) \geq 0$
2.  $P(\Omega) = 1$
3. se  $\{A_n\}_{n=1}^{+\infty}$  è una successione di eventi mutuamente esclusivi, cioè

$$(A_i) \cap A_j = \emptyset \forall i \neq j, \text{ allora } P\left(\bigcup_i^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Quest'ultima proprietà è detta *additività numerabile*, perchè indica che gli elementi hanno intersezione nulla e la somma delle loro probabilità si può portare fuori dal segno di probabilità.

## Note

1. ↑ Durante tutta la trattazione si usa la base 2. Questo perchè qualsiasi insieme finito o infinito proprio può essere messo in relazione con l'insieme dei numeri naturali, e questi possono essere indicizzati con l'utilizzo dei soli simboli  $\{0,1\}$ . Inoltre, lo scopo del corso è permettere l'utilizzo di tecnologie di tipo digitale, che si basano proprio sulla base 2.

Estratto da "[http://it.wikiversity.org/wiki/Modelli\\_probabilistici\\_di\\_fenomeni\\_aleatori](http://it.wikiversity.org/wiki/Modelli_probabilistici_di_fenomeni_aleatori)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 23:17, 15 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Spazi di probabilità

Da Wikiversità, l'università aperta.

## Indice

- 1 Spazi di probabilità discreti
- 2 Spazi di probabilità continui
  - 2.1 Spazio di Borel
  - 2.2 Algebra di Borel
  - 2.3 Metodi per l'introduzione di misure su spazi misurabili
  - 2.4 Probabilità su spazi misurabili

### Definizione: Spazio numerabile

Una coppia  $(\Omega, F)$  dove  $\Omega$  è un insieme e  $F$  una  $\sigma$ -algebra di sottoinsiemi di  $\Omega$  è detta spazio numerabile.

### Definizione: Misura

Una funzione  $m$  definita sulla  $\sigma$ -algebra  $F$ , con

$$m = F \rightarrow [0, \infty]$$

è una misura se

$$\forall A_k \in F, k \in [0, \infty], A_i \cap A_j = \emptyset \quad \forall i \neq j$$

e vale

$$m\left(\bigcup_k A_k\right) = \sum_k m(A_k)$$

dove gli  $A_k$  sono una collezione di insiemi.

### Definizione: Probabilità

Dato uno spazio misurabile  $(\Omega, F)$ , la probabilità  $P$  è una misura sulla  $\sigma$ -algebra  $F$  tale che  $P(\Omega) = 1$ .

### Definizione: Spazio di probabilità

Lo spazio dotato di misura  $(\Omega, F, P)$  è detto spazio di probabilità (o esperimento casuale).

### Proprietà: Proprietà 1

$$B \subseteq A \Rightarrow P(B) \leq P(A)$$

#### Dimostrazione:

$$A = B \cup (A \cap \bar{B}) \Rightarrow P(A) = P(B \cup A \cap \bar{B}) = P(B) + P(A \cap \bar{B})$$

Si sottolinea che le probabilità sono sempre positive.

*Proprietà: Proprietà 2*

$$P(\bar{A}) = 1 - P(A)$$

*Dimostrazione:*

$$A \cup \bar{A} = \Omega, A \cap \bar{A} = \emptyset$$

*da qui si ha*

$$P(A \cup \bar{A}) = P(A) + P(\bar{A}) = P(\Omega) = 1$$

*Proprietà: Proprietà 3*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

*Dimostrazione:*

$$A \cup B = A \cup (\bar{A} \cap B) \rightarrow P(A \cup B) = P(A) + P(\bar{A} \cap B)$$

$$B = (A \cap B) \cup (\bar{A} \cap B) \rightarrow P(B) = P(A \cap B) + P(\bar{A} \cap B)$$

*Da qui si ottiene*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

*Proprietà: Proprietà 4*

$$P(A \cap B) \leq \min(P(A), P(B))$$

*Dimostrazione:*

$$B = (A \cap B) \cup (\bar{A} \cap B)$$

$$P(B) = P(A \cap B) + P(\bar{A} \cap B) \Rightarrow P(A \cap B) \leq P(B)$$

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) \Rightarrow P(A \cap B) \leq P(A)$$

*Proprietà: Proprietà 5**La funzione probabilità è continua.**Dimostrazione:**Sia  $\{A_n\}_{n=1}^{\infty}$  una successione di eventi di  $F$  con*

$$\lim_{n \rightarrow \infty} A_n = A$$

*Allora*

$$\lim_{n \rightarrow \infty} P(A_n) = P(\lim_{n \rightarrow \infty} A_n).$$

#### Definizione alternativa: *Probabilità col metodo statistico*

Assegnata una coppia  $(\Omega, F)$ , si dice *probabilità* la funzione  $P$  su  $F$ , a valori non negativi, tale che:

1.  $P(A) \geq 0 \forall A \in F$ , assioma della non negatività;
2.  $P(\Omega) = 1$ , assioma di normalizzazione
3. La *successione*

$$S \equiv \{A_n\}_{n=1}^{\infty}$$

è una *successione di eventi mutualmente esclusivi*, quindi

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$$

Questi sono gli assiomi di Kolmogorov.

#### Definizione alternativa: *Probabilità col metodo frequentista*

Un'altra definizione è quella che usa la *probabilità relativa*, il rapporto

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

Si può osservare che anche questa definizione soddisfa tutti e tre gli assiomi di Kolmogorov. Il problema è che, in genere, non è possibile ripetere un esperimento per un numero infinito di volte; inoltre, il limite potrebbe non esistere.

#### Definizione alternativa: *Probabilità col metodo classico*

Per far sì che esista sempre un risultato, si può definire la *probabilità* come

$$P(A) = \frac{|A|}{|\Omega|}$$

cioè, il rapporto tra il numero di possibili risultati favorevoli (la cardinalità di  $A$ ) e la cardinalità di  $\Omega$ . Anche in questo caso, sono rispettati gli assiomi di Kolmogorov.

## Spazi di probabilità discreti

Sia

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_k, \dots\}$$

un insieme discreto, con cardinalità finita o infinita numerabile. In tal caso, si può scegliere come  $\sigma$ -algebra la collezione delle parti

$$F = 2^\Omega$$

Un evento è definito come



$$A = \bigcup_{i \in I_A} \{\omega_i\}, \quad \forall A \in \mathcal{F}$$

dove gli  $\omega_i$  sono gli eventi elementari (e disgiunti:  $\omega_i \neq \omega_j \quad \forall i \neq j$ ). Allora, la probabilità viene definita come

$$P(A) = \sum_{i \in I_A} P(\omega_i)$$

Per caratterizzare in modo completo uno spazio di probabilità discreto è sufficiente, quindi calcolare soltanto i  $P(\omega_i)$ ,  $\forall i$ .

*Esempio:*

*Il lancio della moneta.*

$$\Omega = \{T, C\}$$

$$\mathcal{F} = \{\emptyset, \Omega, \{T\}, \{C\}\}$$

$$P(T) = 0,5$$

$$P(C) = 0,5$$

*Scrivendo questa definizione del problema, abbiamo usato l'approccio classico; questo approccio non funziona se si ha a che fare con monete truccate.*

Se la probabilità dei singoli eventi non è uniforme, allora il metodo classico introduce un errore.

*Esempio:*

*Il lancio di due monete.*

$$\Omega = \{TT, CC, TC, CT\}$$

$$P(TT) = P(CC) = P(CT) = P(TC) = 1/4$$

*La probabilità che esca T al primo lancio è 1/2.*

*Esempio:*

*Il lancio di un dado.*

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$F_1 = \{\emptyset, \text{PARI}, \text{DISPARI}, \Omega\}$$

$$F_2 = \{\emptyset, 1, 3, \{1, 3\}, \Omega\}$$

$$F_3 = \{\emptyset, 1, 2, \{1, 4\}, \Omega\}$$

*Tra queste  $F$ , quali sono  $\sigma$ -algebre? soltanto  $F_1$ .*

*Esempio:*

*Dati  $A$  e  $B$  ad una  $\sigma$ -algebra  $\mathcal{F}$ ,  $A, B \in \mathcal{F}$ , si definisce  $(\Omega, \mathcal{F}, P)$ . Quanto vale  $P(A|B)$ ?*

*Soluzione:*

*Si ha*

$$P(A \setminus B) = P(A) - P(A \cap B)$$

## Spazi di probabilità continui

Se  $\Omega$  è un insieme continuo, la sua dimensione  $|\Omega|$  è infinita non numerabile. In questo caso, l'insieme delle parti  $F = 2^\Omega$  è troppo ricco per poter definire una misura di probabilità sul suo contenuto.

Le successioni di insiemi  $\{A_n\}_{n=1}^\infty$  possono essere

- decrescenti (rispetto alla relazione di inclusione), cioè

$$A_n \supseteq A_{n+1} \quad \forall n \in \mathbb{N} \rightarrow \lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n.$$

- crescenti, cioè

$$A_n \subseteq A_{n+1} \quad \forall n \in \mathbb{N} \rightarrow \lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

## Spazio di Borel

Dato

$$\Omega = \mathbb{R} (-\infty, +\infty)$$

si consideri l'insieme  $G$  formato da tutti i sottoinsiemi di  $\mathbb{R}$  dati dalla somma finita di intervalli disgiunti della forma

$$(a, b] = \{x \in \mathbb{R} \text{ t.c. } a < x \leq b\} \text{ con } -\infty < a < b < +\infty$$

Allora

$$A = \bigcup_{i=1}^n (a_i, b_i] \quad n < \infty$$

*Teorema:*

*$G$  è un'algebra, ma non è una  $\sigma$ -algebra.*

*Dimostrazione: Dimostriamo che esistono algebre che non sono  $\sigma$ -algebre*

*Sia*

$$A_n = \left(0, 1 - \frac{1}{n}\right] \in G$$

*In questo caso, si ha*

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{i=1}^n A_n = (0, 1) \notin G$$

*Quindi, si vede che  $G$  non è una  $\sigma$ -algebra.*

*Esempio:*

$$A_n = \left(0, 1 - \frac{1}{N}\right] \in G$$

*Si ha*

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{i=1}^{\infty} A_n = (0, 1)$$

*ma  $(0, 1) \notin G$ , quindi questa non è una  $\sigma$ -algebra.*

## Algebra di Borel

L'algebra di Borel in  $\mathbb{R}$  è definita come la più piccola  $\sigma$ -algebra generata dalla collezione di tutti gli intervalli della forma  $(a, b]$  e viene indicata come  $\mathbb{B}(\mathbb{R})$ .

*Teorema:*

*L'algebra  $\mathbb{B}(\mathbb{R})$  contiene tutti gli intervalli del tipo*

- $[a, b]$
- $(a, b)$
- $(-\infty, b]$
- $\{a\}$
- $(-\infty, b)$

*Dimostrazione:*

*La dimostrazione è lasciata per esercizio. Si segua questa traccia.*

$$[a, b] = \bigcap_n \left(a - \frac{1}{n}, b\right], \quad a < b$$

$$\{a\} = \bigcap_n \left(a - \frac{1}{n}, a\right]$$

Lo spazio di Borel è definito come  $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ . La restrizione di  $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$  ad un insieme  $A \in \mathbb{B}(\mathbb{R})$  è la  $\sigma$ -algebra di tutti gli insiemi della forma  $B \cap A$  con  $B \in \mathbb{B}(\mathbb{R})$ .

*Esempio:*

$([0, 1], \mathbb{B}([0, 1]))$

*dove  $\mathbb{B}$  è uno spazio di Borel. Allora*

$$B \cap [0, 1] \text{ con } B \in \mathbb{B}(\mathbb{R})$$

## Metodi per l'introduzione di misure su spazi misurabili

In generale è più semplice definire una misura di sottoinsiemi di  $\Omega$  e poi estenderla alla  $\sigma$ -algebra  $F$ . Nel caso continuo, si procede definendo la misura di sottoinsiemi di  $\Omega$  appartenenti ad un'algebra tale che  $F$  possa essere generata.

Consideriamo  $(\Omega, F)$ . Sia  $a$  un'algebra di  $\Omega$  tale che  $F = \sigma(a)$ . Allora, si può definire una premisura

$$m_0: a \rightarrow [0, \infty]$$

in modo tale che gli insiemi siano

$$A_k \in a, \quad \forall k = 1, 2, \dots$$

sono disgiunti a coppie e

$$\left( \bigcup_k A_k \in a \right)$$

Allora, la misura dell'algebra è

$$m \left( \bigcup_k A_k \right) = \sum_k m_0(A_k)$$

### *Teorema: Teorema di unicità della misura*

*La misura*

$$m: F \rightarrow [0, \infty] \text{ tale che } m(A) = m_0(A) \quad \forall A \in a$$

*è unica, sotto l'ipotesi che  $m_0$  sia  $\sigma$ -finita, ossia che esista una sequenza di insiemi*

$$B_k \in a, \quad k = 1, 2, \dots$$

*tale che*

- $\bigcup_k B_k = \Omega$
- $m(B_k) < \infty \quad \forall k.$

### *Teorema: Teorema di Caratheodory*

*Si consideri una misura  $\sigma$ -finita su di un'algebra  $a$  di  $\Omega$ . Esiste ed è unica la misura  $m$  che estende  $m_0$ .*

## Probabilità su spazi misurabili

Consideriamo  $\Omega = \mathbb{R}$ ,  $F = \mathbb{B}(\mathbb{R})$ . Definiamo

$$P_0 \mid P(\Omega) = 1$$

sull'algebra  $G$  di tutti i sottoinsiemi di  $\mathbb{R}$  del tipo

$$\bigcup_{k=1}^n (a_k, b_k) \text{ con } (a_i, b_i) \cap (a_j, b_j) = \emptyset \quad \forall i \neq j$$

Nota: si ha  $\sigma(G) = \mathbb{B}(\mathbb{R})$ .

Per definire  $P_0$  definiamo la funzione di distribuzione  $F_X$ .

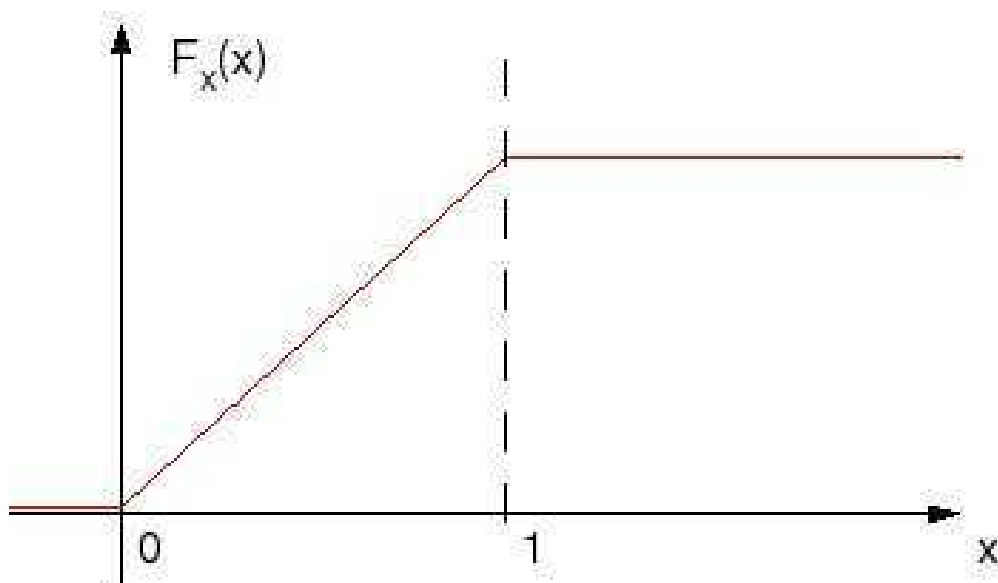
*Definizione: Funzione di distribuzione  $F_X$*

La funzione di distribuzione (o funzione di ripartizione) è definita come

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

tale che:

1.  $F_X$  non è decrescente
2.  $\lim_{x \rightarrow -\infty} F_X(x) = 0$
3.  $\lim_{x \rightarrow +\infty} F_X(x) = 1$
4.  $F_X$  è continua a destra, cioè  $\lim_{\epsilon \rightarrow 0^+} F_X(x + \epsilon) = F(x) \quad \forall x \in \mathbb{R}$
5.  $F_X$  ammette limite sinistro, cioè  $\exists \lim_{\epsilon \rightarrow 0^-} F_X(x + \epsilon) \quad \forall x \in \mathbb{R}$



Definiamo  $P_0: G \rightarrow [0, 1]$  come la probabilità

$$P_0(A) = \sum_{k=1}^m [F(b_k) - F(a_k)] \quad \forall A \in G \mid A = \bigcup_{k=1}^m (a_k, b_k)$$

dove gli  $(a_k, b_k)$  sono disgiunti a coppie. Si può dire che  $P_0$  è  $\sigma$ -finita,

$$P_0(\Omega) = 1$$

Dal teorema di Caratheodory segue che  $P_0$  si estende in modo univoco alla misura  $P$  su  $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$  tale che

$$P((a, b]) = F_X(b) - F_X(a)$$

Anche  $P$  è una probabilità, perché

$$P(\Omega) = F_X(\infty) - F_X(-\infty) = 1$$

La  $F_X(x)$  può essere assegnata come:

$$F_X(x) = \int_{-\infty}^x f_X(\alpha) d\alpha$$

con

$$F(x) \geq 0 \quad \forall x \in \mathbb{R}$$

e dove vale

$$\int_{-\infty}^{\infty} f_X(\alpha) d\alpha = 1$$

La  $f(\cdot)$  è detta funzione di densità di probabilità, o densità della funzione di distribuzione  $F$ .

Se esiste la  $f_X(x)$ , allora si ha

$$P((a, b]) = \int_a^b f(\alpha) d\alpha$$

*Esempio: La funzione di densità di probabilità gaussiana*

La funzione di densità di probabilità gaussiana è definita dall'equazione

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} = N(\mu, \sigma)$$

*Esempio: La funzione di densità di probabilità uniforme*

La funzione di densità di probabilità uniforme è definita dall'equazione

$$f(x) = \frac{1}{b-a} \quad \forall x \in [a, b], \quad a < b$$

*Definizione: Spazio di Borel su  $\mathbb{R}^2$*

Lo spazio di Borel

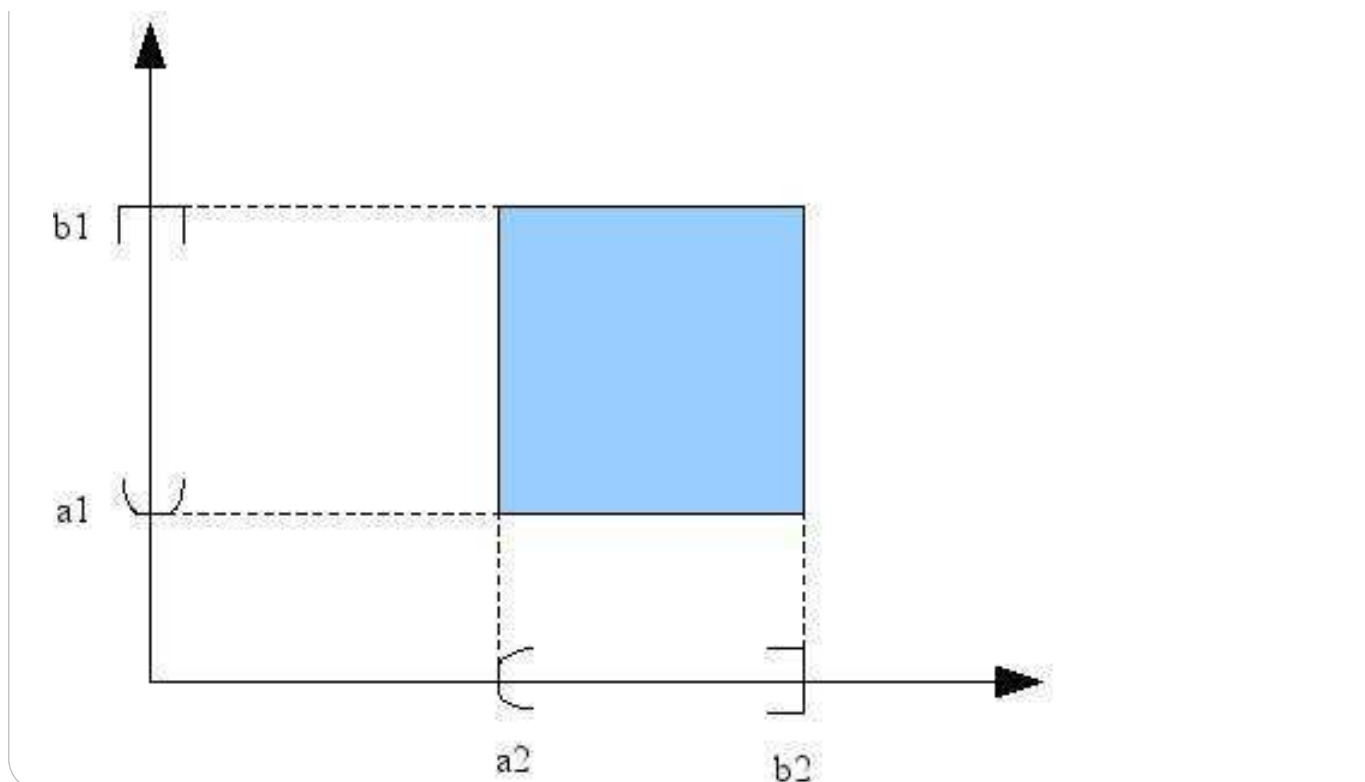
$$(\mathbb{R}^2, \mathbb{B}(\mathbb{R}^2))$$

con

$$\mathbb{R}^2 = \{(\alpha, \beta) | \alpha \in \mathbb{R}, \beta \in \mathbb{R}\}$$

è la più piccola  $\sigma$ -algebra su  $\mathbb{R}^2$  che contiene insiemi del tipo

$$(a_1, b_1] \times (a_2, b_2] = \{(\alpha, \beta) \in \mathbb{R}^2 | \alpha \in (a_1, b_1], \beta \in (a_2, b_2]\}$$



Analogamente al caso monodimensionale, si può definire una misura di probabilità assegnando un'opportuna densità di probabilità

$$f : \mathbb{R}^2 \rightarrow [0, \infty] \text{ è integrabile}$$

tale per cui

$$f(\alpha, \beta) \geq 0 \quad \forall (\alpha, \beta) \in \mathbb{R}^2$$

e con

$$\int_{\mathbb{R}^2} f(\alpha, \beta) d\alpha d\beta = 1$$

Il metodo visto si applica anche ai seguenti casi:

- $(A, \mathbb{B}(A))$  con  $A \in \mathbb{B}(\mathbb{R})$ . In tal caso

$$\int_A f(\alpha) d\alpha = 1$$

- $\mathbb{R}^n, \mathbb{B}(\mathbb{R}^n)$ , con

$$\int_{\mathbb{R}^n} f(\alpha_1, \alpha_2, \dots, \alpha_n) d\alpha_1 d\alpha_2 \dots d\alpha_n = 1$$

- $(A, \mathbb{B}(A))$  con  $A \in \mathbb{B}(\mathbb{R})^n$  e

$$\int_A f(\alpha_1, \alpha_2, \dots, \alpha_n) d\alpha_1 d\alpha_2 \dots d\alpha_n = 1$$

*Esempio: La roulette*

Vogliamo calcolare la probabilità che la pallina cada in un quadrante piuttosto che in un altro. Ipotizziamo che la probabilità che la pallina cada su un numero sia uguale per tutti i numeri; in questo caso  $\theta$  ha una densità di probabilità uniforme. Definiamo lo spazio di probabilità  $(\Omega, \mathcal{F}, P)$  come:

- $\Omega = [0, 2\pi]$
- $\mathcal{F} = \{A \in \Omega \mid A = B \cap \Omega, B \in \mathbb{B}(\mathbb{R})\}$
- $f(\theta) = \begin{cases} \frac{1}{2\pi} & \theta \in [0, 2\pi] \\ 0 & \text{altrove} \end{cases}$

Si definiscono i casi:

- $A_1 = \left[0, \frac{\pi}{2}\right]$
- $A_2 = \left[\pi, 2\pi\right]$
- $A_3 = \left[\frac{\pi}{4}\right]$

Si ha:

- $P(A_1) = \int_0^{\frac{\pi}{2}} f(\theta) d\theta = \int_0^{\frac{\pi}{2}} \frac{1}{2\pi} d\theta = \frac{1}{2\pi} \cdot \frac{\pi}{2} = \frac{1}{4}$
- $P(A_2) = \int_{\pi}^{2\pi} \frac{1}{2\pi} d\theta = \frac{1}{2\pi} (2\pi - \pi) = \frac{1}{2}$
- $P(A_3) = \int_{\frac{\pi}{4}}^{\frac{\pi}{4}} \frac{1}{2\pi} d\theta = 0$

#### Esercizio: Luca e Giorgio

Due amici, Luca e Giorgio, si danno appuntamento al bar tra le 8 e le 8:30; i due arrivano al bar in modo casuale ed indipendente.

1. Definire lo spazio di probabilità;
2. Calcolare la probabilità che Luca arrivi prima di Giorgio;
3. Sapendo che arrivati al bar si fermano per un tempo pari a  $\Delta T_L = 5$  minuti e  $\Delta T_G = 5$  minuti, calcolare la probabilità che si incontrino.

Per la soluzione dell'esercizio, si veda la pagina della soluzione.

Estratto da "http://it.wikiversity.org/wiki/Spazi\_di\_probabilit%C3%A0"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 00:41, 16 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.



# Indipendenza tra eventi

Da Wikiversità, l'università aperta.

*Definizione: Eventi indipendenti*

Dato uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , gli eventi  $A$  e  $B$  si dicono indipendenti se

$$P(A \cap B) = P(A) \cdot P(B)$$

*Teorema:*

Se due eventi  $A$  e  $B$  sono indipendenti, allora sono indipendenti anche

- $A$  e  $\bar{B}$
- $\bar{A}$  e  $B$
- $\bar{A}$  e  $\bar{B}$

*Dimostrazione:  $A$  e  $\bar{B}$  sono indipendenti*

Si ha:

$$A = (A \cap B) \cup (A \cap \bar{B})$$

da cui

$$P(A) = P(A \cap B) + P(A \cap \bar{B})$$

dove

$$P(A \cap \bar{B}) = P(A)P(\bar{B}) = P(A)(1 - P(B))$$

E' importante notare la differenza tra i due concetti di eventi indipendenti ed eventi disgiunti.

*Esempio: Caso con  $A$  e  $B$  disgiunti*

Con  $A$  e  $B$  disgiunti si ha intersezione tra i due insiemi nulla,

$$A \cap B = \emptyset$$

In questo caso, si ha

$$P(A \cap B) = P(\{\emptyset\}) = 0 \neq P(A)P(B)$$

Perché sia  $P(A)$  che  $P(B)$  possono essere non nulli.

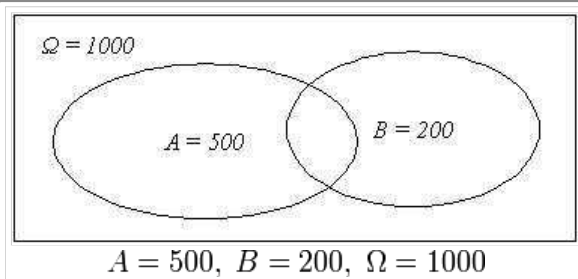
*Esempio: Caso con  $A \cap B \neq \{\emptyset\}$*

In questo caso, a priori non si può dire nulla, perchè potrebbe esserci indipendenza come potrebbe non esserci.

*Definizione: Eventi mutuamente esclusivi*

Quando l'intersezione tra due eventi  $A$  e  $B$  è nulla, i due eventi si dicono mutuamente esclusivi.

*Esempio:*



Si ha, per definizione,

$$P(A \cap B) = \frac{1}{10}$$

$$P(A) = \frac{1}{2}$$

$$P(B) = \frac{1}{5}$$

Allora

$$P(A)P(B) = \frac{1}{2} \cdot \frac{1}{5} = \frac{1}{10}$$

*Siccome il prodotto delle due probabilità coincide con la probabilità dell'intersezione, allora i due eventi sono indipendenti.*

*Esempio: Lancio ripetuto di una moneta*

Lanciamo due volte una moneta. Si ha

- $\Omega = \{TT, TC, CT, CC\}$
- $F = 2^\Omega$  insieme delle parti
- $P(\{TT\}) = P(\{TC\}) = P(\{CT\}) = P(\{CC\})$

Definiamo i due eventi

- $A =$  esce testa al primo lancio  $= \{\{TC\}, \{TT\}\}$
- $B =$  esce testa al secondo lancio  $= \{\{CT\}, \{TT\}\}$

Si ha

$$A \cap B = \{TT\}$$

Vogliamo sapere se i due eventi  $A$  e  $B$  sono indipendenti. Si ha

$$P(A) = P(\{TT\}) + P(\{TC\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} = P(B)$$

$$P(B) = P(\{TT\}) + P(\{CT\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} = P(B)$$

$$P(A \cap B) = P(\{TT\}) = \frac{1}{4}$$

$$P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Da questo, si ha che i due eventi sono effettivamente indipendenti tra loro.

- Cosa cambia se usiamo una moneta truccata?
- Cosa cambia se i lanci non sono indipendenti?

Sia

$$P(\{TT\}) = P(\{TC\}) = P(\{CT\}) = \frac{1}{6}$$

$$P(\{CC\}) = \frac{3}{6}$$

*In questo caso, le tabelle di probabilità non rispettano più l'indipendenza tra eventi, quindi  $A$  e  $B$  non sono indipendenti. Si verifichi per esercizio.*

## Indice

- 1 Eventi indipendenti multipli
- 2 Probabilità condizionata
- 3 Legge della probabilità composta
- 4 Probabilità totale
- 5 Indipendenza condizionata tra eventi
- 6 Spazi di probabilità prodotto
  - 6.1 spazio di probabilità prodotto come modello probabilistico per eventi indipendenti

## Eventi indipendenti multipli

*Definizione: Eventi multipli indipendenti*

Dati  $(\Omega, \mathcal{F}, P)$  e gli eventi

$$A_1, A_2, \dots, A_n \in \mathcal{F}$$

gli eventi  $A_i$  si dicono indipendenti se

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i) \quad \forall \text{ sottoinsieme di indici } I$$

Nota: dalla definizione si deduce che non basta verificare

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = \prod_{i=1}^n P(A_i)$$

ma bisogna controllare che questo sia vero per ogni sottoinsieme di indici  $I$ .

*Esempio: Lancio della moneta*

Siano gli eventi

- $A$  = testa al primo lancio
- $B$  = croce al primo lancio
- $C$  =  $\emptyset$

Si ha

$$P(A \cap B \cap C) = P(\emptyset) = 0 = P(A)P(B)P(C)$$

Da questa prima verifica, potrebbe sembrare che i tre eventi siano indipendenti, ma in realtà non lo sono, perchè

$$P(A \cap B) = P(\emptyset) = 0 \neq P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Quindi,  $A$ ,  $B$  e  $C$  sono sono indipendenti.

## Probabilità condizionata

*Definizione: Probabilità condizionata*

Dato lo spazio di probabilità  $(\Omega, \mathcal{F}, P)$  si considerino  $A, B \in \mathcal{F}$  con  $P(B) \neq 0$ . Si definisce probabilità condizionata di  $A$  dato  $B$  la probabilità

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

*Teorema:*

Fissato  $B \in \mathcal{F}$  con  $P(B) \neq 0$ , la funzione

$$P(\cdot|B) : \mathcal{F} \rightarrow [0, 1]$$

definisce una misura di probabilità di  $\mathcal{F}$ .

*Dimostrazione:*

1. Siano  $A_1, A_2, \dots, A_k \in \mathcal{F}$  disgiunti a coppie. Allora

$$P\left(\bigcup_k A_k | B\right) = \frac{P\left(\bigcup_k A_k \cap B\right)}{P(B)} = \sum_k \frac{P(A_k \cap B)}{P(B)}$$

2. Si ha

$$P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

**Teorema:**Stiano gli eventi  $A, B_1, B_2 \in \mathcal{F}$  e sia  $P(B_1 \cap B_2) \neq 0$ . Si ha

$$P(A|B_1, B_2) = P(A|B_1 \cap B_2) = \frac{P(A \cap B_1 \cap B_2)}{P(B_1 \cap B_2)}$$

**Teorema:**Se gli eventi  $A$  e  $B$  sono indipendenti, si ha

$$P(A|B) = \frac{P(A)P(B)}{P(B)} = P(A)$$

cioè, la probabilità a priori e a posteriori sono identiche.

**Esempio: Il lancio del dado**

Calcolare la probabilità che sia uscito il 6, sapendo che è uscito un numero pari. Si ha

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$F = 2^\Omega$$

$$P(k) = \frac{1}{6} \quad \forall k = 1, 2, \dots, 6$$

Si ha

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

Si noti che si è ottenuto  $P(A|B) = \frac{1}{3} > P(A) = \frac{1}{6}$ . Si dice che l'evento  $A$  è attratto dall'evento  $B$ . Al contrario, con l'evento  $C = \{1\}$ , si ha

$$P(C|B) = \frac{P(C \cap B)}{P(B)} = 0$$

In questo caso, si è ottenuto che  $P(C) > P(C|B)$ , quindi si dice che  $C$  è respinto da  $B$ .**Esempio: Esempio di applicazione della probabilità condizionata**

Si veda la pagina dell'esempio.

## Legge della probabilità composta

**Definizione:**Dati  $(\Omega, \mathcal{F}, P)$  e  $A, B \in \mathcal{F}$ , si ha

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

La relazione vale anche se una delle due probabilità è nulla, infatti

$$P(A \cap B) \leq \min(P(A), P(B))$$

**Teorema: Regola della catena**Dati  $(\Omega, \mathcal{F}, P)$  e gli eventi

$$A_k \in \mathcal{F}, \quad k = 1, 2, \dots, n$$

si ha

$$\begin{aligned} P(A_1 \cap A_2 \cap \dots \cap A_n) \\ = P(A_n | A_{n-1} \cap A_{n-2} \cap \dots \cap A_2 \cap A_1) \cdot P(A_{n-1} | A_{n-2} \cap A_{n-3} \cap \dots \cap A_2 \cap A_1) \cdot \dots \cdot P(A_2 | A_1) \cdot P(A_1) \end{aligned}$$

**Dimostrazione:**

Coincide con l'imporre che

$$B = (A_{n-1} \cap A_{n-2} \cap \dots \cap A_1)$$

e dire

$$P(A_n, B) = P(A_n|B)P(B)$$

iterativamente. L'equazione finale si ricava per passi successivi.

## Probabilità totale

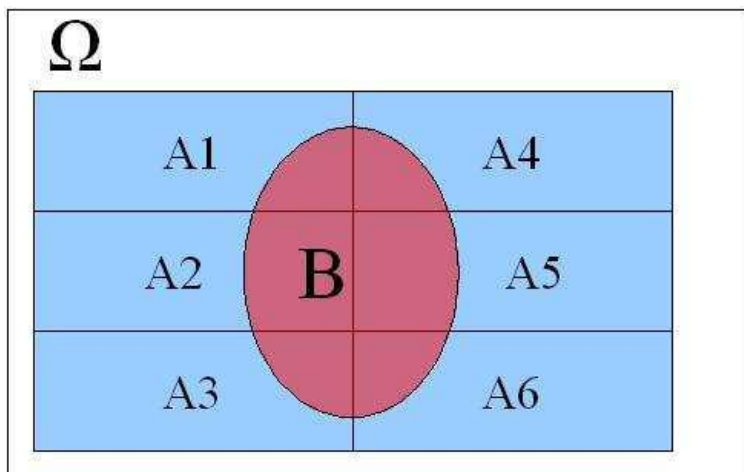
**Teorema: Teorema della probabilità totale**

Sia  $(\Omega, \mathcal{F}, P)$  e siano  $A_1, A_2, \dots, A_n \in \mathcal{F}$  eventi mutuamente esclusivi (o disgiunti a coppie,  $A_i \cap A_j = \emptyset \forall i \neq j$ ) di una  $\sigma$ -algebra. Sia  $B \in \mathcal{F}$  un evento tale che

$$B \subseteq \bigcup_{i=1}^n A_i$$

Allora, si ha

$$P(B) = \sum_{i=1}^n P(B|A_i) \cdot P(A_i)$$



Nota: non è necessario, per il teorema della probabilità totale, che

$$\bigcup_{i=1}^n A_i = \Omega$$

**Teorema: Teorema di Bayes**

Sia  $(\Omega, \mathcal{F}, P)$  uno spazio di probabilità,  $A, B \in \mathcal{F}$  con  $P(B) \neq 0$ . Allora si ha

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Se consideriamo  $A_1, A_2, \dots, A_n \in \mathcal{F}$  eventi disgiunti a coppie e tali che

$$B \subseteq \bigcup_{i=1}^n A_i$$

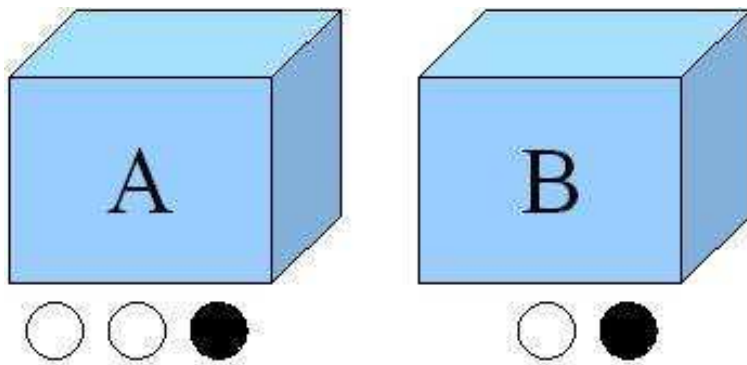
allora, si ha

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)}$$

**Esempio: Scatole e palline**

Ci sono due scatole A e B:

- in A ci sono due palline bianche ed una pallina nera;
- in B ci sono una pallina bianca ed una pallina nera.



Studiare il sistema; qual è la probabilità di pescare una pallina bianca?

Per la soluzione, si veda la pagina di soluzione.

## Indipendenza condizionata tra eventi

*Definizione: Eventi condizionatamente indipendenti*

Dato  $(\Omega, \mathcal{F}, P)$ , gli eventi  $A, B \in \mathcal{F}$  si dicono condizionatamente indipendenti da un evento  $C \in \mathcal{F}$  se, dato

$$P(C) \neq 0$$

si ha

$$P(A \cap B | C) = P(A | C) \cdot P(B | C)$$

E' da notare che l'indipendenza condizionale non implica l'indipendenza tra  $A$  e  $B$ ; vale anche il viceversa. Banalmente, si verifica se  $C = \Omega$ .

*Esempio:*

Si hanno due lanci consecutivi di una moneta. Vediamo che

$$A, B \text{ indipendenti} \not\Rightarrow A, B \text{ indipendenti} | C$$

Si ha:

- $\Omega = TC, CT, TT, CC$
- $\mathcal{F} = 2^\Omega$
- $P(\{s\}) = 1/4 \forall s \in \Omega$

Si ha

- $A = \{TC, TT\}$  testa al primo lancio
- $B = \{CT, TT\}$  testa al secondo lancio

Allora,

$$P(A \cap B) = P(\{TT\}) = 1/4 \text{ e } P(A) = 1/2 = P(B)$$

da cui si ha che  $A$  e  $B$  sono statisticamente indipendenti  $\left(\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}\right)$ .

Sia  $C = \{TC, CT\}$  evento in cui esce una sola testa.  $P(C) = \frac{1}{2}$ . Verifichiamo se  $P(A \cap B | C) = P(A | C)P(B | C)$ .

Si ha:

- $P(A \cap B | C) = \frac{P(A \cap B \cap C)}{P(C)} = \frac{P(\emptyset)}{1/2} = 0$
- $P(A | C) = \frac{P(A \cap C)}{P(C)} = \frac{1/4}{1/2} = \frac{1}{2} \neq 0$
- $P(B | C) = \frac{P(B \cap C)}{P(C)} = \frac{1/4}{1/2} = \frac{1}{2} \neq 0$

Quindi, concludendo,  $P(A | C)P(B | C) = 1/4 \neq 0 = P(A \cap B | C)$ . Ne risulta che  $A$  e  $B$  non sono indipendenti dato  $C$ .

## Spazi di probabilità prodotto

**Definizione: Spazio misurabile prodotto**

Dati gli spazi misurabili  $(\Omega_k, F_k)$ ,  $k = 1, 2, \dots, n$ , si dice spazio misurabile prodotto lo spazio

$$(\Omega_1 \times \Omega_2 \times \dots \times \Omega_n, F_1 \otimes F_2 \otimes \dots \otimes F_n)$$

dove

1. il prodotto diretto

$$\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$$

è l'insieme delle  $n$ -uple

$$(s_1, \dots, s_n) | s_k \in \Omega_k, k = 1, \dots, n$$

2. il prodotto diretto

$$(F_1 \otimes F_2 \otimes \dots \otimes F_n)$$

è la più piccolo  $\sigma$ -algebra definita sull'insieme

$$\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$$

che contiene tutti gli insiemi nella forma

$$A_1 \times A_2 \times \dots \times A_n = \{(s_1, \dots, s_n) | s_k \in A_k\}$$

con  $A_k \in F_k$  e  $k = 1, 2, \dots, n$ .

**Esempio:**

Considero  $(\mathbb{R}^n, \mathbb{B}(\mathbb{R}^n))$  dove

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$$

$n$  volte e

$$\mathbb{B}(\mathbb{R}^n) = \mathbb{B}(\mathbb{R}) \otimes \mathbb{B}(\mathbb{R}) \otimes \dots \otimes \mathbb{B}(\mathbb{R})$$

$n$  volte.

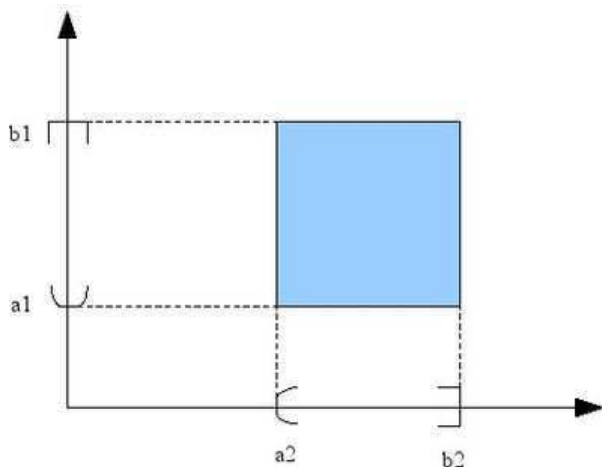
Nel caso  $n = 2$ , si ha

$$\mathbb{R}^2 = \{(\alpha, \beta) | \alpha \in \mathbb{R}, \beta \in \mathbb{R}\}$$

$$\mathbb{B}(\mathbb{R}^2) = \mathbb{B}(\mathbb{R}) \otimes \mathbb{B}(\mathbb{R})$$

è la più piccolo  $\sigma$ -algebra che contiene tutti i rettangoli i cui lati sono dei borelliani.

$$I = (a_1, b_1] \times (a_2, b_2]$$

**Definizione: Misura prodotto**

Si considerino gli spazi dotati di misura  $\sigma$ -finita

$$(\Omega_k, F_k, m_k), k = 1, 2, \dots, n$$

Sia

$$(\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n, F_1 \otimes F_2 \otimes \cdots \otimes F_n)$$

lo spazio misurabile prodotto. Definiamo su di esso la misura prodotto

$$m_1 \cdot m_2 \cdot m_3 \cdots m_n.$$

Consideriamo il caso  $n=2$  per semplicità. Sia  $G$  l'algebra generata dagli insiemi ottenuti come unione finita di insiemi disgiunti di forma  $A_1 \times A_2$  con  $A_1 \in F_1, A_2 \in F_2$ . Allora

$$A_i \in G \Rightarrow \bigcup_{k=1}^P (A_{1k} \times A_{2k}), \quad A_{1k} \in F_1, \quad A_{2k} \in F_2, \quad k = 1, \dots, P$$

Inoltre,

$$(A_{1i} \times A_{2i}) \cap (A_{1j} \times A_{2j}) = \emptyset \quad \forall i \neq j$$

Si ha  $\sigma(G) = F$ , perchè

- $A_1 \times A_2 \in \sigma(G)$
- $F_1 \otimes F_2$  è la più piccola che contiene insiemi della forma  $A_1 \times A_2$

Consideriamo

$$m_0 : G \rightarrow [0, +\infty] \mid m_0(A) = \sum_{k=1}^P m_1(A_{1k}) \cdot m_2(A_{2k})$$

Si può dimostrare che questa è una premisura  $\sigma$ -finita su  $G$ . L'estensione di  $m_0$  sulla  $\sigma$ -algebra  $F_1 \otimes F_2 = \sigma(G)$  è detta misura prodotto ed è indicata con  $m_1 \times m_2$ .

*Definizione: Spazio prodotto dotato di misura*

*Dati  $n$  spazi dotati di misura  $\sigma$ -finita,*

$$(\Omega_k, F_k, m_k), \quad k = 1, \dots, n$$

*lo spazio prodotto dotato di misura*

$$(\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n, F_1 \otimes F_2 \otimes \cdots \otimes F_n, m_1 \times m_2 \times \cdots \times m_n)$$

### spazio di probabilità prodotto come modello probabilistico per eventi indipendenti

*Teorema:*

*Consideriamo due esperimenti casuali  $(\Omega_1, F_1, P_1)$  e  $(\Omega_2, F_2, P_2)$ . Per descrivere congiuntamente i due esperimenti, consideriamo lo spazio di probabilità prodotto  $(\Omega_1 \times \Omega_2, F_1 \otimes F_2, P_1 \times P_2)$ . Tale spazio è adatto a descrivere congiuntamente eventi indipendenti.*

*Dimostrazione:*

*Consideriamo*

$$A_1 \times \Omega_2, \quad \Omega_1 \times A_2 \in F_1 \otimes F_2$$

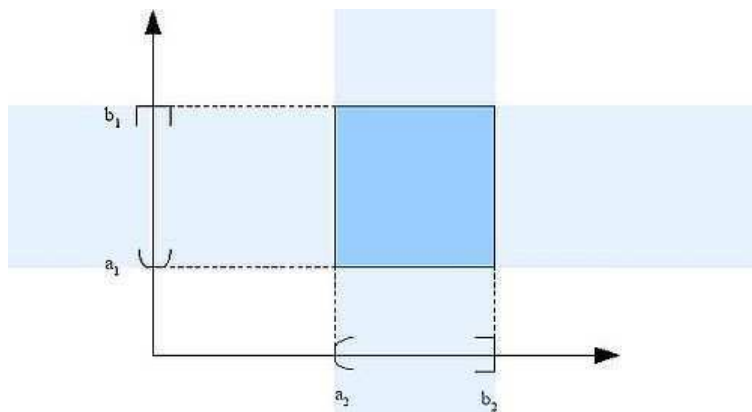
*e dimostriamo che sono indipendenti, il che equivale a dire*

$$P((A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)) = P((A_1 \times \Omega_2)) \cdot P((\Omega_1 \times A_2))$$

*Si ha:*

- $P(A_1 \times \Omega_2) = P(A_1) \cdot P(\Omega_2) = P(A_1)$
- $P(\Omega_1 \times A_2) = P(\Omega_1) \cdot P(A_2) = P(A_2)$
- $P((A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)) = P(A_1 \times A_2) = P(A_1) \cdot P(A_2)$





*Questo verifica che i due eventi sono indipendenti, infatti il prodotto delle marginali è sufficiente calcolare la probabilità totale soltanto in questo caso.*

Estratto da "[http://it.wikiversity.org/wiki/Indipendenza\\_tra\\_eventi](http://it.wikiversity.org/wiki/Indipendenza_tra_eventi)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 02:57, 16 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Prove ripetute

Da Wikiversità, l'università aperta.

Le **prove ripetute** sono un caso particolare di prove indipendenti, infatti si ripetono  $n$  prove indipendenti dello stesso esperimento casuale  $(\Omega_1, F_1, P_1)$ .

Lo spazio di probabilità è  $(\Omega, F, P)$  dove

- $\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$
- $F = F_1 \times F_2 \times \dots \times F_n$
- $P = P_1 \times P_2 \times \dots \times P_n$

## Prove bernoulliane

Le prove bernoulliane sono un sottocaso delle prove ripetute che rispondono alla domanda *"l'evento si verifica o no?"*. Nel settore delle telecomunicazioni, potrebbe essere per esempio *"l'errore si verifica o no?"*.

Le prove bernoulliane sono delle prove ripetute in cui si pone l'attenzione su un evento  $A \in F_1$ , chiamato successo, che si verifica con probabilità

$$P_1(A) = p$$

in ogni singola prova.  $\bar{A}$  è detto insuccesso e si verifica con probabilità

$$P_1(\bar{A}) = 1 - p = q$$

Dato che  $A$  è l'evento di interesse, si può considerare

$$F_1 = \{\emptyset, A, \bar{A}, \Omega_1\}$$

con

- $P_1(A) = p$
- $P_1(\bar{A}) = q$

In generale, si vuole determinare la probabilità  $P_n(k)$  che, su  $n$  prove, l'evento  $A$  si verifichi  $k$  volte in un qualunque ordine.

Consideriamo lo spazio di probabilità prodotto  $(\Omega, F, P)$  associato alle  $n$  prove. L'evento di interesse è

$$B = \{B_1 \times B_2 \times \dots \times B_n\} \text{ t.c. } B_i \in A, i \in I, B_j \in \bar{A}, j \in \bar{I}$$

dove  $I$  è un qualsiasi insieme di indici con cardinalità  $k$ .

*Esempio:*

Prendiamo  $B_1 \times B_2 \times B_3$ , cioè ( $k = 3$ ). Allora vogliamo

$$B_1, B_2, B_3 \in A$$

mentre i successivi

$$B_4, B_5, B_6 \in \bar{A}$$

Questa proprietà non dipende dall'ordine con cui si susseguono gli eventi, ma soltanto dal numero di eventi positivi alla fine dei 6 esperimenti. Si considera l'evento di interesse.

*Esempio:*

Cerchiamo  $n = 3, k = 2$ . Si hanno

- $A \times \bar{A} \times \bar{A}$
- $\bar{A} \times \bar{A} \times A$
- $\bar{A} \times A \times A$

Sono 3 configurazioni possibili per i risultati. Allora, la probabilità è  $3/8$ .

Nel caso di prove ripetute bernoulliane, il numero di configurazioni con  $k$  volte  $A$  e  $(n - k)$  volte  $\bar{A}$  è pari al valore

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Il numero delle possibili combinazioni di questo tipo è pari a  $\binom{n}{k}$ . Indichiamo con  $B_i$  una di queste possibili configurazioni e  $B(k)$  l'evento che contiene tutte le possibili combinazioni favorevoli,

$$B(k) = \bigcup_{i=1}^{\binom{n}{k}} B_i$$

La probabilità di questo  $B(k)$  è

$$P((A \times A \times \dots \times A \times \bar{A} \times \bar{A} \times \dots \times \bar{A})) = p^k \cdot q^{n-k}$$

cioè

$$P_n(k) = P(B(k)) = \sum_{i=1}^{\binom{n}{k}} P(B_i) = \sum_{i=1}^{\binom{n}{k}} p^k \cdot q^{n-k} = \binom{n}{k} \cdot p^k \cdot q^{n-k}$$

Gli eventi  $B(k), k = 0, 1, \dots, n$  sono tra di loro disgiunti, con

$$\Omega = \bigcup_{k=0}^n B(k)$$

Dato che  $P_n(K)$  soddisfa l'assioma di probabilità

$$1 = P(\Omega) = \sum_k P(B_k) = \sum_k \binom{n}{k} p^k q^{n-k}$$

allora si ha che

$$\sum_k P_n(k) = 1$$

Concludendo,  $P_n(k)$  segue una legge di probabilità binomiale.

#### *Esercizio: Codice di Hamming*

*Il codice di hamming(7,4) è un codice per la correzione degli errori sui canali binari che rappresenta una parola di 4 bit con una parola di 7 bit. Permette di correggere un errore e rilevare fino ad un massimo di due errori.*

*Vogliamo sapere la probabilità che la parola sia esatta, con probabilità di errore sul bit singolo  $P_\varepsilon(b) = 10^{-3}$ .*

*Si suppone un canale binario simmetrico indipendente.*

Per la soluzione dell'esercizio, vedere la pagina di soluzione.

#### *Esercizio: Un mazzo di carte*

*Prendete un mazzo di 52 carte ed estraete 3 carte in maniera indipendente, con reinserimento (ogni volta ci sono 52 carte).*

- 1. Costruire il modello probabilistico.*
- 2. Determinare la probabilità di pescare esattamente 2 cuori;*
- 3. Determinare la probabilità che almeno una carta sia di cuori.*

*Si risolve sulla falsa riga dell'esercizio precedente.*

Per la soluzione dell'esercizio, vedere la pagina di soluzione.

#### *Teorema: Teorema di Moivre Laplace*

*Sia  $0 < p < 1$ . Se  $npq \gg 1$ , allora*

$$P_n(k) = \binom{n}{k} p^n q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-\left(\frac{(k-np)^2}{2npq}\right)}$$

*Questa proprietà è valida per  $k$  in un intorno di larghezza  $\sqrt{npq}$  del valore  $npq$ .*

#### *Considerazioni:*

- 1. Per  $n \rightarrow \infty$ , si ha*

$$P_n(\lfloor np \rfloor) \rightarrow \text{MANCA per } n \rightarrow \infty$$

- 2. Se  $q = p = 1/2$  (per esempio, una moneta), si ha*

$$\blacksquare A = 5 \text{ teste su } 10 \text{ lanci} \Rightarrow P(A) = 0,25$$

- $B = 50 \text{ teste su } 100 \text{ lanci} \Rightarrow P(A) = 0,08$

- $C = 500 \text{ teste su } 1000 \text{ lanci} \Rightarrow P(A) = 0,025$

3. La probabilità si accumula in un intorno  $[np]$  al tendere di  $n \rightarrow \infty$ .

- $D = 4 \text{ teste su } 6 \text{ lanci} \Rightarrow P(A) = 0,66$

- $E = 40 \text{ teste su } 60 \text{ lanci} \Rightarrow P(A) = 0,96$

- $F = 400 \text{ teste su } 600 \text{ lanci} \Rightarrow P(A) = 0,99$

Estratto da "[http://it.wikiversity.org/wiki/Prove\\_ripetute](http://it.wikiversity.org/wiki/Prove_ripetute)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

---

- Ultima modifica per la pagina: 14:44, 16 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Variabili casuali

Da Wikiversità, l'università aperta.

Le **variabili casuali** sono funzioni che permettono di associare un numero al risultato di un esperimento. Questo viene fatto per riportare tutto in uno spazio matematico, quindi si possono usare tutti gli strumenti matematici noti.

*Esempio:*

Consideriamo come esempio il caso in cui si vuole monitorare l'apertura e la chiusura di 100 porte di un centro commerciale. Se vediamo lo spazio di probabilità classico, avremo, considerando una sola porta,  $\Omega = \{\text{aperta}, \text{chiusa}\}$ . Queste porte potrebbero essere indipendenti.

Considerando lo spazio a 100 dimensioni,

$$\Omega' = \{(P_1, P_2, \dots, P_{100}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{100}\}$$

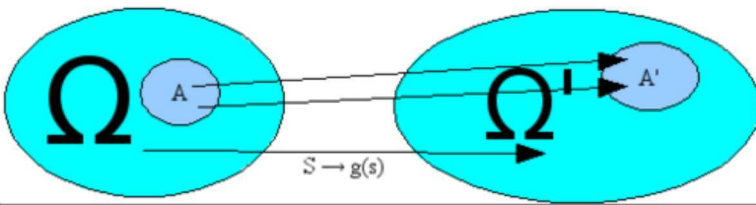
Se voglio solo conoscere il numero medio di aperture, basta definire una funzione, la nostra variabile casuale, che, dato il vettore di 100 elementi, ci restituisce il numero di porte aperte. Abbiamo spostato il problema da uno spazio complicato a dei semplici valori in  $\mathbb{R}$  di solito, in questo caso in  $\mathbb{N}$ .

*Definizione: Funzione misurabile*

Dati due spazi misurabili  $(\Omega, F)$  e  $(\Omega', F')$ , una funzione

$$g: \Omega \rightarrow \Omega'$$

si dice misurabile se  $\forall A' \in F'$ , la controimmagine di  $A'$  attraverso  $g(\cdot)$  appartiene ad  $F$ .



Un boreliano  $B \in \mathbb{B}(\mathbb{R})$  è ad esempio un insieme della forma  $(a,b), [a,b], (a,b)$  con  $a < b$  in generale. La condizione di misurabilità

$$\{s \in \Omega \text{ t.c. } X(s) \in B'\} \in F \quad \forall B' \text{ in } \mathbb{B}(\mathbb{R})$$

consente di attribuire una probabilità agli eventi specificati dai valori assunti dalla variabile casuale.

*Esempi:*

1.  $X < b$
2.  $a < X < b$

Questo perché  $X \leq B$  va inteso come

$$\{s \in \Omega \text{ t.c. } X(s) \leq b\}$$

e la condizione di misurabilità di  $X$  sommato al fatto che  $(-\infty, b]$  è un boreliano assicurano che  $\{s \in \Omega \text{ t.c. } X(s) \leq b\}$  sia un evento in  $F$ . Allora, si potrà scrivere che

$$P(X \leq b) \hat{=} P(\{s \in \Omega \text{ t.c. } X(s) \in (-\infty, b]\})$$

Questo è ben definito, perché

$$P: F \rightarrow [0, 1] \text{ e } A \in F.$$

Se la funzione misurabile è effettivamente una variabile casuale, per ogni evento nello spazio di arrivo possiamo trovare una controimmagine nello spazio di probabilità originario. Nella controimmagine  $A$  abbiamo definito la probabilità, quindi è possibile trovare sempre un valore di probabilità associato al valore della variabile casuale e viceversa. Quello che non deve succedere è che, tornando indietro, venga generata una controimmagine che non appartiene ad  $F$ .

*Definizione: Funzione indicatrice*

Dato lo spazio di probabilità  $\{\Omega, F, P\}$ , consideriamo l'evento  $A \in F$  come l'unico evento che ci interessa. Definita la funzione

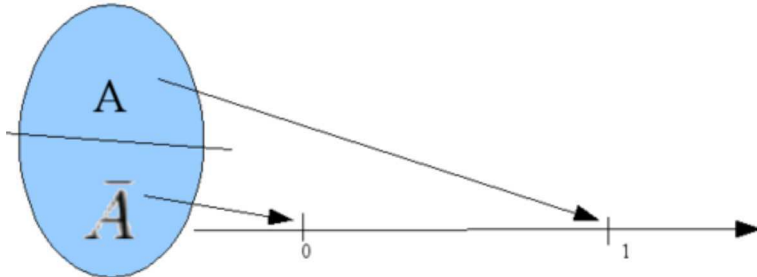
$$X : \Omega \rightarrow \mathbb{R}$$

come

$$X(s) = \begin{cases} 1 & s \in A \\ 0 & s \in \bar{A} \end{cases}$$

dobbiamo verificare che  $X$  è una variabile casuale.

Consideriamo il generico boreliano  $B' \in \mathbb{B}(\mathbb{R})$  e determiniamo la controimmagine data da  $\{s \in \Omega \text{ t.c. } X(s) \in B'\}$ . Graficamente...



Avremo che

$$B = \{s \in \Omega \text{ t.c. } X(s) \in B'\} = \begin{cases} A & 0 \notin B', 1 \in B' \\ \bar{A} & 0 \in B', 1 \notin B' \\ \emptyset & 0, 1 \notin B' \\ \Omega & 0, 1 \in B' \end{cases}$$

$X(s)$  è indicata anche come  $I_A(s)$  ed è detta funzione indicatrice dell'evento  $A$ .

A conclusione di questo, dato che  $B \in \mathcal{F} \forall B' \in \mathcal{F}'$ , allora  $I_A(\cdot)$  è una variabile casuale.

Notare che l'immagine di  $\Omega$  attraverso  $I_A(\cdot)$  è un insieme finito, quindi  $X$  è una variabile casuale discreta.

*Esempio: Lancio una moneta*

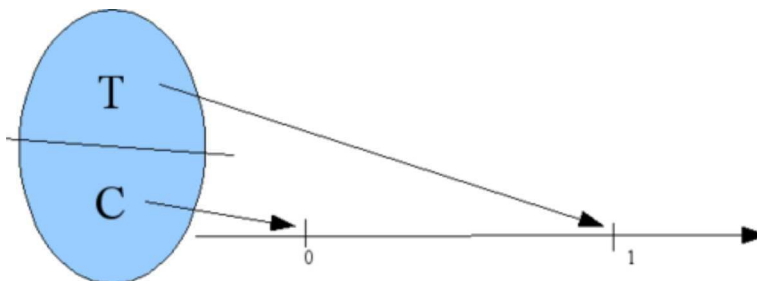
$$\Omega = \{T, C\}$$

$$\mathcal{F} = \{T, C, \Omega, \emptyset\}$$

$$P: \mathcal{F} \rightarrow [0, 1] \text{ con } P(\{T\}) = P(\{C\}) = 1/2$$

Consideriamo  $X$  come:

$$X(s) = \begin{cases} 1 & s = T \\ 0 & s = C \end{cases}$$



$$B = \{s \in \Omega \text{ t.c. } X(s) \in B'\} = \begin{cases} \Omega & 0, 1 \in B' \\ \emptyset & 0, 1 \notin B' \\ T & 1 \in B', 0 \notin B' \\ C & 0 \in B', 1 \notin B' \end{cases}$$

$B \in \mathcal{F} \forall B' \in \mathbb{B}(\mathbb{R}) = \mathcal{F}' \Rightarrow X$  una variabile casuale

Come si può verificare che  $X : \Omega \rightarrow \mathbb{R}^n$  definita su uno spazio di probabilità sia una variabile casuale? La verifica basata sulla definizione appena vista è onerosa, perché bisogna considerare tutti i boreliani  $B' \in \mathbb{B}(\mathbb{R})$  e calcolare le controimmagini attraverso  $X$ , verificando che appartengano a  $F$ . Il seguente lemma permette di restringere la verifica ad un sottoinsieme di boreliani.

**Lemma:**

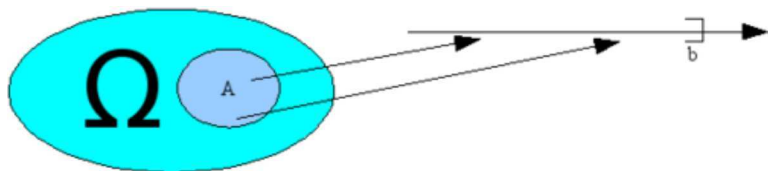
Sia  $\xi$  una collezione di insiemi in  $\mathbb{R}$  tali che  $\sigma(\xi) = \mathbb{B}(\mathbb{R}) = F'$ . Dato  $\{\Omega, F, P\}$ , condizione necessaria sufficiente affinché  $X : \Omega \rightarrow \mathbb{R}^n$  sia una variabile casuale è che

$$\{s \in \Omega \text{ t.c. } X(s) \in E\} \in F \quad \forall E \in \xi$$

Se uno spazio è discreto, basta definire le probabilità sugli eventi elementari della  $\sigma$ -algebra. Se  $\Omega$  è continuo, è impossibile definire le probabilità per ogni elemento della  $\sigma$ -algebra. Avevamo definito una probabilità su un'algebra che ci permette di generare la  $\sigma$ -algebra di interesse, posso estendere la misura di probabilità alla  $\sigma$ -algebra. La stessa cosa accade qui: andiamo a trovare la controimmagine per un sottoinsieme di boreliani, e possiamo poi estendere a tutti gli eventi che sono nell'insieme boreliano. È la stessa cosa. Usiamo la funzione  $X$  per generare il nostro sottoinsieme di eventi, verifichiamo che esistano le controimmagini sul sottoinsieme definito ed abbiamo finito.

Abbiamo visto che  $\mathbb{B}(\mathbb{R})$  può essere costruita come la più piccola  $\sigma$ -algebra che contiene insiemi del tipo  $(-\infty, b] \in \mathbb{R}$ . Dati quindi  $\{\Omega, F, P\}$  e  $X : \Omega \rightarrow \mathbb{R}$ ,  $X$  è una variabile casuale se

$$\{s \in \Omega \text{ t.c. } X(s) \leq b\} \in F \quad \forall b \in \mathbb{R}$$



**Esempio:**

$$\{\Omega, F, P\}$$

$$A \in F$$

$$X(s) = I_A(s) = \begin{cases} 1 & s \in A \\ 0 & s \in \bar{A} \end{cases}$$

$$B = \{s \in \Omega \text{ t.c. } X(s) \leq b\} = \begin{cases} \Omega & b \geq 1 \\ A & 0 \leq b < 1 \\ \emptyset & b < 0 \end{cases}$$

**Esercizio: Esercizio per lo studente**

Dato  $(\mathbb{R}, \mathbb{B}(\mathbb{R}), P)$ , verificare che  $X(s) = s^2$  è una variabile casuale, usando quest'ultimo metodo. La soluzione dell'esercizio si trova alla pagina soluzione.

Se abbiamo una variabile casuale  $n$ -dimensionale, allora abbiamo  $n$  variabili casuali e viceversa. Questo non vuol dire che se abbiamo la densità congiunta di due variabili casuali, queste sono indipendenti. Bisogna prima calcolare le marginali e poi lavorare su quelle.

**Teorema:**

Consideriamo la variabile casuale  $n$ -dimensionale  $X$  su  $\{\Omega, F, P\}$  |  $X : \Omega \rightarrow \mathbb{R}^n$  misurabile da  $(\Omega, F)$  a  $(\mathbb{R}, \mathbb{B}(\mathbb{R}^n))$ ,  $X$  ha  $n$  componenti lungo gli assi coordinati  $X = (X_1, X_2, \dots, X_n)$  con  $X_i : \Omega \rightarrow \mathbb{R}$ , allora ogni componente  $X_i$  è una variabile casuale.

**Dimostrazione:**

Sia  $B \in \mathbb{B}(\mathbb{R})$ . Dimostriamo che  $\{s \in \Omega | X(s) \in B\} \in F \rightarrow (l, F, P)$ .

$$X_i(s) \in B \Leftrightarrow X_1(s) \in \mathbb{R}, X_2(s) \in \mathbb{R}, \dots, X_{i-1}(s) \in \mathbb{R}, X_i(s) \in B, X_{i+1}(s) \in \mathbb{R}, \dots, X_n(s) \in \mathbb{R}$$

Questo è vero se e solo se

$$X(s) = (X_1(s), X_2(s), \dots, X_n(s)) \in \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$$

Da cui si deve avere

$$B = \mathbb{R} \times \dots \times B \times \dots \times \mathbb{R} \in \mathbb{B}(\mathbb{R}^n) \Rightarrow \{s \in \Omega | X(s) \in B\} \in F$$



cioè

$$\{s \in \Omega | X_i(s) \in B\} = \{s \in \Omega | X(s) \in B'\} \in F$$

da cui segue la tesi.

*Dimostrazione del viceversa:*

Se consideriamo  $n$  variabili casuali

$$X_1, X_2, \dots, X_n$$

definite su

$$(\Omega, F, P) | X_i : \Omega \rightarrow \mathbb{R}$$

misurabili da  $(\Omega, F)$  a  $(\mathbb{R}, \mathbb{B}(\mathbb{R}^n))$ , allora il vettore

$$(X_1, X_2, \dots, X_n)$$

è una variabile casuale  $n$ -dimensionale. Questa verifica è banale ed è lasciata per esercizio.

## Indice

- 1 Distribuzioni delle variabili casuali
  - 1.1 Variabili casuali notevoli continue
    - 1.1.1 Esponenziale di parametro  $\lambda > 0$
    - 1.1.2 Esponenziale bilatera, o di Laplace
    - 1.1.3 Variabile casuale di Rayleigh di parametro  $b > 0$
    - 1.1.4 Mixture gaussiane
  - 1.2 Variabili casuali notevoli discrete
    - 1.2.1 Variabile casuale di Bernoulli
    - 1.2.2 Binomiale di parametri  $p$  e  $n$
    - 1.2.3 Variabile casuale poissoniana

## Distribuzioni delle variabili casuali

*Definizione: Distribuzione di una variabile casuale*

Dato  $(\Omega, F, P)$ , sia  $X$  una variabile casuale,  $X : \Omega \rightarrow \mathbb{R}$ . Si definisce funzione di distribuzione della variabile casuale  $X$

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

data da

$$F_X(\alpha) = P(X \leq \alpha) = P(\{s \in \Omega | X(s) \leq \alpha\})$$

È esattamente la stessa cosa che avevamo fatto per la funzione di distribuzione delle distribuzioni di probabilità.

Nella maggior parte dei casi la  $F_X$  si definisce direttamente nella definizione di variabile casuale.

$F_X$  soddisfa le seguenti proprietà:

1.  $F_X$  è non decrescente
2.  $F_X(-\infty) = \lim_{\alpha \rightarrow -\infty} F_X(\alpha) = 0$
3.  $F_X(\infty) = \lim_{\alpha \rightarrow \infty} F_X(\alpha) = 1$
4.  $F_X(\alpha^+) = \lim_{\epsilon \rightarrow 0^+} F_X(\alpha + \epsilon) = F_X(\alpha) \forall \alpha \in \mathbb{R}$ , cioè  $F_X$  è continua a destra;
5.  $F_X$  ammette limite sinistro, ossia
 
$$\exists \lim_{\epsilon \rightarrow 0^+} F_X(\alpha - \epsilon) = F_X(\alpha^-)$$

Tutte le funzioni che soddisfano queste proprietà sono funzioni di distribuzione di  $X$ .

Abbiamo visto che ad ogni funzione di distribuzione  $F_X$  su  $\mathbb{R}$  è associata una ed una sola misura di probabilità  $P$  che soddisfa  $P((a, b)) = F(b) - F(a)$ . Per estensione, indichiamo con  $P_X$  la misura di probabilità associata alla  $F_X$ ,

$$(P_X, \mathbb{B}(\mathbb{R})) \rightarrow [0, 1]$$

è la misura di probabilità su  $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$  tale che  $P_X((a,b]) = F_X(b) - F_X(a)$  e  $P_X((-a,b)) = F_X(b)$ .

Proprietà di  $P_X$  e  $F_X$

- $P_X(\{a\}) = F_X(a) - F_X(a^-)$
- $P_X([a, b]) = F_X(b) - F_X(a^-)$
- $P_X((a, b)) = F_X(b^-) - F_X(a)$
- $P_X([a, b)) = F_X(b^-) - F_X(a^-)$

**Teorema: Teorema di esistenza della variabile casuale**

Data una funzione di distribuzione  $F$  su  $\mathbb{R}$ , si può definire su  $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$  una misura di probabilità  $P$  e su  $(\mathbb{R}, \mathbb{B}(\mathbb{R}), P)$  una variabile casuale  $X$  che ammette  $F$  come funzione di distribuzione.

*Dimostrazione:*

Sia  $P$  la misura di probabilità su  $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$  tale che

$$P((a, b]) = F(b) - F(a) \text{ con } a, b \in \mathbb{R} \text{ } a < b$$

Sappiamo che essa esiste ed è unica. Definiamo

$$X : \mathbb{R} \rightarrow \mathbb{R} \text{ come } X(s) = s \quad \forall s \in \mathbb{R}$$

$X$  è una variabile casuale su  $(\mathbb{R}, \mathbb{B}(\mathbb{R}), P)$ ; inoltre,

$$F_X(b) = P(\{s \in \mathbb{R} | X(s) = s \leq b\}) = P((-\infty, b]) \quad \forall b \in \mathbb{R}$$

**Definizione: Funzione di distribuzione continua**

Ad  $F$  posso associare  $(\Omega, \mathcal{F}, P)$  spazio di probabilità con

- $\Omega = \mathbb{R}$
- $\mathcal{F} = \mathbb{B}(\mathbb{R})$
- $P(P((a, b]) = F(b) - F(a) \quad \forall a, b \in \mathbb{R}, \text{ } a < b.$

Sappiamo che tale  $P$  esiste ed è unica. Allora  $X(s) = s \quad \forall s \in \mathbb{R}$  è una variabile casuale che ammette  $F$  come distribuzione. Se esiste  $f : \mathbb{R} \rightarrow \mathbb{R}^+$  tale che

$$F(b) = \int_{-\infty}^b f(\alpha) d\alpha$$

allora  $F$  è detta assolutamente continua ed  $f$  è detta funzione di densità di probabilità (pdf).

Consideriamo l'evento

- $B \in \mathbb{B}(\mathbb{R})$
- $P(X \in B) = P(\{s \in \Omega | X(s) \in B\}) = \int_B f(\alpha) d\alpha$

Solitamente, si caratterizza la variabile casuale  $X$  indicando la densità di probabilità.

**Definizione: Variabile casuale uniforme**

Una variabile casuale uniforme  $X$  è uniformemente distribuita in  $(a, b)$ ,  $-\infty < a < b < \infty$  e vale

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b) \\ 0 & \text{altrove} \end{cases}$$

**Definizione: Variabile casuale gaussiana**

La variabile casuale gaussiana è

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

con  $\sigma > 0$ . Si ha

$$F(x) = \int_{-\infty}^x f(\alpha) d\alpha = G\left(\frac{x-\mu}{\sigma}\right)$$

con

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{\beta^2}{2}} d\beta$$

*Definizione: Variabile casuale  $Q(x)$*

$Q(x) = 1 - G(x)$  è la coda della gaussiana, usata per calcolare le probabilità di errore.

*Definizione: Funzione di distribuzione discreta*

La funzione di distribuzione discreta è una  $F$  costante a tratti, con:

- $P_i = F(x_i) - F(X_i^-)$
- $\sum P_i = 1$
- $F(x) = \sum_{i: x_i \leq x} P_i$

In questo caso conviene introdurre  $\Omega = \{x_1, x_2\}$ ,  $F = 2^\Omega$  e  $P(\{x_i\}) = P_i$ .

Consideriamo lo spazio  $\{\Omega, F, P\}$  e definiamo la funzione

$$X : \Omega \rightarrow \mathbb{R} | X(s) = s \quad \forall s \in \Omega$$

$X$  è una variabile casuale discreta con funzione di distribuzione  $F_X = F$ ,  $P(x = x_i) = P_i$ . A questo punto, possiamo scrivere che

$$F_X(\alpha) = P(x \leq \alpha) = P(\{s \in \Omega | X(s) \leq \alpha\}) = P(\{s \in \Omega | s \leq \alpha\}) = \begin{cases} P_1 + P_2 & \alpha > x_2 \\ P_1 & x_1 \leq \alpha < x_2 \\ 0 & \alpha < x_1 \end{cases} = \sum_{i: x_i \leq \alpha} P_i$$

*Esempio:*

Banalmente, se avete il lancio di una moneta:

- $X(T) = 1$
- $X(C) = 0$

con  $P(T) = p$  e  $P(C) = q$ . Allora la funzione di distribuzione è

$$F_X(\alpha) = \begin{cases} 1 & \alpha > 1 \\ q & 0 \leq \alpha < 1 \\ 0 & \alpha < 0 \end{cases}$$

*Esercizio:*

Dato lo spazio di probabilità  $\{\Omega, F, P\}$ , lancio di un dado non truccato, con  $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$ , considerare la seguente variabile casuale e caratterizzarla:

$$X(\omega_1) = 2$$

$$X(\omega_2) = 10$$

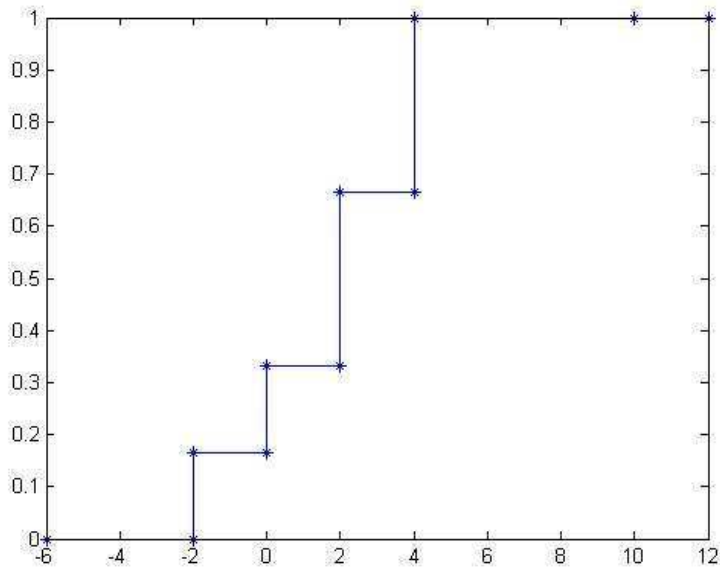
$$X(\omega_3) = 2$$

$$X(\omega_4) = 4$$

$$X(\omega_5) = 0$$

$$X(\omega_6) = -2$$

Il risultato è quello in figura:



## Variabili casuali notevoli continue

### Esponenziale di parametro $\lambda > 0$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \forall x \geq 0 \\ 0 & \text{altrove} \end{cases} \quad \lambda > 0$$

```

> x1=[-10:0.01:0];
> x2=[0.01:0.01:10];
> lambda = 1; % valore scelto arbitrariamente
> f_x = lambda .* exp(- lambda .* x2);
> plot([x1 x2], [zeros(size(x1) f_x)])

```

### Esponenziale bilatera, o di Laplace

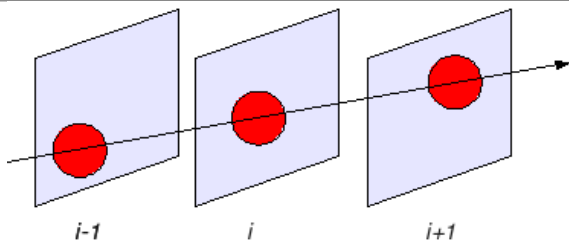
$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|} \quad \lambda > 0$$

```

> x=[-10:0.001:10];
> lambda = 1; % valore scelto arbitrariamente
> f_x = lambda/2 .* exp(- lambda .* abs(x));
> plot(x, f_x)

```

#### Esempio: La codifica predittiva



La codifica predittiva utilizza parecchio la variabile casuale di Laplace. Si vuole comprimere un video, una serie di fotogrammi  $[i - 1, i, i + 1]$ . In generale, la funzione di densità di probabilità dei pixel sarà uniforme; al contrario, però, la funzione di densità di probabilità di

$$E = I_i - I_{i-1}$$

sarà del tipo di Laplace. La funzione  $E$  è detta il predittore: migliore è il predittore, più i valori si concentrano attorno allo zero con distribuzione di Laplace.

### Variabile casuale di Rayleigh di parametro $b > 0$

$$f(x) = \begin{cases} \frac{2x}{b} \cdot e^{-\frac{x^2}{b}} & x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

```
> x1=[-10:0.01:0];
> x2=[0.01:0.01:10];
> b = 1; % valore scelto arbitrariamente
> f_x = x2 .* 2./b .* exp(-x.^2./b)
> plot([x1 x2], [zeros(size(x1)) f_x])
```

## Mixture gaussiane

Sono le variabili casuali le cui densità di probabilità si possono scrivere come combinazione lineare di gaussiane pesate.

### Esempio:

Abbiamo due contenitori di gas  $C_1$  e  $C_2$ , collegati tra loro da un tubo con rubinetto. I due gas sono a pressioni diverse. L'energia delle molecole dei due contenitori è distribuita secondo due gaussiane,

- $N(\mu_1, \sigma_1)$
- $N(\mu_2, \sigma_2)$

Aperto il rubinetto, la densità di probabilità dell'energia delle molecole sarà una combinazione delle due gaussiane, quindi una mistura gaussiana.

### Esempio: Query by sample

Abbiamo tante immagini in un database e vogliamo cercare quelle che sono simili ad un'immagine di esempio. Quello che si può fare per confrontare queste immagini è usare l'istogramma del colore, indipendente dalle dimensioni delle immagini e che è rappresentabile come somma di un certo numero di gaussiane. Se per esempio rappresento l'istogramma con 4 gaussiane, con 8 numeri  $\mu_i, \sigma_i$  sono in grado di rappresentare immagini magari di  $3000 \times 3000$  punti, su cui fare calcoli di distanza sarebbe un lavoro molto oneroso.

## Variabili casuali notevoli discrete

### Variabile casuale di Bernoulli

La variabile casuale di Bernoulli è di parametro  $p$ .

$$p^x \cdot (1 - p)^{1-x} = p\delta(x - 1) + (1 - p)\delta(x)$$

ed è valida soltanto per  $x = 0, 1$

```
> x=[0,1];
> p = .2; % valore scelto arbitrariamente
> f_x = [(1-p), p];
> stem(x, f_x)
```

### Binomiale di parametri $p$ e $n$

$$f_X(x) = \begin{cases} \binom{n}{x} \cdot p^x (1 - p)^{n-x} & x \\ 0 & \text{altrimenti} \end{cases}$$

$$f_X(x) = \sum_{k=0, n} \binom{n}{x} p^k (1 - p)^{n-k}$$

### Variabile casuale poissoniana

$$f_X(x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} & x \\ 0 & \text{altrimenti} \end{cases}$$

da cui

$$f_X(x) = \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \delta(x - k)$$

Questa variabile casuale è utile per le prove ripetute con  $np = \lambda$  e  $n \rightarrow \infty$

Estratto da "http://it.wikiversity.org/wiki/Variabili\_casuali"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 10:13, 20 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

## Trasformazione di variabili casuali

Da Wikiversità, l'università aperta.

Consideriamo due casi:

- date  $X$ , una variabile casuale, e  $Y = g(X)$ , con  $g(\cdot)$  opportuna, bisognerà caratterizzare la nuova variabile casuale  $Y$ , cioè calcolarne  $F_Y$  e  $f_Y$ .
- date  $X$  e  $Y$  variabili casuali, bisognerà calcolare la  $g(\cdot)$  opportuna in modo tale che  $Y = g(X)$ .

Cominciamo con definire i requisiti della trasformazione  $g(\cdot)$  in modo tale da ottenere ancora una variabile casuale. Si consideri una variabile casuale  $n$ -dimensionale  $X$  definita su uno spazio di probabilità  $\{\Omega, F, P\}$  con

$$X : \Omega \rightarrow \mathbb{R}^n$$

misurabile da  $(\Omega, F)$  a  $(\mathbb{R}^n, \mathbb{B}(\mathbb{R}^n))$ . Sia  $g(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  una funzione misurabile da  $(\mathbb{R}^n, \mathbb{B}(\mathbb{R}^n))$  a  $(\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$ , detta funzione di Borel; definiamo

$$Y = g(X)$$

dove

$$Y : \Omega \rightarrow \mathbb{R}^m$$

tale che

$$Y(s) = g(X(s)) \quad \forall s \in \Omega$$

**Teorema:**

Se  $Y$  è ottenuta come composizione delle funzioni  $X$  e  $g(\cdot)$ , allora  $Y$  è una variabile casuale.

*Dimostrazione:*

$Y$  è una funzione,

$$Y : \Omega \rightarrow \mathbb{R}^m$$

e si ha

$$X : \Omega \rightarrow \mathbb{R}^n$$

Consideriamo l'evento  $B \in \mathbb{B}(\mathbb{R}^m)$ . Allora,

$$Y(s) = g(X(s)) \in B \Leftrightarrow X(s) \in \{\alpha \in \mathbb{R}^n \mid g(\alpha) \in B\} = B'$$

Si ha che  $B' \in \mathbb{B}(\mathbb{R}^n)$ , poiché la  $g(\cdot)$  è una funzione di Borel da  $\mathbb{R}^n$  in  $\mathbb{R}^m$ . Quindi, si arriva alla conclusione

$$\{s \in \Omega \mid Y(s) \in B\} = \{s \in \Omega \mid X(s) \in B'\} \in F$$

**Esempio: Funzioni di Borel**

- $g(\alpha) = \alpha$
- $g(\alpha) = \alpha^2$
- $g(\alpha) = \max(\alpha, 0)$
- $g(\alpha) = \min(\alpha, 0)$
- $g(\alpha) = |\alpha|$
- ...

Noi vedremo trasformazioni di tipo  $1 \rightarrow 1, 2 \rightarrow 2$  e  $2 \rightarrow 1$ .

### Indice

- 1 Trasformazioni di tipo  $1 \rightarrow 1$ 
  - 1.1 Calcolo della densità di probabilità
    - 1.1.1 Con variabili casuali discrete
    - 1.1.2 Con variabili casuali continue
      - 1.1.2.1 monotona crescente
      - 1.1.2.2 monotona decrescente
      - 1.1.2.3 monotona
      - 1.1.2.4 non monotona
- 2 Trasformazioni di tipo  $2 \rightarrow 2$
- 3 Trasformazioni di tipo  $2 \rightarrow 1$ 
  - 3.1 Primo metodo, somma di variabili casuali
  - 3.2 Metodo 2, variabile ausiliaria
- 4 Determinazione di  $g(\cdot)$
- 5 Indipendenza tra variabili casuali

### Trasformazioni di tipo $1 \rightarrow 1$

[[Matematica voce|Definizione|Trasformazioni di tipo  $1 \rightarrow 1$ ]] Sia data una variabile casuale  $X$  con funzione di distribuzione  $F_X(x)$ . Si consideri la trasformazione  $Y = g(X)$  con

$$g(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$$

una funzione misurabile. Allora, si ha

$$F_Y = P(Y \leq y) = P(g(X) \leq y)$$

con  $y \in \mathbb{R}$ .

*Esempio: La trasformazione lineare*

Sia

$$Y = aX + b$$

con  $a > 0$ . Allora,

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(aX + b \leq y) = P(aX \leq y - b) = P\left(X \leq \frac{y - b}{a}\right) = P(X \leq x)$$

Allora, si è ottenuto che

$$F_Y(y) = F_X\left(\frac{y - b}{a}\right)$$

Se  $L_X$  è assolutamente continua, ossia ammette derivata

$$f_X(x) = \frac{dF_X(x)}{dx}$$

allora si ha

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X\left(\frac{y-b}{a}\right)}{dy} = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

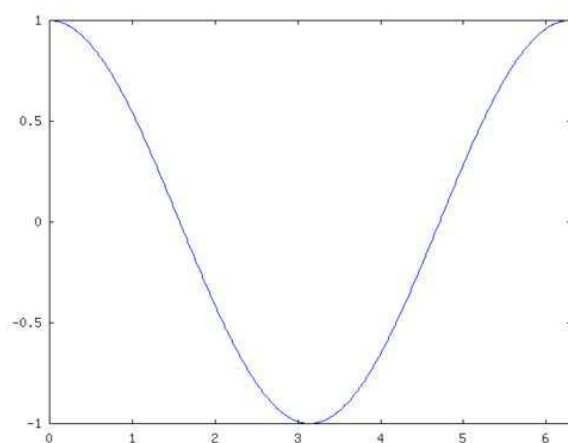
dove si è preso

$$\frac{y - b}{a} = x \Rightarrow y = ax + b \Rightarrow dy = a \cdot dx$$

*Esempio: La funzione coseno*

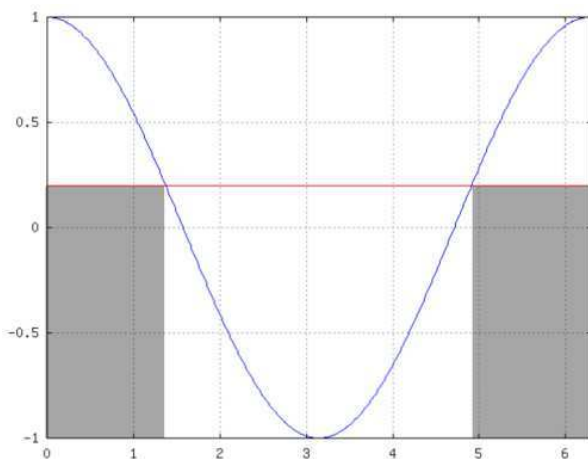
Si ha la funzione

$$y = \cos(x)$$



con distribuzione  $X \sim U(0, 2\pi)$  e con

$$F_Y(y) = \begin{cases} 0 & y < -1 \\ 1 & y \geq 1 \\ ? & -1 \leq y < 1 \end{cases}$$



Fissato un  $\bar{y}$ , abbiamo due soluzioni  $x_1$  e  $x_2$ . Si ha



$$\bar{y} = \cos(x) = \begin{cases} x_1 = \arccos(\bar{y}) & 0 \leq x < \pi \\ x_2 = 2\pi - \arccos(\bar{y}) & \pi \leq x < 2\pi \end{cases}$$

cioè, si possono spezzare i due contributi attorno a  $x = \pi$ , per semplicità. Bisogna calcolare, a questo punto, la probabilità

$$P(Y \leq \bar{y}) = P(-1 < Y \leq \bar{y})$$

Si ottiene

$$F_Y(y) = P(x_1 < X < x_2)$$

Siccome la  $X$  è una variabile casuale uniforme, si ottiene

$$F_Y(y) = P(x_1 < X < x_2) = \frac{x_2 - x_1}{2\pi} = 1 - \frac{\arccos(y)}{\pi}$$

La funzione risultante è continua in  $-1 < y < 1$ , quindi è derivabile e si ottiene

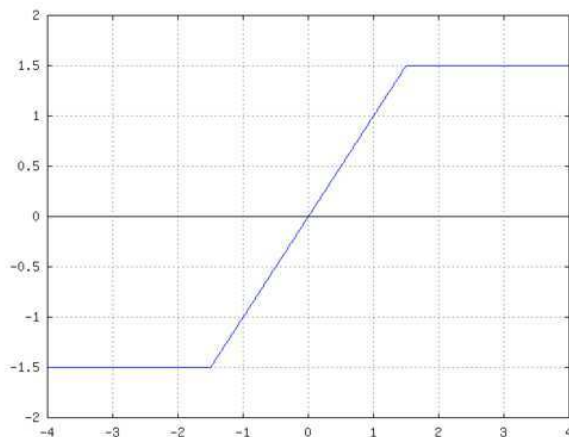
$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{d(1 - \arccos(y))}{\pi dy} = \begin{cases} 0 & y \leq -1, y \geq 1 \\ \frac{1}{\pi\sqrt{1-y^2}} & -1 < y < 1 \end{cases}$$

#### Esercizio: Amplificatore con saturazione

Un amplificatore con saturazione è un oggetto che implementa la funzione

*ciaobepi*

Calcolare la  $f_Y(y)$  e la  $F_Y(y)$  nel caso di  $F_X(x)$  generica lungo tutto l'asse di  $x$ .



Soluzione:

- $y < -aT_X \Rightarrow F_Y(y) = 0$
- $y > aT_X \Rightarrow F_Y(y) = 1$
- $-aT_X \leq y \leq aT_X \Rightarrow F_Y(y) = P(Y \leq y) = P(X \leq x)P\left(x \leq \frac{y}{a}\right) = F_X\left(\frac{y}{a}\right)$

In generale, la  $X$  può avere supporto in  $(-\infty, \infty)$  e potrà essere discontinua.

Immagine: TFA esercizio amplificatore con saturazione F X.png

Si ha

$$f_Y(y) = \frac{dF_X\left(\frac{y}{a}\right)}{dy} + F_X(-T_X)\delta(y + aT_X) + (1 - F_X(T_X))\delta(y - aT_X)$$

Immagine: TFA esercizio amplificatore con saturazione f X.png

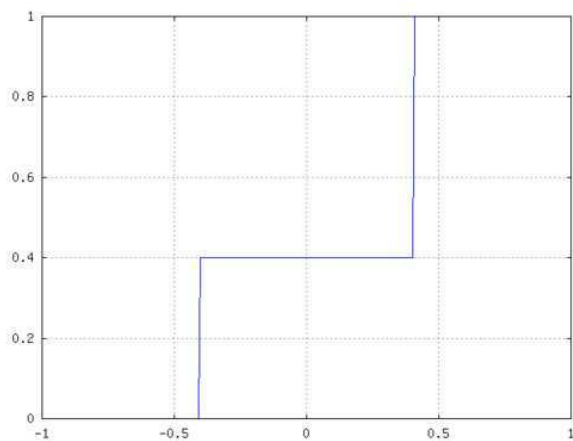
In questo caso, la funzione di densità di probabilità non è né continua né discontinua, ma è mista.

#### Esempio: Il limitatore ideale

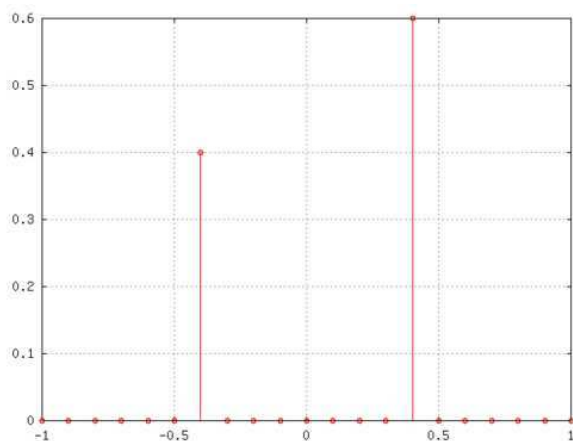
$$g(x) = \begin{cases} T_Y & x \geq 0 \\ -T_Y & x < 0 \end{cases}$$

Immagine: TFA esercizio sul limitatore ideale.png

$$F_Y(y) = \begin{cases} 0 & y < -T_Y \\ 1 & y \geq T_Y \\ P(Y \leq y) = P(X < 0) = F_X(0) & -T_Y \leq y < T_Y \end{cases}$$



La funzione di densità di probabilità discreta è:



$$f_Y(y) = F_X(0)\delta(y + T_Y) - (1 - F_X(0))\delta(y - T_Y)$$

### Calcolo della densità di probabilità

Vediamo un metodo che permette di calcolare la densità di probabilità di una variabile casuale  $Y$  ottenuta dalla trasformazione

$$Y = g(X)$$

con  $X$  variabile casuale e  $g(\cdot)$  funzione misurabile.

#### Con variabili casuali discrete

Sia  $X$  una variabile casuale discreta, con

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

Allora, dato che  $y = g(x)$ , si ha

$$y = \sum_i p_i \delta(y - g(x_i))$$

Può darsi che due  $x_i$  vadano nella stessa  $y_i = g(x_i)$ , quindi si possono anche ottenere meno valori possibili, nella  $Y$ .

#### Con variabili casuali continue

Sia  $X$  una variabile casuale continua con  $f_X(x)$  continua. Sia  $g(X)$  continua e derivabile. Consideriamo tutta la probabilità nell'intervallo  $(a, b)$ :

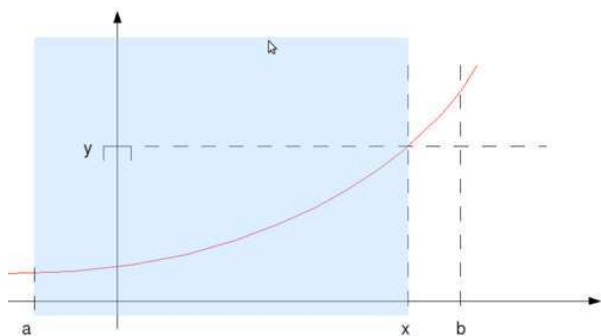
$$P(a < x < b) = 1$$

Identifichiamo vari sottocasi:

##### $g(\cdot)$ monotona crescente

Se  $g(\cdot)$  è monotona crescente, allora ci interessano i valori di probabilità in  $(a, x]$ :

$$F_Y(y) = P(Y \leq y) = P(a < X \leq x) = \int_a^x f_X(\alpha) d\alpha$$



Si definisce la funzione  $x = \psi(y) = g^{-1}(y)$  inversa di  $g(\cdot)$ . Allora,

$$F_Y(y) = \int_a^{\psi(y)} f_X(\alpha) d\alpha.$$

Sappiamo che

$$f_Y(y) = \frac{dF_Y(y)}{dy} = f_X(\psi(y)) \cdot \frac{d\psi(y)}{dy}$$

dove

$$\frac{d\psi(y)}{dy} = \left[ \frac{dg}{dx} \cdot [g^{-1}(y)] \right]^{-1}$$

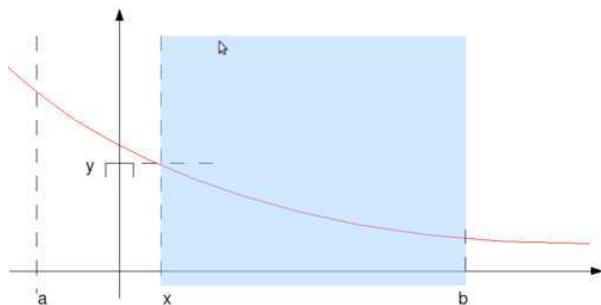
che è la derivata inversa, che coincide con l'inversa della derivata, calcolata in  $g^{-1}(y)$ . Quindi, abbiamo ottenuto

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{\frac{dg}{dx}(g^{-1}(y))}$$

$g(\cdot)$  **monotona decrescente**

Nel caso in cui  $g(\cdot)$  è decrescente, ci interessa il supporto  $[x, b]$ . Si ha

$$F_Y(y) = \int_{g^{-1}(y)}^b f_X(\alpha) d\alpha = \dots \Rightarrow f_Y(y) = -f_X(g^{-1}(y)) \cdot \frac{1}{\frac{dg}{dx}(g^{-1}(y))}$$



$g(\cdot)$  **monotona**

In generale, se  $g(\cdot)$  è monotona, si ha

$$F_Y(y) = f_X(g^{-1}(y)) \cdot \frac{1}{\left| \frac{dg}{dx}(g^{-1}(y)) \right|}$$

Questa è l'espressione generale per le funzioni di tipo monotono. Si ha che  $f_Y = 0$  se

$$Y \notin \text{codominio di } g(\cdot)$$

$g(\cdot)$  **non monotona**

Ipotizziamo che sia possibile partizionare  $(a, b)$  in un numero finito di sottointervalli  $(x_i, x_{i+1})$  in cui  $g(\cdot)$  è monotona (crescente o decrescente che sia). Allora,

$$F_Y(y) = P(Y \leq y) = P(X \in \Delta_1) + P(X \in \Delta_2) + \dots$$

dove

$$P(X \in \Delta_1) = \int_{x_1}^{x_2} f_X(\alpha) d\alpha$$

$$P(X \in \Delta_2) = \int_{x_3}^{x_4} f_X(\alpha) d\alpha$$

Derivando in  $dy$ , si ha

$$\frac{d}{dy} \left( \int_{x_1}^{x_2} f_X(\alpha) d\alpha + \int_{x_3}^{x_4} f_X(\alpha) d\alpha \right)$$

da cui si ottiene

$$\frac{d}{dy} \left( \int_{x_1}^{x_2} f_X(\alpha) d\alpha \right) = f_X(x_2(y)) \cdot \frac{1}{\frac{dg(x_2)}{dx}} - f_X(x_1(y)) \cdot \frac{1}{\frac{dg(x_1)}{dx}}$$

da cui si ha la formula fondamentale (da ricordare)

$$F_Y(y) = \sum_{i=1}^N \frac{f_X(x_i)}{\frac{dg(x_i)}{dx}}$$

Questa è la formula più generale, dove

$$N = |\{g^{-1}(y)\}|$$

cioè il numero di risultati restituiti da  $g^{-1}(y)$ .

*Esempio: La trasformazione lineare*

$$g(x) = ax + b = y \Rightarrow x = g^{-1}(y) = \frac{y-b}{a}$$

Si ha

$$\frac{dg(x)}{dx} = a \Rightarrow F_Y(y) = \frac{f_X(g^{-1}(y))}{\left| \frac{dg(g^{-1}(y))}{dx} \right|} = \frac{f_X(g^{-1}(y))}{|a|} = \frac{f_X\left(\frac{y-b}{a}\right)}{|a|}$$

*Esempio: La trasformazione quadratica*

Nel caso

$$g(x) = ax^2 + b$$

si ottiene

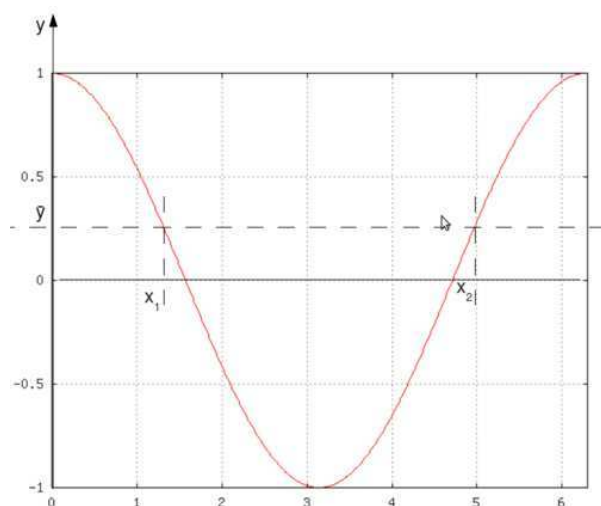
$$F_Y(y) = \frac{f_X\left(\frac{y-b}{a}\right)}{\left| \frac{a}{2} \left(\frac{y-b}{a}\right) \right|}$$

Questo caso si differenzia da quello lineare, perchè la  $x$  compare anche al denominatore di  $F_Y(y)$ , non essendo eliminato dalla derivata.

*Esempio: La trasformazione coseno*

Sia  $X$  uniforme in  $[0, 2\pi]$  (si indica come  $X \sim U[0, 2\pi]$ ).

$$y = \cos(x)$$



Si ha

- $x_1 = \arccos(y) = g_1^{-1}(y)$
- $x_2 = 2\pi - \arccos(y) = g_2^{-1}(y)$

con

$$\frac{dg}{dx} = -\sin(x)$$

da cui

$$F_Y(y) = \sum \frac{f_X(\arccos(y))}{|\cdot|} = \frac{\frac{1}{2\pi}}{|\sin(\arccos(y))|} + \frac{\frac{1}{2\pi}}{|\sin(2\pi - \arccos(y))|} = \frac{1}{\pi} \frac{1}{|\sin(\arccos(y))|} \frac{1}{\pi|\sqrt{1-y^2}|} = \frac{1}{\pi\sqrt{1-y^2}}$$

Questa soluzione è valida per  $-1 < y < 1$ . Si ha

$$f_Y(y) = \begin{cases} 0 & |y| \geq 1 \\ \frac{1}{\pi\sqrt{1-y^2}} & |y| < 1 \end{cases}$$

## Trasformazioni di tipo 2→2

Abbiamo un vettore di variabili casuali in partenza e vogliamo avere un vettore anche per la destinazione,

$$\begin{cases} Y_1 = g_1(X_1, X_2) \\ Y_2 = g_2(X_1, X_2) \end{cases}$$

*Definizione: Funzione di densità di probabilità congiunta*

Si dice funzione di densità probabilità congiunta la funzione

$$f_{X_1, X_2}(x_1, x_2)$$

Siano le  $g_i(\cdot)$  continue e derivabili, per cui esistono le derivate parziali

$$\exists \frac{\partial g_1}{\partial x_i}, \frac{\partial g_2}{\partial x_i}$$

Consideriamo l'intervallo

$$(a_1, b_1) \times (a_2, b_2)$$

che contiene tutta la probabilità

$$P((x_1, x_2) \in (a_1, b_1) \times (a_2, b_2)) = 1$$

Se  $\exists g^{-1}(y)$ , allora  $f_Y(y_1, y_2) = 0$

Supponiamo che esista una partizione di  $(a_1, b_1) \times (a_2, b_2)$  all'interno della quale le  $g_i(\cdot)$  siano monotone nei sottointervalli individuati. Siano le  $(x_{1i}, x_{2i})$ ,  $i = 1, 2, \dots, n$  tali che

$$\{y_1 = g_1(x_{1i}, x_{2i}), y_2 = g_2(x_{1i}, x_{2i}), i = 1, 2, \dots, n\}$$

Con tutte queste premesse, si ha la funzione di densità di probabilità congiunta

$$f_{Y_1, Y_2}(y_1, y_2) = \sum_{i=1}^n \frac{f_{X_1, X_2}(x_{1i}, x_{2i})}{|\det [J_g(x_{1i}, x_{2i})]|}$$

dove  $J_g$  è lo jacobiano

$$J_g(x_1, x_2) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{bmatrix} (x_1, x_2)$$

*Nota:*

Se  $g(\cdot)$  è invertibile, allora si ha

$$\begin{cases} X_1 = \psi_1(Y_1, Y_2) \\ X_2 = \psi_2(Y_1, Y_2) \end{cases}$$

Allora si ha

$$J_g(x_1, x_2) = [J_\psi(Y_1, Y_2)]^{-1}$$

da cui si ha

$$\det(J_g(x_1, x_2)) = \frac{1}{\det(J_\psi(Y_1, Y_2))}$$

*Esercizio:*

Provare a calcolare la  $f_Z, W(z, w)$  con

$$\begin{cases} Z = X + Y \\ W = X - Y \end{cases}$$

e con la funzione di densità di probabilità uniforme su  $[0, 1] \times [0, 1]$

Non si sfrutti il fatto che le due variabili casuali sono indipendenti.

*Esercizio:*

Come il precedente, ma con supporto triangolare

$$D = [0, m] \times [0, n]$$

dove  $m + n = 1$ .

## Trasformazioni di tipo 2→1

Siano  $X_1, X_2$  due variabili casuali con densità di probabilità congiunta

$$f_{X_1, X_2}(x_1, x_2)$$

continua e definita su  $\{\Omega, F, P\}$ . Sia

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}$$

una funzione misurabile continua e derivabile. Data la variabile casuale

$$Z = g(X_1, X_2)$$

qual è la sua distribuzione  $f_Z(z)$ ?

### Primo metodo, somma di variabili casuali

Si ha

$$F_Z(z) = P(Z \leq z) = P(g(X_1, X_2) \leq z) = P((X_1, X_2) \in D_Z)$$

dove

$$D_Z = \{(x_1, x_2) \in \mathbb{R}^2 \mid g(x_1, x_2) \leq z\}$$

Si ha

$$F_Z(z) = \int \int_{D_Z} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2$$

Partendo da qui, si può derivare per ottenere la funzione di densità di probabilità  $f_Z(z)$ .

#### Esempio: Somma di variabili casuali

Sia

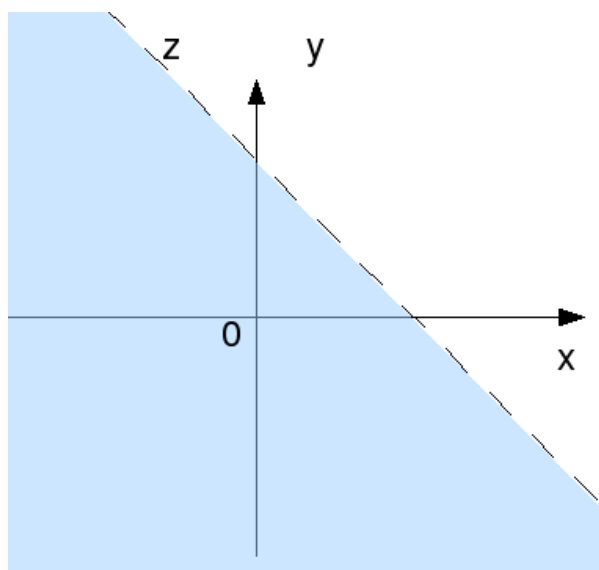
$$Z = X + Y$$

con  $f_{X,Y}(x, y)$  nota e continua. Si ha

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z)$$

Fissato  $z$ , si ha

$$D_Z = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq z\}$$



Si ha

$$F_Z(z) = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{z-x} f_{X,Y}(x, y) dy \right) dx$$

da cui si ha

$$f_Z(z) = \frac{dF_Z(z)}{dz} = \int_{-\infty}^{+\infty} f_{X,Y}(x, z-x) dx = \int_{-\infty}^{+\infty} f_{X,Y}(y-z, y) dy$$

### Metodo 2, variabile ausiliaria

Si aggiunge una variabile casuale ausiliaria in modo da portare la trasformazione in una  $2 \rightarrow 2$ :

$$\begin{cases} y = g(x_1, x_2) \\ z = h(x_1, x_2) \text{ variabile casuale ausiliaria} \end{cases}$$

Si applica a questo punto la tecnica risolutiva per le trasformazioni  $2 \rightarrow 2$  ed otteniamo  $f_{Y,Z}(y, z)$

In alternativa, si determina  $f_Y(y)$  integrando la  $f_{YZ}(y, z)$  rispetto a  $Z$ ,

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{YZ}(y, \alpha) d\alpha$$

*Esempio: Somma di variabili casuali*

Si ha

$$\begin{cases} Z = X + Y \\ W = X \text{ variabile casuale ausiliaria} \end{cases}$$

Si ha

$$|\det(J_g)| = 1$$

quindi posso trovare l'inversa

$$\begin{cases} Y = Z - W \\ X = W \end{cases}$$

Si ha quindi

$$f_{Z,W}(z, w) = \frac{f_{X,Y}(w, z - w)}{1}$$

da cui

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,W}(z, w) dw = \int_{-\infty}^{+\infty} f_{X,Y}(w, z - w) dw = \int_{-\infty}^{+\infty} f_{X,Y}(z - \beta, \beta) d\beta$$

Osservazione: il secondo metodo è più semplice da utilizzare rispetto al primo, a patto che si scelga la variabile ausiliaria migliore possibile.

*Esercizio:*

Sia

$$Z = \sqrt{X^2 + Y^2}$$

qual è la  $W$  migliore?

La soluzione è

$$W = \arctg\left(\frac{X}{Y}\right)$$

## Determinazione di $g()$

Sia data una variabile casuale  $X$  con distribuzione  $F_X(x)$ . Determinare la funzione

$$g : \mathbb{R} \rightarrow \mathbb{R}$$

tale che la

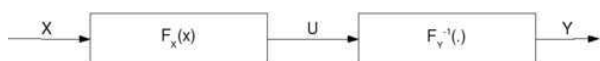
$$Y = g(X)$$

abbia la distribuzione assegnata  $F_Y(y)$ .

La soluzione di questo problema si articola in due passi:

1. passaggio dalla  $F_X(x)$  monotona alla variabile casuale  $U$  uniforme in  $(0, 1)$
2. passaggio da  $U$  alla variabile casuale  $Y$  con  $F_Y(y)$  monotona.

Con uno schema a blocchi:



Si verifica che:

- $U = F_X(x)$  è uniforme.

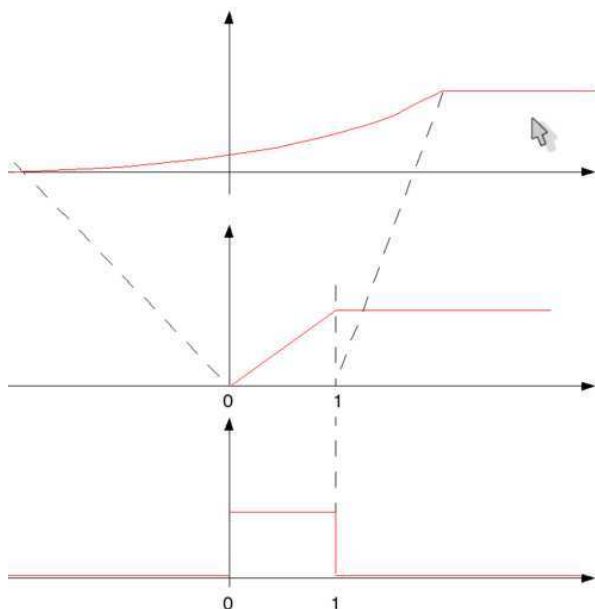
Infatti, si ha

$$F_U(u) = P(U \leq u) = P(F_X(x) \leq u) = P(X \leq x) = U$$

- $Y = F_Y^{-1}(u)$

Infatti, si ha

$$F_Y(y) = P(Y \leq y) = P(F_Y^{-1}(u) \leq y) = P(U \leq F_Y(y)) = F_U(F_Y(y)) = F_Y(y)$$



## Indipendenza tra variabili casuali

**Definizione:** *Variabili statisticamente indipendenti*

Date due variabili casuali  $X$  e  $Y$  definite su  $(\Omega, \mathcal{F}, P)$ , queste si dicono statisticamente indipendenti se:

$$P(x \in B_1, y \in B_2) = P(x \in B_1) \cdot P(y \in B_2) \quad \forall B_1, B_2 \in \mathbb{B}(\mathbb{R})$$

**Teorema:**

$$F_{XY}(x, y) = F_X(x)F_Y(y) =$$

*Dimostrazione:*

$$F_{XY}(\alpha_1, \alpha_2) = P(x \in (-\infty, \alpha_1), y \in (-\infty, \alpha_2)) = P(x \in (-\infty, \alpha_1)) \cdot P(y \in (-\infty, \alpha_2)) = P(X \leq \alpha_1) \cdot P(Y \leq \alpha_2) = F_X(\alpha_1) \cdot F_Y(\alpha_2)$$

**Teorema:**

Se  $F_{XY}(\cdot)$  ammette densità di probabilità, questa è

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$$

**Teorema:**

Se  $X$  e  $Y$  sono indipendenti e

$$g, h : \mathbb{R} \rightarrow \mathbb{R}$$

sono funzioni di Borel, allora

- $Z = g(X)$
- $W = h(Y)$

sono indipendenti.

Osservazione: abbiamo fatto trasformazioni  $\mathbb{1} \rightarrow \mathbb{1}$ , in cui si mantiene l'indipendenza.

**Teorema:**

Dati

- $R_Z = \{x \in \mathbb{R} | g(x) \leq z\}$
- $R_W = \{y \in \mathbb{R} | h(y) \leq w\}$
- $R_Z, R_W \in \mathbb{B}(\mathbb{R})$

con  $g$  e  $h$  funzioni misurabili, allora

$$F_{Z,W}(z, w) = P(Z \leq z, W \leq w)$$

*Dimostrazione:*

Si ha

$$F_{Z,W}(z, w) = P(x \in R_Z, y \in R_W) = P(x \in R_Z) \cdot P(y \in R_W) = P(Z \leq z) \cdot P(W \leq w) = F_Z(z) \cdot F_W(w)$$

**Teorema:**

Dati  $X$  e  $Y$  variabili casuali indipendenti, allora la variabile casuale  $Z = X + Y$  ha una funzione di densità di probabilità

$$f_Z(z) = f_X(x) \cdot f_Y(y)$$

*Dimostrazione:*



$$f_Z(z) = \int_{-\infty}^{+\infty} f(\alpha, z - \alpha) d\alpha = \int_{-\infty}^{+\infty} f(\alpha) \cdot f(z - \alpha) d\alpha$$

*Questa è la definizione di convoluzione. La separabilità delle due funzioni di densità di probabilità è data dal fatto che le variabili casuali sono indipendenti.*

Estratto da "[http://it.wikiversity.org/wiki/Trasformazione\\_di\\_variabili\\_casuali](http://it.wikiversity.org/wiki/Trasformazione_di_variabili_casuali)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 13:59, 22 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Caratterizzazione sintetica delle variabili casuali

Da Wikiversità, l'università aperta.

In molti casi pratici interessa conoscere alcune caratteristiche sintetiche, dette momenti, di una variabile casuale. Queste caratteristiche possono essere calcolate dalla funzione di densità di probabilità.

## Indice

- 1 Valor medio delle variabili casuali
- 2 Varianza delle variabili casuali
- 3 Momenti delle variabili casuali
  - 3.1 Proprietà
    - 3.1.1 Somma di variabili casuali
    - 3.1.2 Prodotto di variabili casuali
- 4 Variabili casuali incorrelate

## Valor medio delle variabili casuali

*Definizione: Valor medio*

Sia  $X$  una variabile casuale con  $f_X(x)$ . Si definisce il valor medio (o valore atteso) di  $X$  la quantità

$$\mu_X = E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx$$

Nel caso di variabili casuali discrete, invece, è definito come

$$\mu_X = E[X] = \sum_i x_i p_i$$

visto che si ha

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

*Esempio: Funzione indicatrice*

Siano  $\{\Omega, \mathcal{F}, P\}$ , con  $A \in \mathcal{F}$  e  $X = I_A$ . Si ha

$$E[X] = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = 0 \cdot (1 - P(A)) + 1 \cdot P(A) = P(A)$$

Quindi, il valor medio della funzione indicatrice è la probabilità di successo.

Si noti che il valor medio della funzione può non essere un risultato possibile per l'esperimento.

*Esempio: Variabile casuale di Bernoulli*

Si ha

$$f_X(x) = p \cdot \delta(x - 1) + (1 - p) \cdot \delta(x)$$

da cui

$$E[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

*Esempio: Variabile casuale di Poisson*

Si ha

$$f_X(x) = \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta(x - k) \quad \lambda > 0$$

da cui si ha

$$E[X] = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!}$$

Siccome vale la proprietà

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot \frac{de^\lambda}{d\lambda} = e^\lambda = \sum_{k=0}^{\infty} k \frac{\lambda^{k-1}}{k!}$$

allora si ottiene

$$E[X] = \left( \sum_{k=0}^{\infty} k \frac{\lambda^{k-1}}{k} e^{-\lambda} \lambda \right) = e^\lambda e^{-\lambda} \lambda = \lambda$$

*Esempio: Variabile casuale uniforme*

Si ha

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b) \\ 0 & \text{altrove} \end{cases}$$

Allora, il valore atteso sarà

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^2}{2} \right]_a^b = \frac{(b^2 - a^2)}{b-a} \cdot \frac{1}{2} = \frac{(b-a)(b+a)}{(b-a)(2)}$$

*Esercizio per casa: La variabile casuale gaussiana*

Trovare il valor medio  $E[X]$  di una variabile casuale gaussiana,

$$X \sim N(\mu_X, \sigma_X)$$

*Teorema: Teorema fondamentale del valore atteso*

Data la variabile casuale  $\underline{X}$   $n$ -dimensionale con funzione di densità di probabilità  $f_{\underline{X}}(\cdot)$  e la trasformazione

$$g : \mathbb{R}^n \rightarrow \mathbb{R} \text{ (una funzione di Borel)}$$

si ha che, per la variabile casuale

$$Y = g(\underline{X})$$

vale

$$E[Y] = \int_{-\infty}^{+\infty} y f_Y(y) dy = \int_{-\infty}^{+\infty} g(x_1, x_2, \dots, x_n) f_{\underline{X}}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$$

Nel caso particolare in cui  $n = 1$  e con una variabile casuale discreta, si ha

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

$$Y = g(X) \text{ ha } f_Y(y) = \sum_i p_i \delta(y - g(x_i)) = \sum_j \left( \sum_{i \mid g(x_i) = y_j} p_i \right) \delta(y - y_j)$$

da cui si ottiene

$$E[Y] = \int_{-\infty}^{+\infty} y f_Y(y) dy = \sum_j y_j \left( \sum_{i | g(x_i)=y_j} p_i \right) = \sum_i g(x_i) p_i = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

#### Proprietà: Proprietà di linearità

Una variabile casuale  $n$ -dimensionale può essere vista come combinazione lineare di  $n$  variabili casuali monodimensionali pesate

$$Y = a_1 g_1(x) + a_2 g_2(x) + \dots + a_n g_n(x)$$

Questo porta alla proprietà

$$E[Y] = \sum_{i=1}^n a_i E[g_i(x)]$$

se poi si impone  $a_i = 1 \forall i$ , allora si ha che  $X = (X_1, X_2, \dots, X_n)$  e vale

$$Y = X_1 + X_2 + \dots + X_n$$

da cui

$$E[Y] = \sum_{i=1}^n E[X_i]$$

## Varianza delle variabili casuali

#### Definizione: Varianza di una variabile casuale

Data una variabile casuale  $X$  con una funzione di densità di probabilità  $f_X(x)$ , si dice varianza la grandezza

$$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f_X(x) dx$$

con

$$\mu_X = E[X]$$

La varianza si può definire anche come

$$\sigma_X^2 = E[(x - \mu_X)^2]$$

Se la variabile casuale è discreta anziché continua, allora si ha

$$\sigma_X^2 = \sum_i p_i (x_i - \mu_X)^2$$

con

$$\mu_X = \sum_i x_i p_i$$

#### Esempio: Variabile casuale binomiale

Si hanno  $n$  prove ripetute indipendenti di una variabile casuale binomiale, con  $\{\Omega, F, P\}$ ,  $F = 2^\Omega$  e con

- $P(\{1\}) = p$
- $P(\{0\}) = (1 - p) = q$

Si definisce lo spazio prodotto

$$\bar{\Omega} = \Omega \times \Omega \times \dots \times \Omega$$

$$\bar{F} = F \otimes F \otimes \dots \otimes F$$

$$\bar{P} = P \times P \times P \times \dots \times P$$

Se  $X$  rappresenta il numero di successi su  $n$  prove, si ha

$$P(X = k) = \binom{n}{k} p^k q^{n-k}$$

da cui

$$E[X] = \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^n k \frac{n!}{(n-k)!k!} p^k q^{n-k} = \sum_{k=1}^n \frac{n!}{(n-k)!(k-1)!} p^k q^{n-k} = \sum_{k=0}^{n-1} \frac{n!}{(n-k)!} p^{k+1} q^{n-k-1}$$

Si ha:

$$\sigma_X^2 = E[(X - \mu_X)^2] = E[(X^2 - 2\mu_X X + \mu_X^2)] = E[X^2] - \mu_X^2 + 2E[X]\mu_X - \mu_X^2 = E[X^2] - \mu_X^2 = E[X^2] - (E[X])^2$$

*Esempio: Variabile casuale di Bernoulli*

Si ha

$$f_X(x) = p\delta(x-1) + q\delta(x)$$

da cui

- $E[X] = p$
- $E[X^2] = 1^2 \cdot p + 0 \cdot q = p$

di conseguenza

$$\sigma_X^2 = E[X^2] - E^2[X] = p - p^2 = p(1-p) = pq$$

*Esempio: Variabile casuale uniforme*

Sia  $X \sim U(a, b)$ . Allora:

- $E[X] = \frac{b+a}{2}$
- $E[X^2] = \int_a^b \frac{1}{b-a} x^2 dx = \frac{1}{b-a} \left[ \frac{x^3}{3} \right]_a^b = \dots = \frac{1}{3}(b^2 + ab + a^2)$

da cui si ottiene

$$\sigma_X^2 = E[X^2] - E^2[X] = \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(a^2 + ab + b^2) = \frac{1}{12}(a^2 + ab + b^2)$$

*Esempio: Variabile casuale gaussiana*

Si ha una variabile casuale con distribuzione

$$N(\mu, \sigma) \Rightarrow f_X(\alpha) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\alpha-\mu)^2}{2\sigma^2}}$$

Allora, vale

$$\sigma_X^2 = E[(X - \mu_X)^2] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (\alpha - \mu)^2 e^{-\frac{(\alpha-\mu)^2}{2\sigma^2}} d\alpha$$

Si impone  $\alpha - \mu = \beta$  e si ha

$$\sigma_X^2 = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \beta^2 e^{-\frac{\beta^2}{2\sigma^2}} d\beta$$

Siccome vale  $\int gh' = gh - \int g'h$ , allora, con

- $g = \beta$
- $h' = \frac{\beta^2}{\sigma^2} e^{-\frac{\beta^2}{2\sigma^2}}$
- $h = e^{-\frac{\beta^2}{2\sigma^2}}$

si ha

$$\sigma_X^2 = -\frac{\sigma}{\sqrt{2\pi}} \left\{ \left[ \beta e^{-\frac{\beta^2}{2\sigma^2}} \right] - \int_{-\infty}^{+\infty} e^{-\frac{\beta^2}{2\sigma^2}} d\beta \right\} = \sigma^2 \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{\beta^2}{2\sigma^2}} d\beta = \sigma^2$$

## Momenti delle variabili casuali

*Definizione: Momenti di una variabile casuale*

Sia  $X$  una variabile casuale con funzione di densità di probabilità  $f_X(x)$ . Si dice momento di ordine  $n$  della variabile casuale  $X$  la quantità

$$m_{n,X} = E[X^n] = \int_{-\infty}^{+\infty} \alpha^n f_X(\alpha) d\alpha$$

*Definizione: Momento centrale di una variabile casuale*

Data una variabile casuale  $X$  con funzione di densità di probabilità  $f_X(x)$ , si dice momento centrale di ordine  $n$  di  $X$  la quantità

$$m_{n,X}^c = E[(X - \mu_X)^n] = \int_{-\infty}^{+\infty} (\alpha - \mu_X)^n f_X(\alpha) d\alpha$$

## Proprietà

Il valore atteso di una funzione di due variabili casuali  $X$  e  $Y$ , con

- densità di probabilità  $f_{X,Y}(x,y)$
- $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  (funzione di Borel)

è dato da

$$E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\alpha, \beta) f_{X,Y}(\alpha, \beta) d\alpha d\beta$$

Inoltre, se vale

$$g(\alpha, \beta) = ag_1(\alpha, \beta) + bg_2(\alpha, \beta)$$

allora

$$E[g(X, Y)] = aE[g_1(\alpha, \beta)] + bE[g_2(\alpha, \beta)]$$

## Somma di variabili casuali

Siano  $X$  e  $Y$  due variabili casuali. Allora il valore atteso è

$$E[X + Y] = E[X] + E[Y] = \mu_X + \mu_Y$$

mentre la varianza è

$$E[((X+Y) - (\mu_X + \mu_Y))^2] = E[(X - \mu_X)^2] + E[(Y - \mu_Y)^2] + 2E[(X - \mu_X)(Y - \mu_Y)] = \sigma_X^2 + \sigma_Y^2 + 2E[(X - \mu_X)(Y - \mu_Y)]$$

### Prodotto di variabili casuali

Siano  $X$  e  $Y$  due variabili casuali. Allora il valore atteso del loro prodotto è

$$E[XY] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \alpha\beta f_{XY}(\alpha, \beta) d\alpha d\beta \neq E[X] \cdot E[Y]$$

Se poi  $X$  e  $Y$  sono indipendenti, allora si ha

$$E[XY] = E[X] \cdot E[Y]$$

La varianza, invece, è

$$E[(X + Y - (\mu_X + \mu_Y))^2] = \sigma_X^2 + \sigma_Y^2 + 2E[(X - \mu_X)(Y - \mu_Y)]$$

Ancora una volta, nel caso di indipendenza, si ha

$$E[(X + Y - (\mu_X + \mu_Y))^2] = \sigma_X^2 + \sigma_Y^2$$

### Variabili casuali incorrelate

*Definizione: Variabili casuali incorrelate*

Due variabili casuali  $x$  e  $Y$  con densità di probabilità  $f_{XY}$  si dicono incorrelate se

$$E[XY] = E[X] \cdot E[Y]$$

Si ha

$X, Y$  indipendenti  $\Rightarrow$  incorrelate

$X, Y$  incorrelate  $\not\Rightarrow$  indipendenti

Si ricordano le definizioni:

- incorrelazione  $\Rightarrow E[XY] = E[X] \cdot E[Y]$
- indipendenza  $\Rightarrow f_{XY}(\alpha, \beta) = f_X(\alpha) \cdot f_Y(\beta)$

*Esempio:*

Siano  $X$  e  $Y$  due variabili casuali con

$$f_{XY}(\alpha, \beta) = \begin{cases} \frac{1}{4} & (\alpha, \beta) = (1, 1) \text{ oppure } (\alpha, \beta) = (-1, 1) \\ \frac{1}{2} & (\alpha, \beta) = (0, 0) \\ 0 & \text{altrimenti} \end{cases}$$

Allora, si ha

$$E[XY] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \alpha\beta f_{X,Y}(\alpha, \beta) d\alpha d\beta = (-1) \cdot (-1) \cdot \frac{1}{4} + 0 \cdot 0 \cdot \frac{1}{2} + (-1) \cdot (1) \cdot \frac{1}{4} = 0$$

$$f_X(\alpha) = \int_{-\infty}^{+\infty} f_{XY}(\alpha, \beta) d\beta = \dots = \begin{cases} 1/4 & x = -1 \\ 1/2 & x = 0 \\ 1/4 & x = 1 \\ 0 & \text{altrove} \end{cases}$$

Allo stesso modo, si ha

$$f_Y(\beta) = \begin{cases} 1/2 & y = 0 \\ 1/2 & y = 1 \\ 0 & \text{altrove} \end{cases}$$

Si ha

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx = 0$$

$$E[Y] = \int_{-\infty}^{+\infty} y f_Y(y) dy = \frac{1}{2}$$

da cui si deduce che le due variabili casuali sono incorrelate. Non si può dire, però che vi sia indipendenza.

Se due variabili casuali sono congiuntamente gaussiane e sono incorrelate, allora sono anche indipendenti; vale il viceversa.

**Definizione: Variabili casuali ortogonali**

Due variabili casuali  $X$  e  $Y$  con funzione di densità di probabilità congiunta  $f_{XY}$  si dicono ortogonali se

$$E[XY] = 0$$

e si indicano con  $X \perp Y$ .

Si ha:

- $X, Y$  incorrelate  $\not\Rightarrow X \perp Y$
- $X \perp Y \not\Rightarrow X, Y$  incorrelate
- $\begin{cases} X, Y \text{ incorrelate} \\ \mu_X = 0 \text{ e/o } \mu_Y = 0 \end{cases} \Rightarrow X \perp Y$
- $\begin{cases} X \perp Y \\ \mu_X = 0 \text{ e/o } \mu_Y = 0 \end{cases} \Rightarrow X, Y \text{ incorrelate}$
- $X, Y$  indipendenti  $\not\Rightarrow X \perp Y$
- $\begin{cases} X, Y \text{ indipendenti} \\ \mu_X = 0 \text{ e/o } \mu_Y = 0 \end{cases} \Rightarrow X \perp Y$
- $X, Y$  indipendenti  $\Rightarrow (X - \mu_X) \perp (Y - \mu_Y)$

**Definizione: Covarianza di variabili casuali**

Siano  $X$  e  $Y$  due variabili casuali con funzione di densità di probabilità congiunta  $f_{XY}$ . Si definisce la covarianza come

$$C_{X,Y} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dx dy$$

dove si ha

- $\mu_X = E[X]$
- $\mu_Y = E[Y]$

Una definizione alternativa è

$$C_{X,Y} = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X E[Y] + \mu_X \mu_Y - \mu_Y E[X] = E[XY] - E[X] \cdot E[Y]$$

Si ha che due variabili casuali  $X$  e  $Y$  sono incorrelate quando  $C_{X,Y} = 0$ .

**Definizione: Coefficiente di correlazione**

Date  $X$  e  $Y$  due variabili casuali con funzione di densità di probabilità congiunta  $f_{XY}$ , si dice coefficiente di correlazione di  $X$  e  $Y$  la grandezza

$$r = \frac{C_{X,Y}}{\sigma_X \sigma_Y}$$



Matematica voce|Esempio||Sia  $X$  una variabile casuale con funzione di densità di probabilità  $f_X$  e sia

$$Y = aX + b$$

Mostrare che  $r = 1$  se  $a > 0$  e  $r = -1$  se  $a < 0$ .

*Soluzione:*

*Il coefficiente di correlazione è una misura della forza della relazione lineare che intercorre tra  $x$  e  $Y$ . Si ha*

■  $r = 0 \Rightarrow X, Y$  incorrelate

■  $r = 1 \Rightarrow X, Y$  non incorrelate

Si ha

$$C_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

dove

$$\mu_Y = E[Y] = E[aX + b] = a\mu_X + b$$

da cui

$$C_{XY} = E[(X - \mu_X)(aX + b - (a\mu_X + b))] = E[(X - \mu_X)a(X - \mu_X)] = aE[(X - \mu_X)^2] = a \cdot \sigma_X^2$$

Si ha

$$\sigma_Y^2 = E[(Y - \mu_Y)^2] = E[a^2(X - \mu_X)^2] = a^2 \sigma_X^2$$

da cui si può calcolare il coefficiente di correlazione

$$r = \frac{C_{XY}}{\sigma_X \sigma_Y} = \frac{a \sigma_X^2}{\sigma_X |a| \sigma_X} = \frac{a}{|a|} = \begin{cases} 1 & a > 0 \\ -1 & a < 0 \end{cases}$$

Estratto da "[http://it.wikiversity.org/wiki/Caratterizzazione\\_sintetica\\_delle\\_variabili\\_casuali](http://it.wikiversity.org/wiki/Caratterizzazione_sintetica_delle_variabili_casuali)"

Categorie: [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

- Ultima modifica per la pagina: 12:28, 6 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

## Distribuzione e densità condizionata

Da Wikiversità, l'università aperta.

**Definizione: Distribuzione condizionata**

Data una variabile casuale  $X$  definita su  $(\Omega, \mathcal{F}, P)$  con  $A \in \mathcal{F}$  e con  $P(A) \neq 0$  si dice distribuzione di  $X$  condizionata all'evento  $A$  la funzione

$$F_{X|A} : \mathbb{R} \rightarrow [0, 1]$$

definita come

$$F_{X|A} = P(X \leq x | A)$$

dove

$$P(X \leq x | A) = P(\{s \in \Omega \mid X(s) \leq x\} \cap A) = \frac{P(\{s \in \Omega \mid X(s) \leq x\} \cap A)}{P(A)}$$

Nota:  $F_{X|A}$  può essere vista come la funzione di distribuzione definita su  $(\Omega, \mathcal{F}, P)$  con  $P_A(B) = P(B | A)$ .

**Definizione: Funzione di densità di probabilità condizionata**

Sia  $X$  una variabile casuale su  $(\Omega, \mathcal{F}, P)$  con  $A \in \mathcal{F}$  e  $P(A) \neq 0$ . Se esiste la funzione

$$f_{X|A} : \mathbb{R} \rightarrow \mathbb{R}^+$$

tale che

$$F_{X|A}(x|A) = \int_{-\infty}^{+\infty} f_{X|A}(\alpha|A) d\alpha \quad \forall x \in \mathbb{R}$$

allora si dice che  $F_{X|A}$  ammette densità e  $f_{X|A}$  è la funzione di densità di probabilità condizionata di  $X$  dato l'evento  $A$ .

**Teorema: Teorema della probabilità totale**

Sia  $X$  una variabile casuale su  $(\Omega, \mathcal{F}, P)$  con

$$\Omega = \bigcup_{i=1}^n A_i$$

con gli  $A_i \in \mathcal{F}$  disgiunti a coppie. Allora, la probabilità dell'evento

$$\{s \in \Omega \mid X(s) \leq x\} \in \mathcal{F}$$

è data da

$$P(X \leq x) = P(X \leq x | A_1) + P(X \leq x | A_2) + \dots + P(X \leq x | A_n) \text{ ossia } F_X(x) = F_{X|A_1}(x|A_1) \cdot P(A_1) + \dots + F_{X|A_n}(x|A_n) \cdot P(A_n)$$

Se poi la variabile casuale  $X$  ammette densità, allora si ha

$$f_X(x) = f_{X|A_1}(x|A_1) \cdot P(A_1) + f_{X|A_2}(x|A_2) \cdot P(A_2) + \dots + f_{X|A_n}(x|A_n) \cdot P(A_n)$$

**Esempio:**

Sia  $X$  una variabile casuale con funzione di densità di probabilità  $f_X$  definita su  $(a, b)$ . Siano gli  $A_i$  una partizione dell'insieme  $(a, b)$ . Data

$$Y = g(X)$$

si ha

$$f_Y(y) = f_{Y|A_1}(y|A_1) \cdot P(A_1) + f_{Y|A_2}(y|A_2) \cdot P(A_2) + \dots + f_{Y|A_n}(y|A_n) \cdot P(A_n) = \sum_{i=1}^n \frac{f_{X|A_i}(g^{-1}(y|A_i)) \cdot P(A_i)}{\left| \frac{dg}{dx}(g^{-1}(y)) \right|}$$

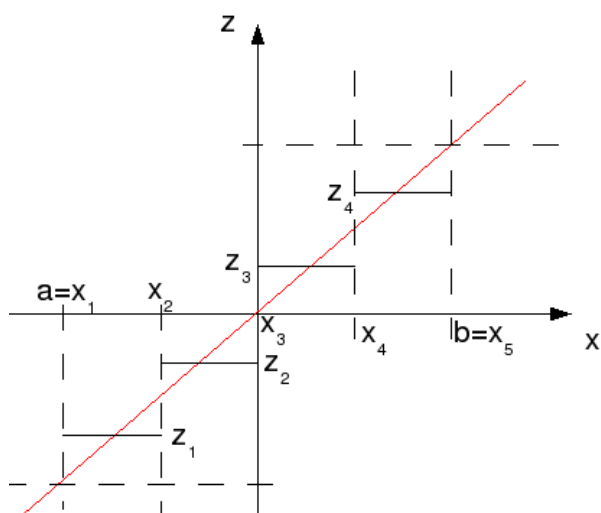
**Esempio: Calcolo della densità di probabilità dell'errore di quantizzazione**

Sia  $X$  una variabile casuale con funzione di densità di probabilità  $f_X$  su  $(a, b)$ . Sia  $E$  la variabile casuale errore, calcolare la sua  $f_E$  con

$$E = X - Z = X - h(X) = g(X)$$

dove

$$g(x) = \begin{cases} x - z_1 & x \in (x_1, x_2] \\ x - z_2 & x \in (x_2, x_3] \\ x - z_3 & x \in (x_3, x_4] \\ x - z_4 & x \in (x_4, x_5] \end{cases}$$



In generale, per determinare la  $f_{X|A}$  si deve conoscere lo spazio di probabilità  $(\Omega, \mathcal{F}, P)$  su cui è definita la variabile casuale  $X$ , dal momento che bisogna calcolare la probabilità

$$P(X \leq x | A)$$

con  $A \in \mathcal{F}$ . Se però  $A$  è espresso in funzione di  $X$ , per esempio

$$A = \{s \in \Omega \mid X(s) \in B\} \text{ con } B \in \mathcal{B}(\mathbb{R})$$

allora la  $f_{X|A}$  può essere calcolata a partire da  $F_X$ .

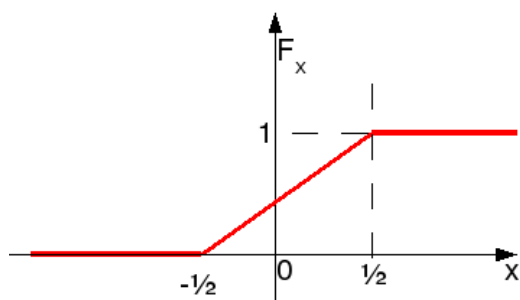
*Esempio:*

Sia  $X$  una variabile casuale con distribuzione di probabilità  $F_X$  e sia

$$A = \{s \in \Omega \mid a < X(s) < b\}$$

Allora, si ha

$$F_{X|A}(x|A) = F_{X|A}(x|a < X \leq b) = P(X \leq x | a < X \leq b) = \frac{P(X \leq x \cap a < X \leq b)}{P(a < X \leq b)}$$



Si ha

$$F_{X|A}(x | a < X \leq b) = \begin{cases} 0 & x < a \\ \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} & x \in [a, b) \\ 1 & x \geq b \end{cases}$$

Se la  $F_X$  ammette anche densità, allora

$$f_{X|a < X \leq b}(x | a < X \leq b) = \frac{dF_{X|A}}{dx}(x|A) = \begin{cases} 0 & x \leq a, x > b \\ \frac{f_X(x)}{F_X(b) - F_X(a)} & x \in (a, b) \end{cases}$$

*Esercizio:*

Sia  $X$  una variabile casuale con  $F_X$  e  $f_X$ . Calcolare la  $F_{X|X>a}$  e la  $f_{X|X>a}$ .

## Indice

- 1 Valore atteso condizionato ad un evento
- 2 Densità condizionata di variabile casuale, data un'altra variabile casuale
- 3 Valore atteso condizionato
- 4 Stima con la massima probabilità a posteriori (MPP) e con la massima verosimiglianza (MV)
  - 4.1 Massima probabilità a posteriori
  - 4.2 Massima verosimiglianza

### Valore atteso condizionato ad un evento

Siano  $(\Omega, \mathcal{F}, P)$  con  $A \in \mathcal{F}$  e con  $P(A) \neq 0$ . Data la variabile casuale  $X$  definita su  $(\Omega, \mathcal{F}, P)$  che ammette densità condizionata  $f_{X|A}$ , allora

$$E[X|A] = \int_{-\infty}^{+\infty} x f_{X|A}(x|A) dx$$

Se esiste una funzione  $g(\cdot)$  per cui  $Y = g(X)$ , con  $g : \mathbb{R} \rightarrow \mathbb{R}$  funzione di Borel, allora si ha

$$E[Y|A] = \int_{-\infty}^{+\infty} g(x) f_{X|A}(x|A) dx$$

*Teorema: Teorema di Bayes*

Data una variabile casuale  $X$  definita su uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , con  $A \in \mathcal{F}$  e con  $x \in \mathbb{R} \mid P(X \leq x) \neq 0$ . Allora, si ha

$$P(A|X \leq x) = \frac{P(X \leq x|A)P(A)}{P(X \leq x)}$$

*Esempio:*

Sia  $X$  una variabile casuale con distribuzione  $F_X$  definita su  $(\Omega, \mathcal{F}, P)$ , con  $A \in \mathcal{F}$ . Si ha

$$P(A|x_1 < X \leq x_2) = \frac{P(x_1 < X \leq x_2|A)P(A)}{P(x_1 < X \leq x_2)} = \frac{[F_{X|A}(x_2|A) - F_{X|A}(x_1|A)] \cdot P(A)}{F_X(x_2) - F_X(x_1)}$$

Vediamo ora come è possibile calcolare la probabilità di un evento, condizionata ad una variabile casuale. Se si ha  $X$  come variabile casuale discreta, si ha

$$P(A|X = x) = \frac{P(X = x|A) \cdot P(A)}{P(X = x)}$$

con  $P(X=x) \neq 0$ . Se  $X$  è una variabile casuale continua, al contrario, si ha che  $P(X = x) = 0 \forall x \in \mathbb{R}$ ; osserviamo che vale

$$\int_B f_{X|A}(x|A) P(A) dx = P(x \in B, A) \forall B \in \mathbb{B}(\mathbb{R})$$

La probabilità di un evento  $A$  condizionata a  $X=x$  è definita come quella funzione

$$g_A : \mathbb{R} \rightarrow \mathbb{R}^+$$

tale per cui

$$\int_B g_A(x) f_X(x) = P(x \in B, A) \forall B \in \mathbb{B}(\mathbb{R})$$

La grandezza  $P(A|X=x)$  può essere calcolata da

$$P(A|X = x) = \begin{cases} \frac{f_{X|A}P(A)}{f_X(x)} & f_X(x) \neq 0 \\ 0 & \text{altrimenti} \end{cases}$$



**LAVORI IN CORSO! - WORK IN PROGRESS!**

**Davide Morellato**  si sta occupando di questo testo; non apportare modifiche se l'ultima modifica ([http://it.wikiversity.org/w/index.php?title=Distribuzione\\_e\\_densit%C3%A0\\_condizionata&action=history](http://it.wikiversity.org/w/index.php?title=Distribuzione_e_densit%C3%A0_condizionata&action=history)) è recente.

*Teorema: Teorema della probabilità totale*

Data una variabile casuale  $X$  definita su  $(\Omega, \mathcal{F}, P)$ , con  $A \in \mathcal{F}$  e con funzione di densità di probabilità  $f_X(x)$ , allora si ha che

$$P(A) = \int_{-\infty}^{+\infty} P(A|X=x) f_X(x) dx$$

*Dimostrazione:*

Dalla definizione precedente, si ha

$$\int_B P(A|X=x) f_X(x) dx = P(x \in B, A)$$

Se l'evento  $B$  è tutto  $\Omega$ , allora si ha

$$P(\Omega, A) = P(\Omega \cap A) = P(A)$$

**Teorema: Teorema di Bayes per funzioni di densità di probabilità**

Si ha lo spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , con  $\mathcal{A} \in \mathcal{F}$  come  $\sigma$ -algebra, e con  $P(A) \neq 0$ . Si ha una variabile casuale  $X$  con funzione di densità di probabilità  $f_X$ . Allora, vale

$$f_{X|A}(x|A) = \frac{P(A|X=x) f_X(x)}{\int_{-\infty}^{+\infty} P(A|X=x) f_X(x) dx} = \frac{P(A|X=x) f_X(x)}{P(A)}$$

*Dimostrazione:*

E' lasciata allo studente...

### Densità condizionata di variabile casuale, data un'altra variabile casuale

Sappiamo calcolare la probabilità

$$P(Y \in B | X = x)$$

come un caso particolare di  $P(A | X = x)$ . In altri termini, si ha

$$P(Y \in B | X = x) = \begin{cases} \frac{f_{X|Y \in B}(x|Y \in B) P(Y \in B)}{f_X(x)} & \forall x | f_X(x) \neq 0 \\ 0 & \text{altrimenti} \end{cases}$$

**Definizione: Distribuzione di  $Y$  condizionata a  $X = x$**

Data la variabile casuale  $X$  con funzione di densità  $f_X$ , si ha

$$F_{Y, X=x}(y | X = x) = P(Y \leq y | X = x)$$

Questa è la funzione di distribuzione (se  $f_X(x) \neq 0$ ) perché la funzione

$$m_X: \mathcal{F} \rightarrow [0, 1]$$

definita da

$$m_X(A) = P(A | X = x)$$

è una misura di probabilità se esiste la funzione di densità di probabilità  $f_{Y|X}(y | x)$  tale per cui

$$F_{Y|X=x}(y | X = x) = \int_{-\infty}^{+\infty} P_{Y|X}(B|x) dB$$

da cui si ha che la  $f_{Y|X}$  è una densità.

**Proposizione:**

$$f_{Y|X} = \begin{cases} \frac{f_{XY}(x,y)}{f_X(x)} & x | f_X(x) \neq 0 \\ 0 & \text{altrimenti} \end{cases}$$

Questo risultato assomiglia alla probabilità condizionata di un evento, che è

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A, B)}{P(B)}$$

cioè, la densità di probabilità congiunta non è altro che l'intersezione delle due variabili casuali di partenza. } }

**Teorema:**

Si ha

$$P(y \in C | X = x) = \int_C f_{Y|X}(y, x) dy \quad \forall C \in \mathbb{B}(\mathbb{R})$$

con  $f_{Y|X}$  dato dalla proposizione precedente.

*Dimostrazione:*

Per definizione,

$$\int_B P(Y \in C | X = x) f_X(x) dx = P(Y \in C, X \in B) \quad \forall B \in \mathbb{B}(\mathbb{R})$$

Si ha

$$\int_B \left( \int_C f_{Y|X}(y|x) dy \right) f_X(x) dx = \int_B \int_C f_{Y|X} f_X(x) dy dx = \int_B \int_C f_{X,Y}(x, y) dx dy = P(X \in B, Y \in C) \quad \forall I$$

Osservazioni:

- $f_{Y|X}(y|\bar{x})$  è la sezione di  $f_{X,Y}(x,y)$  con il piano  $x = \bar{x}$ ;
- se  $X$  e  $Y$  sono variabili casuali indipendenti, allora si ha

$$f_{Y|X}(y|x) = f_Y(y)$$

Infatti, se le due variabili sono indipendenti, la conoscenza di  $Y$  non è in alcun modo influenzata dalla conoscenza di  $X$ .

## Valore atteso condizionato

*Definizione: Valore atteso condizionato*

Date due variabili casuali  $X$  e  $Y$ , con funzione di densità di probabilità  $f_{X,Y}(x,y)$ , si definisce il valore atteso di  $Y$  condizionato da  $X = x$

$$E[Y|X = x] = \int_{-\infty}^{+\infty} y f_{Y|X}(y|x) dy$$

Anche qui, vale il teorema del valore atteso che dice che, data una variabile casuale  $Z = g(Y)$ , si ha

$$E[Z|X = x] = E[g(Y)|X = x] = \int_{-\infty}^{+\infty} g(y) \cdot f_{Y|X}(y|x) dy$$

*Proprietà: Proprietà del valore atteso condizionato:*

- Se le variabili casuali  $X$  e  $Y$  sono indipendenti, allora  $f_{Y|X} = f_Y$ , da cui si ha

$$E[Y|X = x] = \int y f_{Y|X}(y|x) dy = \int y f_Y(y) dy = E[Y]$$

- Se  $Y = g(X)$ , allora si ha dipendenza totale e vale

$$E[Y|X = x] = E[g(X)|X = x] = \int_{-\infty}^{+\infty} y \delta(y - g(x)) dy = g(x)$$

Questo è abbastanza intuitivo: se si ha dipendenza completa  $Y = g(X)$ , allora vale

$$f_{Y|X}(y|x) = \delta(y - g(x))$$

cioè, si ha una  $\delta$  quando  $y = g(x)$ , non si ha alcuna incertezza sul risultato.

- $E[g(X,Y) | X = x] = E[g(x,Y) | X = x]$

Se poi  $g(x,Y)$  è separabile, cioè

$$g(x, Y) = g_1(x) \cdot g_2(Y)$$

allora si ha

$$E[g_1(x) \cdot g_2(Y) | X = x] = g_1(x) \cdot E[g_2(Y) | X = x]$$

Consideriamo il solito spazio di probabilità  $\{\Omega, \mathcal{F}, P\}$  con le variabili casuali  $X, Y$   $f_{X,Y}$  e con un'altra variabile casuale  $Z : \Omega \rightarrow \mathbb{R}$  tale che

$$Z(s) = E[Y|X = X(s)] \quad \forall s \in \Omega$$

La variabile casuale  $Z$  ha un valore atteso condizionato; in generale, la variabile casuale  $Z$  rappresenta il valore atteso condizionato

$$E[Y|X]$$

*Teorema:*

*Si ha*

$$E[Z] = E[E[Y|X]] = E[Y]$$

*Dimostrazione:*

*Si ha*

$$E[Z] = \int_{-\infty}^{+\infty} E[Y|X = x] f_X(x) dx$$

*Applicando il teorema del valore atteso, si ha*

$$E[Z] = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{+\infty} y \cdot f_{Y|X}(y|x) dy \right) \cdot f_X(x) dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y \cdot f_{Y|X}(y|x) dx dy = \int_{-\infty}^{+\infty} y \cdot f_Y(y) dy = E[Y]$$

## Stima con la massima probabilità a posteriori (MPP) e con la massima verosimiglianza (MV)

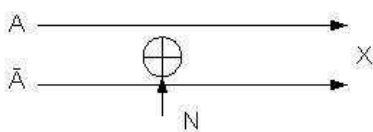
Dato lo spazio di probabilità  $\{\Omega, F, P\}$  con  $A, \bar{A} \in F$  e  $A \cap \bar{A} = \emptyset$   $A \cup \bar{A} = \Omega$ , si definisce la variabile casuale

$$X f_X(x) = f_{X|A}(x|A) \cdot P(A) + f_{X|\bar{A}}(x|\bar{A}) \cdot P(\bar{A})$$

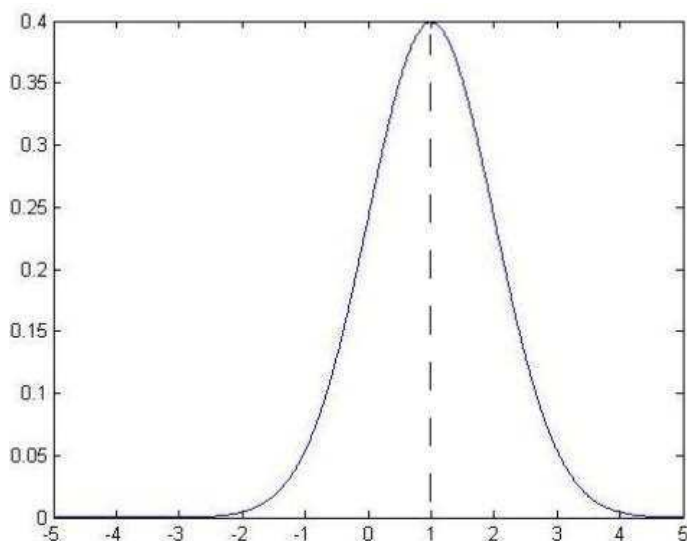
Quello che vogliamo fare è, dato  $\bar{x}$ , capire a quale delle due popolazioni  $f_{X|A}$  oppure  $f_{X|\bar{A}}$  appartiene. Se si prende l'esempio delle molecole di gas, la funzione di densità di probabilità dell'energia è una combinazione lineare di altre funzioni di densità di probabilità; qui è la stessa cosa, si vuole calcolare la densità di uscita come *mixture* di altre densità.

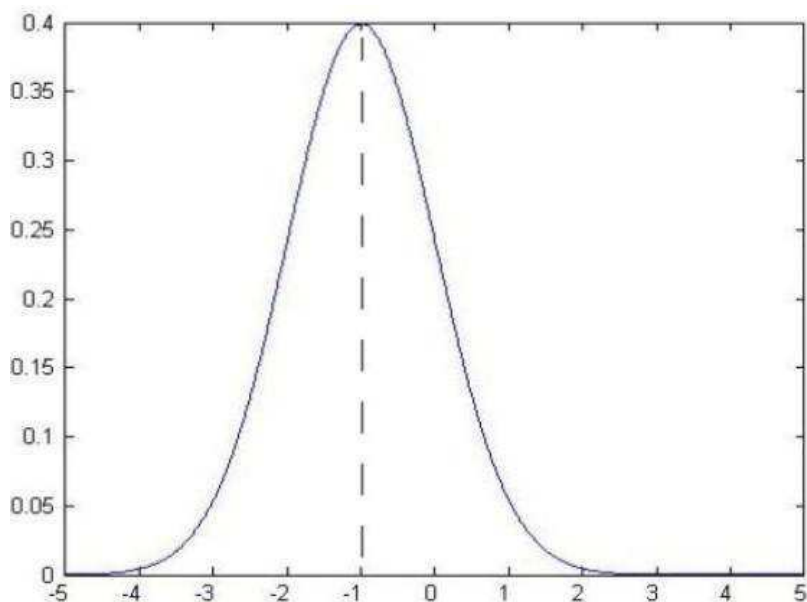
*Esempio: Canale discreto gaussiano*

*Si ha un canale discreto con rumore gaussiano.*



*I simboli di partenza possono essere  $A$  oppure  $\bar{A}$ , il rumore è di tipo gaussiano  $N(\mu, \sigma)$  e la variabile casuale di interesse è la  $X$  di uscita. Se si trasmette il segnale  $A$ , la funzione di densità di probabilità sarà la  $f_{X|A}$ ; al contrario, se si trasmette  $\bar{A}$ , allora la densità sarà  $f_{X|\bar{A}}$ .*





Si ha

- $A \Rightarrow f_{X|A}(x|A) = N(\mu + A, \sigma)$
- $\bar{A} \Rightarrow f_{X|\bar{A}}(x|\bar{A}) = N(\mu + \bar{A}, \sigma)$

Da qui, si ha che la  $f_X$  sarà una combinazione lineare di  $f_{X|A}$  e  $f_{X|\bar{A}}$ . In questo esempio, la somma sarà pesata con le probabilità dei due simboli,  $P(A)$  e  $P(\bar{A})$ . Lo scopo finale di questo lavoro è osservare il risultato della variabile casuale  $X$  per capire se è stato trasmesso  $A$  o  $\bar{A}$ .

*Esempio:*

Data un'altezza umana, si tratta dell'altezza di un uomo o di una donna?

### Massima probabilità a posteriori

La massima probabilità a posteriori è definita come

$$P(A|X = x) = \begin{cases} > P(\bar{A}|X = x) & A \\ < P(\bar{A}|X = x) & \bar{A} \end{cases}$$

Se vale  $P(A|X = x) > P(\bar{A}|X = x)$ , allora decidiamo che è stato trasmesso  $A$  e non  $\bar{A}$ .

Il criterio a massima probabilità a posteriori, quindi, è

$$f_{X|A}(x|A) \cdot P(A) \begin{cases} < \\ > \end{cases} f_{X|\bar{A}}(x|\bar{A}) \cdot P(\bar{A})$$

### Massima verosimiglianza

Nel caso in cui  $P(A) = P(\bar{A})$ , il criterio di massima probabilità a posteriori diventa un criterio di massima verosimiglianza.

$$f_{X|A}(x|A) \cdot P(A) \begin{cases} < \\ > \end{cases} f_{X|\bar{A}}(x|\bar{A}) \cdot P(\bar{A}) = f_{X|A}(x|A) \begin{cases} < \\ > \end{cases} f_{X|\bar{A}}(x|\bar{A})$$

Il criterio a massima verosimiglianza è meno potente di quello a massima probabilità a posteriori, perché quest'ultimo sfrutta la conoscenza della probabilità dei simboli prima di essere trasmessi, mentre il secondo metodo si limita ad osservare il risultato finale e la funzione di densità di probabilità, considerando tutti i simboli equiprobabili. Quindi, il criterio a massima verosimiglianza è meno potente, perché sfrutta meno informazioni.

*Esercizio al calcolatore:*

Stimare la MPP e la MV del sistema dell'esempio, con sorgente  $A \sim U[0, 1]$  e rumore  $N(\mu = 0, \sigma)$





**LAVORI IN CORSO! - WORK IN PROGRESS!**

**Daide Morellato** si sta occupando di questo testo; non apportare modifiche se l'ultima modifica ([http://it.wikiversity.org/w/index.php?title=Distribuzione\\_e\\_densit%C3%A0\\_condizionata&action=history](http://it.wikiversity.org/w/index.php?title=Distribuzione_e_densit%C3%A0_condizionata&action=history)) è recente.

Estratto da "[http://it.wikiversity.org/wiki/Distribuzione\\_e\\_densit%C3%A0\\_condizionata](http://it.wikiversity.org/wiki/Distribuzione_e_densit%C3%A0_condizionata)"

Categorie: [WIP](#) | [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

---

- Ultima modifica per la pagina: 13:57, 12 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Processi stocastici, stazionarietà ed ergodicità e densità spettrale di potenza

Da Wikiversità, l'università aperta.

Un processo casuale potrebbe essere, per esempio, quando si acquisisce un segnale con un oscilloscopio collegato ad un generatore di forme d'onda; la sinusoide generata sarà  $\sin(2\pi ft + \Phi)$ , dove  $\Phi$  è una variabile casuale.

## Indice

- 1 Processi casuali
- 2 Caratterizzazione statistica del processo
  - 2.1 Realizzazioni
  - 2.2 Variabili casuali
  - 2.3 Densità di probabilità
- 3 Caratterizzazione statistica di un processo
- 4 Ergodicità dei processi casuali

## Processi casuali

Un processo casuale  $\{X(t), t \in T\}$  è una collezione di variabili casuali indicizzate dal tempo  $t \in T$ . Quindi,  $X(\bar{t}) \forall \bar{t} \in T$  è una variabile casuale definita nello spazio di probabilità  $\{\Omega, \mathcal{F}, P\}$ . Fissato l'istante temporale  $\bar{t} \in T$ , si ha una funzione misurabile

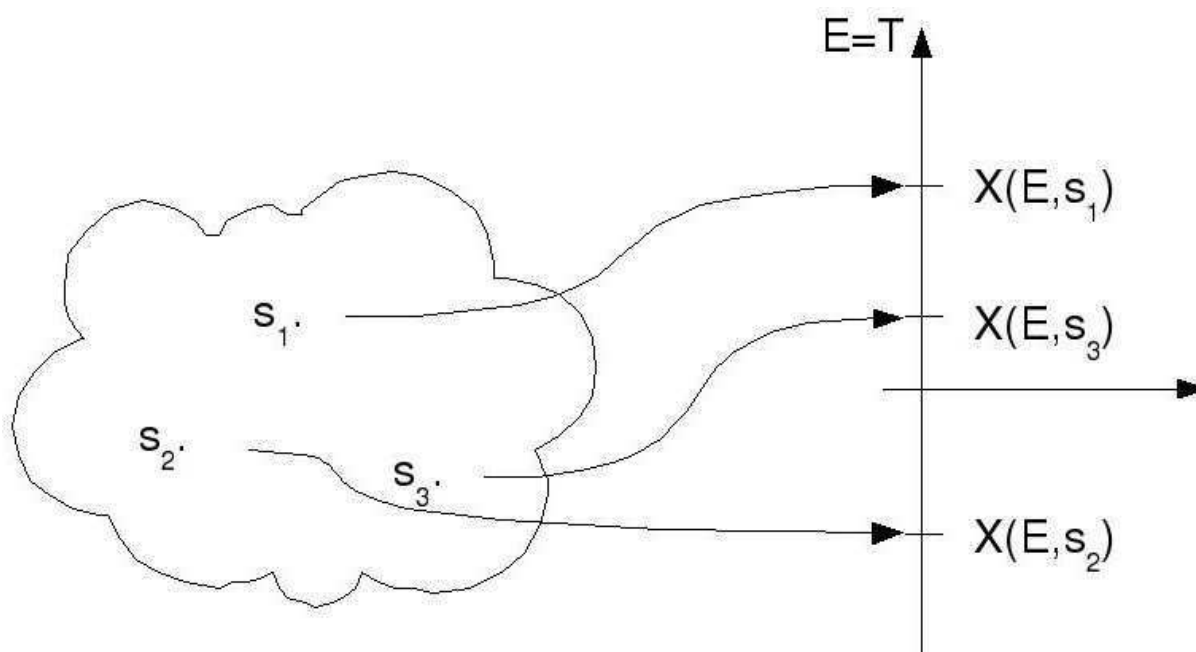
$$X(\bar{t}) : \Omega \rightarrow \mathbb{R}$$

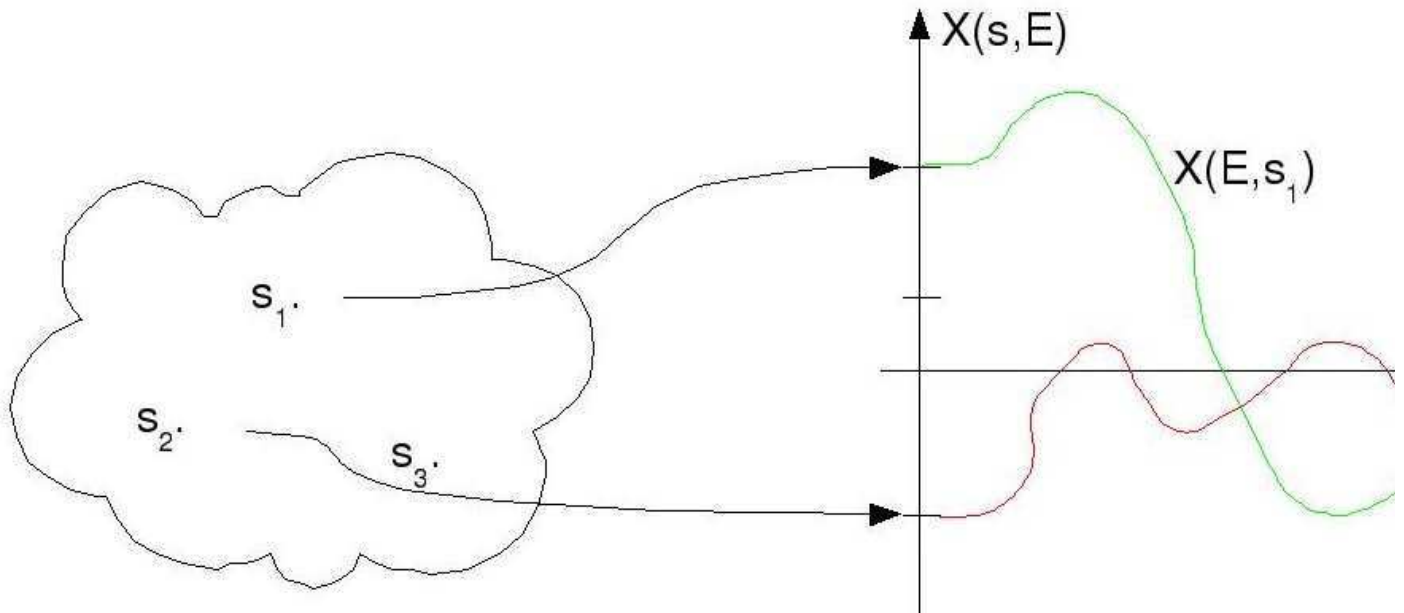
che parte dallo spazio  $(\Omega, \mathcal{F})$  per giungere allo spazio  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . E' possibile indicare un processo stocastico anche con la notazione

$$X(t, s) : \Omega \times T \rightarrow \mathbb{R}$$

dove  $X(t, s)$  è misurabile rispetto ad  $s$ .

Se  $T = \mathbb{R}$ , il processo si dice a tempo continuo; al contrario, se  $T = \mathbb{Z}$ , allora il processo si dice a tempo discreto.



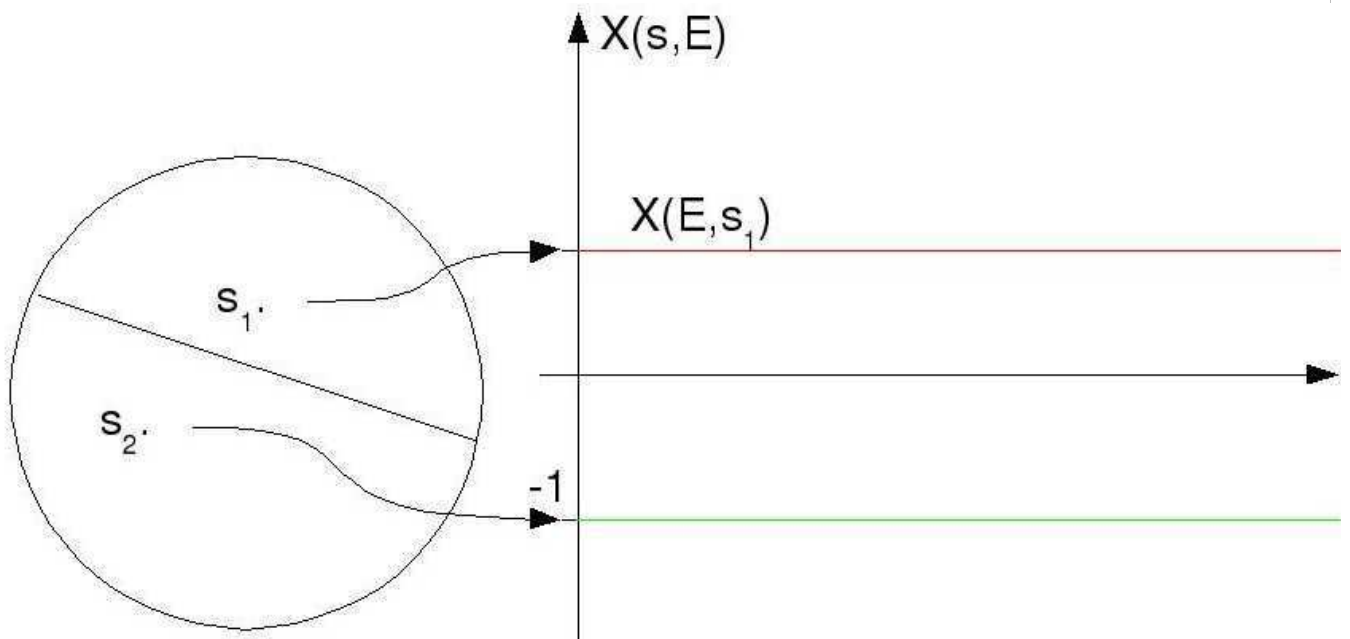


Se fissiamo  $s \in \Omega$ , allora  $X(s)$  è una funzione del tempo, altrimenti detta realizzazione.

*Esempio:*

Sia  $X(t) = V$ , con  $V$  una variabile casuale con

$$f_V(v) = \frac{1}{2}\delta(v - 1) + \frac{1}{2}\delta(v - (-1))$$



Osservazione: fissato un qualunque  $\bar{t} \in T$ , la variabile casuale resta sempre la stessa

$$f_V(\bar{t}) = \frac{1}{2}\delta(\bar{t} - 1) + \frac{1}{2}\delta(\bar{t} - (-1)) \quad \forall \bar{t} \in T$$

*Esercizio:*

Sia  $X(t) = V$  una variabile casuale con  $V \sim U[0,1]$ , disegnare il grafico come per l'esempio precedente.

*Esempio:*

Sia  $X(t)$  la variabile casuale

$$X(t) = V \cdot \sin(\omega t)$$

tale che

- $\omega$  è costante;
- $f_v(v) = \frac{2}{3}\delta(v-1) + \frac{1}{3}\delta(v+1)$

Immagine: TFA processo a tempo continuo esempio sinusoide.jpg

### Caratterizzazione statistica del processo

*Esempio:*

Sia  $X(t,s)$  definito su

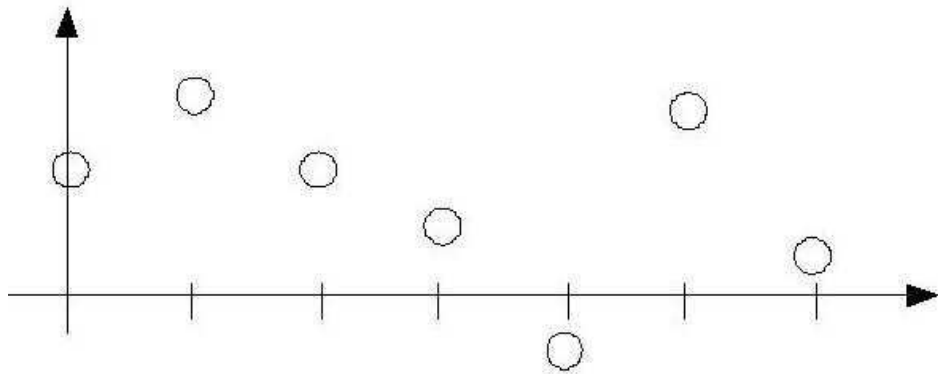
$$(\Omega, F, P) = \{[0, 1], \mathbb{B}([0, 1]), U[0, 1]\}$$

definito come

$$X(t,s) = \begin{cases} 1 & s \in [0, \frac{1}{t}] \\ -1 & s \in (\frac{1}{t}, T] \end{cases} \text{ con } t \in \{1, 2, \dots, \infty\}$$

### Realizzazioni

Fissato  $s$  a scelta, disegniamo una realizzazione del processo nel tempo.



### Variabili casuali

Fissato il tempo, si ha una variabile casuale osservando varie realizzazioni del processo, in quel preciso istante temporale.

*Esempio:*

Nell'esempio, si ha

$$\bar{t} = 1 \Rightarrow X(\bar{t}, s) = \begin{cases} 1 & s \in [0, 1] \\ -1 & s \in \emptyset \end{cases}$$

$$f_V(v, \bar{t} = 1) = \delta(v-1)$$

$$\bar{t} = 2 \Rightarrow X(s, \bar{t}) = \begin{cases} 1 & s \in [0, \frac{1}{2}] \\ -1 & s \in (\frac{1}{2}, 1] \end{cases}$$

$$f_V(v, \bar{t} = 2) = \frac{1}{2}\delta(v-1) + \frac{1}{2}\delta(v+1)$$

$$\bar{t} = 3 \Rightarrow X(s, \bar{t}) = \begin{cases} 1 & s \in [0, \frac{1}{3}] \\ -1 & s \in (\frac{1}{3}, 1] \end{cases}$$

$$f_V(v, \bar{t} = 3) = \frac{1}{3}\delta(v-1) + \frac{2}{3}\delta(v+1)$$

Osservazione: notare che la probabilità di 1 tende a zero per  $E \rightarrow \infty$ , mentre la probabilità di -1 tende a 1 per  $\bar{t} \rightarrow \infty$

### Densità di probabilità

*Esempio:*

Al prim'ordine, si ha

$$f_V(v, t) = \frac{1}{t} \delta(v - 1) + \frac{t - 1}{t} \delta(v + 1)$$

Per il calcolo della densità congiunta del second'ordine, è necessario fissare due istanti temporali  $t_1, t_2, \forall t_1 < t_2$ .

*Esempio:*

Abbiamo tre casi da studiare:

1.  $X(t_1) = 1, X(t_2) = 1$

In questo caso, si ha

$$P(X(t_1) = 1, X(t_2) = 1) = P(\{s \in \Omega \mid X(t_1, s) = 1, X(t_2, s) = 1\}) = P\left(\left\{s \in \left[0, \frac{1}{t_2}\right]\right\}\right)$$

2.  $X(t_1) = 1, X(t_2) = -1$

In questo caso, si ha

$$P(X(t_1) = 1, X(t_2) = -1) = P(\{s \in \Omega \mid X(t_1, s) = 1, X(t_2, s) = -1\}) = P\left(\left\{s \in \left[\frac{1}{t_2}, 1\right]\right\}\right)$$

3.  $X(t_1) = -1, X(t_2) = -1$

In questo caso, si ha

$$P(X(t_1) = -1, X(t_2) = -1) = P(\{s \in \Omega \mid X(t_1, s) = -1, X(t_2, s) = -1\}) = P\left(\left\{s \in \left[1, 1 + \frac{1}{t_2}\right]\right\}\right)$$

Da qui si ottiene che la densità di probabilità congiunta risulta

$$f_{V(t_1)V(t_2)}(v_1, v_2) = \frac{1}{t_1} \delta(v_1 - 1) \delta(v_2 - 1) + \left(\frac{1}{t_1} - \frac{1}{t_2}\right) \delta(v_1 - 1) \delta(v_2 + 1) + \left(1 - \frac{1}{t_1}\right) \delta(v_1 + 1) \delta(v_2 + 1)$$

*Esercizio:*

Calcolare  $f_{V(t_1)}(v_1)$  a partire dalla funzione di densità di probabilità congiunta  $f_{V(t_1)V(t_2)}(v_1, v_2)$

### Caratterizzazione statistica di un processo

Si prendano diversi istanti di tempo  $t_1 < t_2 < \dots < t_n \in T, n \in \mathbb{N}$ . Si definisce la quantità

$$P[X(t_1), X(t_2), \dots, X(t_n) \in B] \forall B \in \mathbb{B}(\mathbb{R}^n)$$

come la probabilità finito-dimensionale del processo  $X(t, x)$ . Dato che  $X(t)$  è una variabile casuale, la quantità appena definita è pari alla probabilità congiunta di un vettore n-dimensionale di variabili casuali, con  $X_i = X(t_i) \forall i = 1, 2, \dots, n$ .

In modo equivalente ai vettori di variabili casuali, un processo è caratterizzato dalla densità di probabilità congiunta, che si indica con

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f_{X(t_1) \dots X(t_n)}(x_1, x_2, \dots, x_n) =_{n=2} f_X(x_1, x_2; t_1, t_2) = f_{X(t_1)X(t_2)}(x_1, x_2)$$

Dalla congiunta è sempre possibile ottenere le marginali, che possono essere pensate come delle congiunte di ordine inferiore.

*Esempio:*

Si ha il processo  $X(t) = V \sin(\omega t)$  con  $\omega$  costante e con una qualsiasi  $f_V(v)$ . Si vogliono calcolare tutte le quantità definite fin'ora. Si ha

$$f_{X(t)} = \begin{cases} \frac{f_V(g^{-1}(x))}{\left|\frac{dg}{dv}(g^{-1}(x))\right|} & \left|\frac{dg}{dv}(g^{-1}(x))\right| \neq 0 \\ 0 & \text{altrimenti} \end{cases} = \begin{cases} \frac{f_V\left(\frac{x}{\sin(\omega t)}\right)}{|\sin(\omega t)|} & |\sin(\omega t)| \neq 0 \\ 0 & \text{altrimenti} \end{cases}$$

Le diverse realizzazioni sono sinusoidi alla stessa frequenza, ma di diversa ampiezza.

*Esercizio:*

Siano le variabili casuali  $X$  e  $Y$ , entrambe distribuite come  $X, Y \sim U[0, 1]$ . Studiare la variabile casuale data dalla relazione

$$Z = \frac{\max(X, Y)}{\min(X, Y)}$$

## Ergodicità dei processi casuali

Supponendo di avere un processo stocastico  $X$  stazionario (in senso lato), andiamo a cercare di dedurre qualcosa dalla sua densità di probabilità  $f_X(x)$  generando una realizzazione nel tempo. Vedremo che se il processo è ergodico, medie d'insieme e temporali coincidono.

Consideriamo un processo  $\{X(t), t \in \mathbb{R}\}$ , con  $X(t, s) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  definito sullo spazio di probabilità  $\{\Omega, \mathcal{F}, P\}$ . Fissato l'esito  $s \in \Omega$ , otteniamo una realizzazione  $X(\cdot, s)$ . Vogliamo capire se (e sotto quali condizioni) è possibile determinare delle caratteristiche di  $X(t)$ , osservando un'unica realizzazione.

Se il processo è stazionario in senso lato, allora

$$\mu_X(t) = \mu_X \quad \forall t \in \mathbb{R}$$

Consideriamo la realizzazione  $X(\cdot, s)$  in una finestra  $\left[-\frac{T}{2}, \frac{T}{2}\right]$ .

Immagine: TFA realizzazione per ergodicità processi.jpg

Definiamo la quantità

$$A_{X,T}(s) = \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} X(t, s) dt$$

che è la media temporale della realizzazione, sul periodo  $T$ . In generale,  $A_{X,T}(s)$  è una funzione dell'esito  $s$ , quindi può essere considerata a sua volta una variabile casuale.

*Definizione: Stimatore del valor medio temporale*

Si dice stimatore del valor medio temporale della realizzazione  $X(\cdot, s)$  la grandezza

$$A_{X,T}(s)$$

La grandezza  $A_{X,T}(\bar{s})$  è una stima, si riferisce ad una particolare realizzazione  $\bar{s}$ .

*Definizione: Stimatore non polarizzato*

Lo stimatore  $A_{X,T}(s)$  è non polarizzato se

$$E[A_{X,T}(s)] = \mu_X$$

ossia se

$$E[A_{X,T}(s)] = \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} E[X(t, s)] dt = \mu_X = E[X(t)]$$

Il valor medio delle medie temporali coincide con il valor medio di insieme.

*Definizione: Stimatore consistente*

Uno stimatore  $A_{X,T}(s)$  è detto consistente se

- $A_{X,T}(s)$  è non polarizzato
- la varianza  $\sigma_{A_{X,T}}$  tende ad annullarsi all'aumentare del tempo di osservazione

$$\sigma_{A_{X,T}} \xrightarrow{T \rightarrow \infty} 0$$

Questo equivale a dire che  $A_{X,T}$  tende, in maniera quadratica, al valor medio del processo, con un tempo di osservazione infinito si può ottenere il valore certo, senza incertezza.

*Definizione: Processo ergodico*

Se  $A_{X,T}$  è uno stimatore consistente di  $\mu_X$ , allora il processo  $X(t)$  è ergodico.

**Teorema: Teorema di Slutsky**

Dato un processo  $\{X(t), t \in \mathbb{R}\}$  WSS (del second'ordine), allora

$$A_{X,T} \xrightarrow{\mathcal{L}^2} \mu_X \Leftrightarrow \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T C_X(\tau) d\tau = 0 \Rightarrow R_X(\tau) \rightarrow \mu_X$$

dove  $\mathcal{L}^2$  indica la convergenza in norma quadratica e  $C_X(\tau)$  è la funzione di autocovarianza del processo.

Il teorema di Slutsky permette di facilitare il compito di verifica dell'ergodicità di un dato processo. Una condizione sufficiente (ma non necessaria) affinché  $X(t)$  sia ergodico è che

$$C_X(\tau) \xrightarrow{\tau \rightarrow \infty} 0$$

**Esempio:**

Sia un processo  $X(t) = V$ , con  $V$  una variabile casuale stazionaria in senso lato. Si ha

$$C_X(\tau) = R_X(\tau) - \mu_V^2 = E[V^2] - \mu_V^2 = \sigma_V^2$$

Dal teorema di Slutsky, si ha

$$\lim_{t \rightarrow \infty} \frac{1}{T} \int_0^T \sigma_V^2 d\tau \neq 0$$

quindi, il processo non è ergodico.

**Esempio:**

Si ha il processo

$$X(t) = A \sin(\omega t) + B \cos(\omega t) + C$$

con  $C$  e  $\omega$  costanti, mentre  $A$  e  $B$  sono variabili casuali incorrelate tra loro ed a media nulla. Si ha

$$\mu_X(t) = C$$

$$C_X(\tau) = \sigma_X^2 \cos(\omega t)$$

Applicando la condizione sufficiente appena vista, si ha

$$C_X(\tau) \not\xrightarrow{\tau \rightarrow \infty} 0$$

Questo fatto non vuol dire che il processo non possa ancora essere ergodico (la condizione è sufficiente, ma non è necessaria).

Dal teorema di Slutsky, si ottiene

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T C_X(\tau) d\tau = 0$$

Da questo, si ottiene che il processo in esame è ergodico (rispetto al valor medio).

In modo alternativo, si può definire l'ergodicità del processo  $\{X(t), t \in T\}$  nel seguente modo:

- se  $X(t)$  è WSS di prim'ordine, allora

$$E[X(t)] = \mu_X(t) = \mu_X$$

Di conseguenza, si ha che  $X(t)$  è un processo ergodico rispetto al valor medio se è verificato

$$P(\{s \in \Omega \mid A_{X,T}(s) = \mu_X\}) = 1$$

Questa si dice *convergenza in probabilità*, o convergenza *qox* (per quasi ogni  $x$ ). Se  $\Omega$  è continuo, possono esistere alcuni  $s$  per cui non è verificata l'equazione, ma se questi punti sono isolati (concetto di *qox*), allora si ha comunque la convergenza cercata. La probabilità di un particolare  $s$ , infatti, è infinitesima (nel caso di processi a tempo continuo, lo stesso non vale per processi a tempo discreto).

- $X(t)$  WSS del second'ordine è ergodico rispetto alla sua funzione di autocorrelazione se vale

$$P(\{s \in \Omega \mid \varphi_{X,T}(\tau, s) = R_X(\tau)\}) = 1$$

Anche questa è una convergenza in probabilità, con

$$\varphi_{X,T}(\tau, s) \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} X(t, s) X(t + \tau, s) dt$$

che è l'autocorrelazione temporale del processo. In termini pratici, questo equivale a dire che

$$R_X(t, t + \tau) = E[X(t)X(t + \tau)] = R_X(\tau) = \varphi_{X,T}(\tau, s)$$

dove  $E[\cdot]$  è la solita stima della funzione di autocorrelazione e  $R_X(\tau)$  è l'autocorrelazione d'insieme, che viene a coincidere con l'autocorrelazione temporale  $\varphi_{X,T}(\tau, s)$ .

Estratto da "[http://it.wikiversity.org/wiki/Processi\\_stocastici,\\_stazionariet%C3%A0\\_ed\\_ergodicit%C3%A0\\_e\\_densit%C3%A0\\_spettrale\\_di\\_potenza](http://it.wikiversity.org/wiki/Processi_stocastici,_stazionariet%C3%A0_ed_ergodicit%C3%A0_e_densit%C3%A0_spettrale_di_potenza)"

Categorie: [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

---

- Ultima modifica per la pagina: 01:37, 14 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.



# Esempi di processi stocastici Processi PAM Processi gaussiani Processi di Markov

Da Wikiversità, l'università aperta.

## Indice

- 1 Processi e catene di Markov
  - 1.1 Densità di probabilità del processo di Markov
  - 1.2 Classificazione
  - 1.3 Catene di Markov a tempo discreto
  - 1.4 Densità spettrale di potenza
    - 1.4.1 Proprietà della densità spettrale di potenza
  - 1.5 Descrizione congiunta dei processi stocastici

## Processi e catene di Markov

Cominciamo col capire cosa sono le catene di Markov. Prendiamo un meteo, in cui si hanno soltanto due stati:

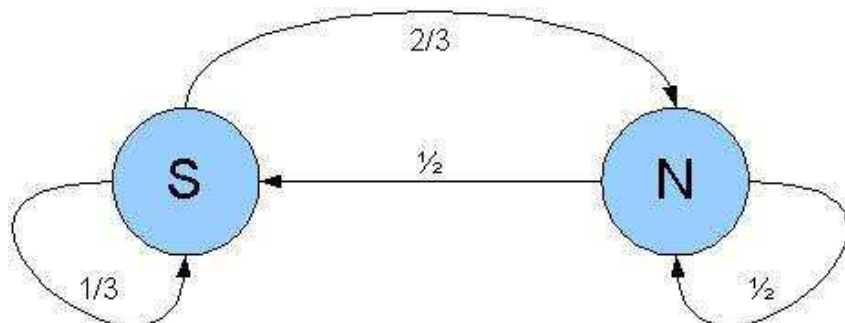
- sole,  $S$ ;
- nuvoloso,  $N$ .

Se oggi c'è il sole, allora domani ci sarà

- sole con probabilità  $\frac{1}{3}$ ;
- nuvoloso con probabilità  $\frac{2}{3}$ .

Al contrario, se oggi è nuvoloso, allora domani ci sarà

- sole con probabilità  $\frac{1}{2}$ ;
- nuvoloso con probabilità  $\frac{1}{2}$ .



Questa è una catena di Markov. Il tempo di domani dipende dal tempo di oggi, e non dal tempo di ieri; inoltre, la probabilità che esce da ogni *stato* è unitaria, quindi anche la probabilità che in un dato istante si sia in un determinato punto soddisferà anch'essa la proprietà di uniterietà.

*Definizione: Processo di Markov*

Il processo  $\{X(t), t \in T\}$  è un processo di Markov se

$$P(X(t_n)|X(t_{n-1}), X(t_{n-2}), \dots, x(t_1)) = P(X(t_n)|X(t_{n-1})) \forall t_1 < t_2 < \dots < t_n \in T$$

In altre parole, possiamo scrivere che una variabile casuale  $X(t_n)$  dipende solo dalla variabile casuale  $X(t_{n-1})$ , mentre è indipendente da tutte le variabili casuali precedenti all'istante  $t_{n-1}$ .

Implicazioni:

1. la probabilità del futuro, dato passato e presente, dipende solo dal presente.
2. la probabilità del futuro, congiunta a quella del passato e conoscendo il presente, è

$$P(\text{futuro}|\text{passato, presente})P(\text{futuro}|\text{presente}) \cdot P(\text{presente}|\text{passato})$$

Gli eventi che soddisfano questa seconda condizione sono detti *condizionalmente indipendenti*, cioè sono indipendenti soltanto se vi è la conoscenza dello stato intermedio. Nel caso in cui non si conosca il presente, allora passato e futuro non sono più indipendenti.

## Densità di probabilità del processo di Markov

$$\begin{aligned} \blacksquare f_{X(t_n)|X(t_{n-1}) \dots X(t_1)}(x_n|x_{n-1}, x_{n-2}, \dots, x_1) &= f_{X(t_n)|X(t_{n-1})}(x_n|x_{n-1}) \\ \blacksquare \begin{cases} f_{X(t_n)|X(t_{n-1}) \dots X(t_1)}(x_n, x_{n-1}, \dots, x_1) = \\ = f_{X(t_n)|X(t_{n-1}), X(t_{n-2}) \dots X(t_1)}(x_n|x_{n-1}, x_{n-2}, \dots, x_1) \times \dots \times f_{X(t_2)|X(t_1)}(x_2|x_1) \times f_{X(t_1)}(x_1) \\ = f_{X(t_n)|X(t_{n-1})}(x_n|x_{n-1}) \times \dots \times f_{X(t_2)|X(t_1)}(x_2|x_1) \times f_{X(t_1)}(x_1) \\ = \left[ \prod_{k=1}^{n-1} f_{X(t_{k+1})|X(t_k)}(x_{k+1}|x_k) \right] \cdot f_{X(t_1)}(x_1) \end{cases} \end{aligned}$$

Per caratterizzare una catena di Markov è sufficiente conoscere la densità di probabilità del second'ordine, non serve scendere fino all' $n$ -esimo ordine. In questo modo, la trattazione diventa molto più semplice, fino a renderla quasi banale.

## Classificazione

I processi di Markov si possono classificare in base allo stato/tempo continuo/discreto:

- stato continuo e tempo continuo: *processo a tempo continuo*;
- stato continuo e tempo discreto: *processo a tempo discreto*;
- stato discreto e tempo continuo: *catena a tempo continuo*;
- stato discreto e tempo discreto: *catena a tempo discreto*.

## Catene di Markov a tempo discreto

Si ha

$$\begin{aligned} \blacksquare T &= \{0, 1, \dots\} \\ \blacksquare S &= \{1, 2, \dots\} \end{aligned}$$

che sono due insiemi discreti. La catena di Markov è caratterizzata da due quantità:

- la probabilità incondizionata  $P_i = P(X(n) = i)$ ;
- la matrice delle probabilità di transizione  $P_{ij}(m, n) = P(X(n) = j | X(m) = i)$ .

Si ha

$$P_i(n) \in [0, 1] \forall i \in S, \forall n \in T \mid \sum_{i \in S} P_i(n) = 1$$

cioè, per ogni istante, la somma delle probabilità di tutti gli stati è unitaria; noto l'alfabeto  $s$ , la probabilità incondizionata si può scrivere come

$$\underline{P}(n) = [P_1(n) P_2(n) \cdots P_{|s|}(n)]$$

La stessa condizione si può esprimere con

$$\underline{P} \cdot e^T = 1 \quad e = [1 \ 1 \ \cdots \ 1]$$

Fissati due istanti temporali  $n < m$  e dato lo stato di partenza  $n$ , la somma delle probabilità degli stati di arrivo è unitaria.

$$\underline{P}(m, n) = \begin{bmatrix} P_{1,1}(m, n) & P_{1,2}(m, n) & \cdots & P_{1,|s|}(m, n) \\ P_{2,1}(m, n) & P_{2,2}(m, n) & \cdots & P_{2,|s|}(m, n) \\ \vdots & \vdots & \ddots & \vdots \\ P_{|s|,1}(m, n) & P_{|s|,2}(m, n) & \cdots & P_{|s|,|s|}(m, n) \end{bmatrix}$$

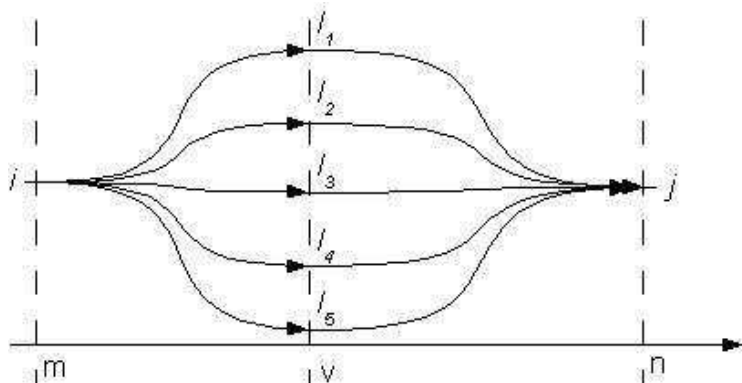
dove la somma dei valori di ogni singola riga è 1

$$\sum_{i=1,2,\dots,|s|} P_{i,j}(m, n) = 1 \quad \forall j \in [0, 1, \dots, |s|]$$

$$\underline{P}(m, n) \cdot e^T = e^T = [1 \ 1 \ \cdots \ 1]$$

Per ogni coppia  $m, n$ , il valore di  $\underline{P}(m, n)$  sarà diverso. Si ha

$$\underline{P}(n) = \underline{P}(m, n) \cdots \underline{P}(m)$$



$$\begin{aligned} P_{i,j}(m, n) &= P(X(n) = j | X(m) = i) \\ &= \sum_{l \in S} P(X(n) = j | X(u) = l, X(m) = i) \cdot P(X(u) = l | X(m) = i) \\ &= \sum_{l \in S} P_{l,j}(u, n) \cdot P_{i,l}(m, u) \end{aligned}$$

Quest'ultima è detta l'equazione di Chapman e Kolmogorov, che in forma matriciale si può scrivere come

$$\underline{P}(m, n) = \underline{P}(m, u) \cdot \underline{P}(u, n)$$

da cui si ha

$$\underline{P}(m, n) = \prod_{k=m}^{n-1} \underline{P}(k, k+1)$$

Per caratterizzare una catena di Markov, basta conoscere:

- $\underline{P}(0)$
- $\underline{P}(k, k+1) \quad \forall k \in T$
- $\underline{P}(n) = \underline{P}(0) \cdot \underline{P}(0, 1) \cdot \underline{P}(1, 2) \cdots \underline{P}(n-1, n)$

**Esempio: Caso 1**

Si ha

$$\underline{P}(1) = \underline{P}(0) \cdot \underline{P} = [1 \ 0] \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [0 \ 1]$$

$$\underline{P}(2) = \underline{P}(0) \cdot \underline{P}^2 = [1 \ 0] \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [1 \ 0]$$

La sequenza degli stati è deterministica,

- $P(2m) = [1 \ 0] \ \forall m \in \mathbb{N}$
- $P(2m + 1) = [0 \ 1] \ \forall m \in \mathbb{N}$

**Esempio: Caso 2**

Si ha

$$\underline{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$P(0) = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

da cui

$$\underline{P}(1) = \underline{P}(0) \cdot \underline{P} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

$$\underline{P}(2) = \underline{P}(0) \cdot \underline{P}^2 = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

da cui si deduce che vale

$$\underline{P}(m) = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \ \forall m \in T$$

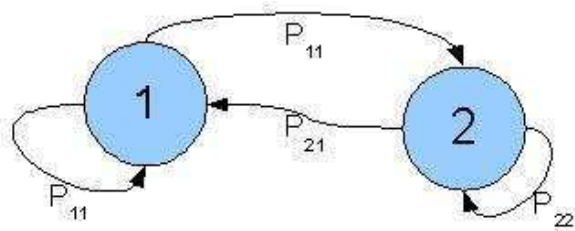
**Definizione: Catena di Markov omogenea**

Una catena di Markov è detta omogenea se

$$\underline{P}(k, k + 1) = \underline{P} \ \forall k \in T$$

ossia, la matrice di probabilità di transizione ad un passo  $(k, k + 1)$  è la stessa per tutti i  $k \in T$ .**Esempio:**

Si ha



$$\underline{P} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

da cui

- $\underline{P}(0) = [P_1 \ P_2]$
- $s = \{1, 2\}$
- $T = \{0, 1, \dots\}$

**Definizione: Distribuzione stazionaria**

Data una catena di Markov omogenea, si definisce distribuzione stazionaria il vettore di probabilità

$$\underline{\Pi} = [\pi_1 \ \pi_2 \ \dots \ \pi_{|s|}]$$

tale che

- $\underline{\Pi} \cdot \underline{P} = \underline{\Pi}$
- $\underline{\Pi} \cdot \underline{e}^T = 1$

da cui, nel caso in cui  $P(0) = \underline{\Pi}$ , allora si ottiene  $P(n) = \underline{\Pi} \ \forall n \in T$ .

In generale, una catena di Markov può avere più distribuzioni stazionarie.

**Esempio:**

Sia

$$\underline{P} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

da cui si ha

- $\Pi_1 = \left[ \frac{1}{2} \ \frac{1}{2} \ 0 \right]$
- $\Pi_2 = [0 \ 0 \ 1]$
- $\Pi_3 = \left[ \frac{1}{4} \ \frac{1}{4} \ \frac{1}{4} \right]$

**Definizione: Catena di Markov irriducibile**

Una catena di Markov si dice irriducibile se non è possibile portare la matrice di probabilità di transizione in una forma diagonale a blocchi, del tipo

$$\begin{bmatrix} \left\{ \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\} & 0 \\ & \left\{ \begin{matrix} 1 & 0 \\ 0 & 0 \end{matrix} \right\} & 0 \\ & & 1 \end{bmatrix}$$

Se una catena di Markov è irriducibile, allora la distribuzione stazionaria esiste ed è unica.

**Definizione: Distribuzione limite**

Una distribuzione stazionaria  $\underline{\Pi}$  si dice distribuzione limite se

$$\lim_{n \rightarrow \infty} \underline{P}(n) = \underline{P} \ \forall P(0)$$

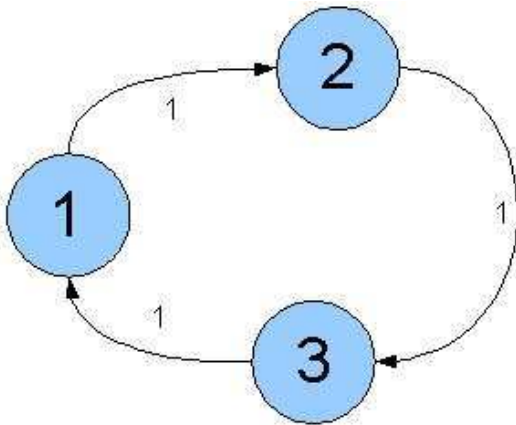
**Definizione: Catena di Markov aperiodica**

Una catena di Markov omogenea ed irriducibile è aperiodica se il massimo comun divisore delle lunghezze di

tutti i cammini chiusi che si possono individuare sul diagramma delle transizioni è pari a 1.

*Esempio:*

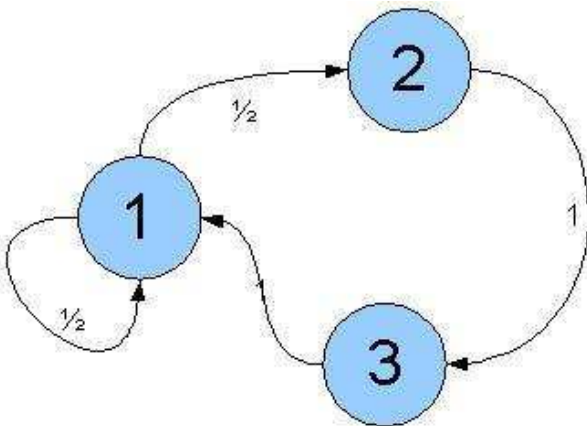
*Catena periodica di periodo 3:*



*In questo caso, MDC = 3.*

*Esempio:*

*Catena di Markov aperiodica:*



*In questo caso, MDC = 1.*

Se una catena di Markov omogenea ed irriducibile è aperiodica, allora la distribuzione stazionaria è anche distribuzione limite. Per queste catene di Markov, a regime la probabilità assoluta è indipendente dal tempo, dando origine a processi stazionari.

Nota: una distribuzione  $\underline{\Pi}$  è una distribuzione limite se

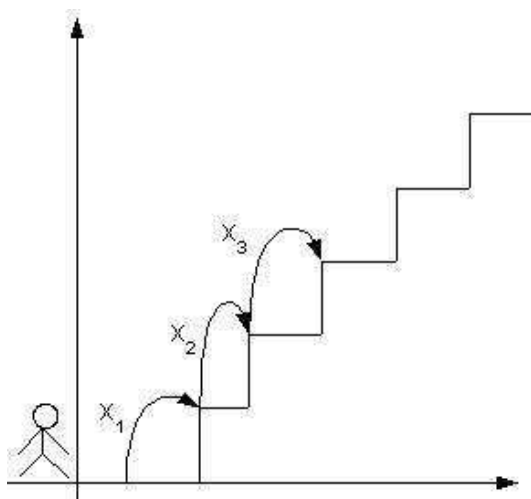
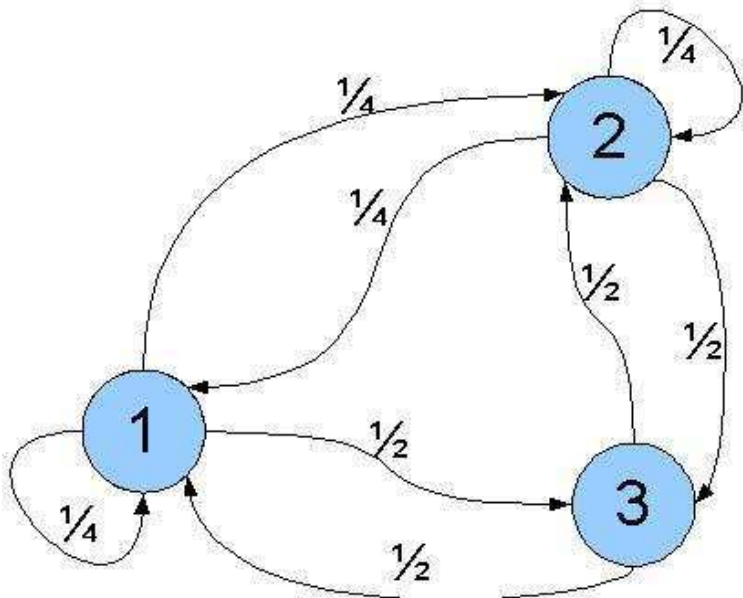
$$\lim_{n \rightarrow \infty} \underline{P}^n = \begin{bmatrix} \underline{\Pi} \\ \underline{\Pi} \\ \underline{\Pi} \end{bmatrix}$$

*Definizione: Matrice doppiamente stocastica*

Una matrice di probabilità di transizione è doppiamente stocastica se la somma degli elementi di ciascuna colonna è unitario; in tal caso, la distribuzione limite risulta essere

$$\underline{\Pi} = \left[ \begin{array}{ccc} \frac{1}{|s|} & \frac{1}{|s|} & \dots & \frac{1}{|s|} \end{array} \right]$$

*Esempio:*



Si ha

- $\underline{P} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$
- $\underline{\Pi} = [\pi_1 \ \pi_2 \ \dots \ \pi_3]$

da cui

$$\begin{cases} \underline{\Pi} \cdot \underline{P} = \underline{\Pi} \\ \underline{\Pi} \cdot \underline{e}^T = 1 \end{cases}$$

$$\begin{cases} \frac{1}{4}\pi_1 + \frac{1}{4}\pi_2 + \frac{1}{2}\pi_3 = \pi_1 \\ \dots = \pi_2 \\ \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2 = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

si ottiene

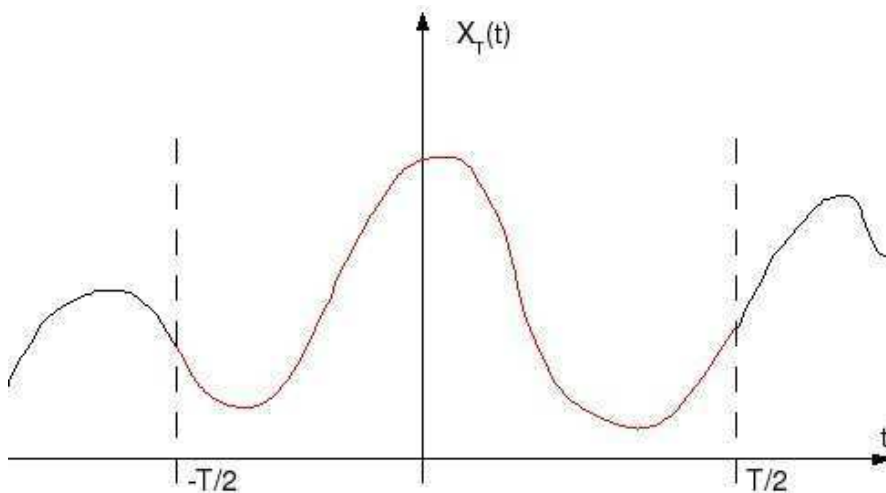
$$\underline{\Pi} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 3 & 3 \end{bmatrix}$$

che è una matrice doppiamente stocastica.

## Densità spettrale di potenza

Consideriamo il processo  $\{X(t), t \in T\}$ . Fissato  $s \in \Omega$  otteniamo una realizzazione che in generale rappresenta un segnale di potenza non periodico, quindi non è possibile dare una caratterizzazione frequenziale attraverso la trasformata di Fourier (in modo diretto). E' tuttavia possibile caratterizzare i processi casuali almeno in termini di *spettro di potenza*. A questo proposito, consideriamo il processo

$$X_T(t, s) = X(t, s) \cdot \text{rect}\left(\frac{t}{T}\right)$$



Limitatamente al periodo  $T$ , il segnale diventa ad energia finita,

$$E = \int_{-\infty}^{+\infty} |X_T(t, s)|^2 dt = \int_{-T/2}^{T/2} |X(t, s)|^2 dt < \infty$$

quindi si ha anche lo spettro di potenza finito,

$$\frac{1}{T} = \int_{-\infty}^{+\infty} |X_T(f, s)|^2 df$$

*Definizione: Densità spettrale di potenza*

Di definisce densità spettrale di potenza del processo  $\{X(t), t \in T\}$  la grandezza

$$S_X(f) = \lim_{T \rightarrow \infty} \frac{1}{T} E [ |X_T(f, s)|^2 ]$$

dove  $X_T(f, s)$  è una variabile casuale ottenuta da

$$X_T(f, s) = \int_{-\infty}^{+\infty} X_T(t, s) e^{-j2\pi ft} dt$$

che è la trasformata di Fourier della variabile casuale  $X_T(t, s)$ .

## Proprietà della densità spettrale di potenza

- $S_X(f) \geq 0 \quad \forall f$



2.  $\frac{|X_T(f, s)|^2}{T}$  è detto *periodogramma* del processo

3. Teorema di Wiener-Kinchine; se un processo stocastico è stazionario in senso lato (del second'ordine), allora si ha

$$S_X(f) = F[R_X(\tau)] = \int_{-\infty}^{+\infty} R_X(\tau) e^{-j2\pi f\tau} d\tau$$

4. Se  $X(t)$  è a valori reali, allora  $S_X(f)$  è pari, di conseguenza si ha che

$$R_X(\tau) = R_X(-\tau)$$

5.  $\int_{-\infty}^{+\infty} S_X(f) df = R_X(0) = P_X$  che è la potenza del processo

6.  $R_X(\tau) = F^{-1}[S_X(f)] = \int_{-\infty}^{+\infty} S_X(f) e^{j2\pi f\tau} df$

#### Esempio:

Sia un processo stocastico  $X(t) = V$  con  $V$  una variabile casuale con  $f_V(v)$  qualsiasi. Si ha:

- $R_X(\tau) = E[X(t)X(t+\tau)] = E[V^2] = \sigma_V^2 + \mu_V^2$
- $S_X(f) = (\sigma_V^2 + \mu_V^2) \cdot \delta(f)$

cioè, l'autocorrelazione  $R_X(\tau)$  è costante, quindi la densità spettrale di potenza del processo è una  $\delta$  nell'origine.

#### Esempio:

Si ha il processo stocastico

$$X(t) = A \sin(\omega t) + B \cos(\omega t) + C$$

con  $C$  e  $\omega$  definiti, con  $A$  e  $B$  variabili casuali incorrelate e a media nulla. Allora, si ha

- $E[X(t)] = C$
- $R_X(\tau) = \sigma^2 \cos(\omega\tau) + C^2$
- $S_X(f) = \frac{\sigma^2}{2} \left[ \delta\left(f + \frac{\omega}{2\pi}\right) + \delta\left(f - \frac{\omega}{2\pi}\right) \right] + C^2 \delta(f)$

## Descrizione congiunta dei processi stocastici

Consideriamo  $\{X(t), t \in T\}$  e  $\{Y(t), t \in T\}$  due processi definiti sullo stesso spazio di probabilità  $(\Omega, F, P)$ . Se volessimo caratterizzare congiuntamente i due processi, dovremmo fare

$$f_{X,Y}(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m; t_1, t_2, \dots, t_n, t'_1, t'_2, \dots, t'_m)$$

e questo dovrebbe valere:

- $\forall t_1, t_2 < \dots < t_n \in T$
- $\forall t'_1, t'_2 < \dots < t'_m \in T$
- $\forall m, n \in \mathbb{N}$

Questa è la densità congiunta finito-dimensionale; passando invece alle descrizioni sintetiche, si possono identificare:

- i valori medi

$$\mu_X(t) = E[X(t)]$$

$$\mu_Y(t) = E[Y(t)]$$

- le funzioni di autocorrelazione

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)]$$

$$R_Y(t_1, t_2) = E[Y(t_1)Y(t_2)]$$

- le funzioni di covarianza

$$C_X(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))]$$

$$C_Y(t_1, t_2) = E[(Y(t_1) - \mu_Y(t_1))(Y(t_2) - \mu_Y(t_2))]$$

- le funzioni di crosscorrelazione

$$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)]$$

$$R_{YX}(t_1, t_2) = E[Y(t_1)X(t_2)]$$

- le funzioni di crosscovarianza

$$C_{XY}(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(Y(t_2) - \mu_Y(t_2))] = R_{XY}(t_1, t_2) - \mu_X(t_1)\mu_Y(t_2)$$

$$C_{YX}(t_1, t_2) = E[(Y(t_1) - \mu_Y(t_1))(X(t_2) - \mu_X(t_2))] = R_{YX}(t_1, t_2) - \mu_Y(t_1)\mu_X(t_2)$$

#### Definizione: *Indipendenza dei processi*

Due processi  $\{X(t), t \in T\}$  e  $\{Y(t), t \in T\}$  definiti sullo stesso spazio di probabilità  $(\Omega, F, P)$  sono indipendenti se

$$f_{X,Y}(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m; t_1, t_2, \dots, t_n, t'_1, t'_2, \dots, t'_m) = f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \cdot f_Y(y_1, y_2, \dots, y_m; t'_1, t'_2, \dots, t'_m)$$

e questo deve valere:

- $\forall m, n \in \mathbb{N}$
- $\forall t_1 < t_2 < \dots < t_n \in T$
- $\forall t'_1 < t'_2 < \dots < t'_m \in T$

#### Definizione: *Processi incorrelati*

Due processi  $\{X(t), t \in T\}$  e  $\{Y(t), t \in T\}$  definiti sullo stesso spazio di probabilità  $(\Omega, F, P)$  sono incorrelati se

$$E[X(t_1)Y(t_2)] = R_{XY}(t_1, t_2) = E[X(t_1)] \cdot E[Y(t_2)] = \mu_X(t_1)\mu_Y(t_2)$$

Nel caso particolare in cui i processi siano delle variabili casuali gaussiane, allora si ha che

incorrelazione  $\Leftrightarrow$  indipendenza

mentre in tutti gli altri casi si ha

incorrelazione  $\not\Leftrightarrow$  indipendenza

indipendenza  $\Rightarrow$  incorrelazione

**Definizione: Processi congiuntamente gaussiani**

Un processo  $\{X(t), t \in T\}$  è congiuntamente gaussiano se tutte le densità di probabilità finito-dimensionali sono congiuntamente gaussiane.

**Definizione: Processi congiuntamente stazionari in senso stretto**

Due processi  $\{X(t), t \in T\}$  e  $\{Y(t), t \in T\}$  si dicono congiuntamente stazionari in senso stretto (SSS) se la densità di probabilità congiunta è invariante alla traslazione temporale.

**Definizione: Processi congiuntamente stazionari in senso lato**

Due processi  $\{X(t), t \in T\}$  e  $\{Y(t), t \in T\}$  si dicono congiuntamente stazionari in senso lato (WSS) se:

$$1. \begin{cases} \mu_X(t) = \mu_X \\ \mu_Y(t) = \mu_Y \\ R_X(t, t + \tau) = R_X(\tau) \\ R_Y(t, t + \tau) = R_Y(\tau) \end{cases}$$

$$2. R_{XY}(t, t + \tau) = R_{XY}(\tau) \quad \forall t, \tau \in T$$

cioè, anche la crosscorrelazione deve essere funzione della sola distanza tra istanti temporali.

Nel caso di processi congiuntamente stazionari in senso lato (WSS), si ha

$$R_{YX}(t, t + \tau) = R_{YX}(0, \tau) = R_{XY}(\tau) \quad \forall t, \tau \in T$$

e vale

$$R_{YX}(\tau) = R_{XY}(-\tau)$$

Inoltre, vale

$$|R_{XY}(0)|^2 \leq R_X(0)R_Y(0)$$

**Esempio:**

Se  $X(t)$  e  $Y(t)$  sono congiuntamente WSS, l'autocorrelazione di  $Z(t) = X(t) + Y(t)$  vale

$$\begin{aligned} R_Z(t, t + \tau) &= E[((X(t_1) + Y(t_1)) \cdot (X(t_2) + Y(t_2)))] \\ &= E[(X(t_1)X(t_2) + Y(t_1)Y(t_2) + X(t_1)Y(t_2) + Y(t_1)X(t_2))] \\ &= R_X(t_1, t_2) + R_Y(t_1, t_2) + R_{XY}(t_1, t_2) + R_{YX}(t_1, t_2) \\ &= R_X(\tau) + R_Y(\tau) + R_{XY}(\tau) + R_{YX}(\tau) \end{aligned}$$

$$\mu_Z(t) = E[X(t) + Y(t)] = \mu_X(t) + \mu_Y(t) = \mu_X + \mu_Y$$

**Esempio:**

Se  $X(t)$  e  $Y(t)$  sono congiuntamente WSS e indipendenti, l'autocorrelazione di  $Z(t) = X(t)Y(t)$  vale

$$\mu_Z(t) = E[X(t)Y(t)] = \mu_X\mu_Y$$

$$\begin{aligned} R_Z(t, t + \tau) &= E[Z(t)Z(t + \tau)] \\ &= E[X(t)Y(t)X(t + \tau)Y(t + \tau)] \\ &= E[(X(t)X(t + \tau))(Y(t)Y(t + \tau))] \\ &= E[X(t)X(t + \tau)] \cdot E[Y(t)Y(t + \tau)] \\ &= R_X(\tau)R_Y(\tau) \end{aligned}$$

### LAVORI IN CORSO! - WORK IN PROGRESS!



**Davide Morellato** si sta occupando di questo testo; non apportare modifiche se l'ultima modifica

(<http://it.wikiversity.org>

[/w/index.php?title=Esempi\\_di\\_processi\\_stocastici\\_Processi\\_PAM\\_Processi\\_gaussiani\\_Processi\\_di\\_Markov&action=history](http://it.wikiversity.org/w/index.php?title=Esempi_di_processi_stocastici_Processi_PAM_Processi_gaussiani_Processi_di_Markov&action=history)) è recente.

Estratto da "<http://it.wikiversity.org>

[/wiki/Esempi\\_di\\_processi\\_stocastici\\_Processi\\_PAM\\_Processi\\_gaussiani\\_Processi\\_di\\_Markov](http://it.wikiversity.org/wiki/Esempi_di_processi_stocastici_Processi_PAM_Processi_gaussiani_Processi_di_Markov)"

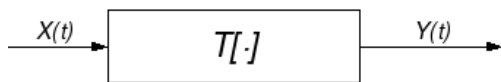
Categorie: [WIP](#) | [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

---

- Ultima modifica per la pagina: 15:38, 15 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Elaborazione lineare e nonlineare di un processo stocastico

Da Wikiversità, l'università aperta.



Si vuole modificare un processo stocastico  $X(t)$  con un sistema  $T[\cdot]$ , che è una trasformazione.

*Definizione: Sistema tempo-invariante*

Un sistema  $T[\cdot]$  si dice tempo-invariante se, dato  $Y(t,s) = T[X(t,s)]$ , si ha che

$$Y(t + \tau, s) = T[X(t + \tau, s)] \quad \forall t, \tau \in T, \quad \forall s \in \Omega$$

## Indice

- 1 Classificazione dei sistemi
- 2 Sistemi tempo-invarianti
  - 2.1 Autocorrelazione del second'ordine
  - 2.2 Covarianza di Y(t)
- 3 Processi bianchi
  - 3.1 Processi bianchi discreti
- 4 Processi ciclostazionari

## Classificazione dei sistemi

- $Y(t) = X^2(t)$  è tempo invariante, non lineare;
- $Y(t) = X(t) - X(t - 1)$  è tempo invariante, lineare;

Se il sistema  $T[\cdot]$  è tempo invariante e  $X(t)$  è stazionario in senso stretto di ogni ordine, allora anche il processo  $Y(t) = T[X(t)]$  è stazionario in senso stretto (SSS) di ogni ordine.

In generale, il fatto che  $X(t)$  sia SSS non vuol dire che  $Y(t) = T[X(t)]$  sia anch'esso SSS.

*Esempio:*

Si ha

$$Y(t) = X(t) - X(t - 1) = Z(t) - W(t)$$

$$f_Y(\beta) = \int_{-\infty}^{+\infty} f_{X(t)-X(t-1)}(\alpha, \beta - \alpha) d\alpha$$

*Esempio:*

Sia  $T[\cdot]$  un sistema tempo invariante,  $X(t)$  WSS e  $Y(t) = X^3(t)$ . Sia

$$X(t) = A \sin(\omega t) + B \cos(\omega t) + C$$

con  $\omega$  e  $C$  costanti, mentre  $A$  e  $B$  due variabili casuali indipendenti e a media nulla. Allora, si ha

$$E[Y(t)] = E[A^3 \sin^3(\omega t) + B^3 \cos^3(\omega t) + \dots] = e(t)$$

Siccome la media è una funzione del tempo, allora il sistema è sicuramente non-lineare; infatti, si ha

$$\mu_Y(0) = E[B^3] \neq \mu_Y\left(\frac{\pi}{2\omega}\right) = E[A^3]$$

## Sistemi tempo-invarianti

I sistemi tempo-invarianti possono essere di due tipi:

- sistemi statici, o *istantanei*;
- sistemi dinamici, o *con memoria*.

*Definizione: Sistema lineare*

Un sistema tempo-invariante definito dalla trasformazione  $T[\cdot]$  è lineare se vale

$$T[aX_1(t) + bX_2(t)] = aT[X_1(t)] + bT[X_2(t)] \quad \forall a, b \in \mathbb{R}, \forall X_1, X_2$$

**Teorema:**

Per un sistema lineare tempo-invariante (LTI) vale

$$E[T[X(t)]] = T[E[X(t)]]$$

$$\mu_Y(t) = T[\mu_X(t)]$$

Nel caso di sistemi LTI, si ha

$$Y(t) = \int_{-\infty}^{+\infty} h(\tau)X(t-\tau) = \int_{-\infty}^{+\infty} X(\tau)h(t-\tau)d\tau = h(t) * X(t) = X(t) * h(t)$$

dove  $h(t) \quad \forall t \in T$  è la risposta all'impulso del sistema  $T[\cdot]$ . Si ha, inoltre,

$$H(f) = \int_{-\infty}^{+\infty} h(t)e^{-j2\pi ft} dt$$

che è la risposta in frequenza del sistema. Se  $X(t)$  è stazionario in senso lato, si può affermare che  $Y(t)$  è ancora stazionario in senso lato; infatti, si ha

$$\begin{aligned} \mu_Y(t) &= E[Y(t)] \\ &= E[h(t) * X(t)] \\ &= E\left[\int_{-\infty}^{+\infty} h(\tau)X(t-\tau)d\tau\right] \\ &= \int_{-\infty}^{+\infty} h(\tau)E[X(t-\tau)]d\tau \\ &= \int_{-\infty}^{+\infty} h(\tau)\mu_X d\tau \\ &= \mu_X \int_{-\infty}^{+\infty} h(\tau)d\tau = \mu_X H(0) \end{aligned}$$

dove  $\mu_X(t) = \mu_X$  perchè  $X(t)$  è, per ipotesi, stazionario.

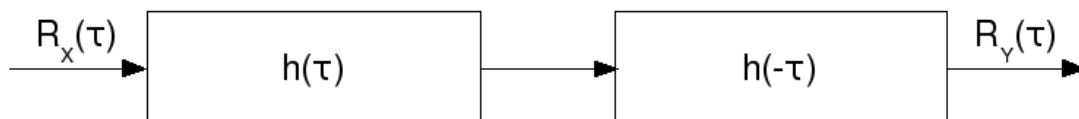
$H(0)$  è la risposta all'impulso del sistema alla frequenza  $f=0$ , il *guadagno di sistema*.

### Autocorrelazione del second'ordine

Per i sistemi LTI, vale

$$\begin{aligned} R_Y(t, t + \tau) &= E[Y(t)Y(t + \tau)] \\ &= E\left[\int_{-\infty}^{+\infty} h(t')X(t-t')dt' \cdot \int_{-\infty}^{+\infty} h(t'')X(t + \tau - t'')dt''\right] \\ &= R_X(\tau) * h(\tau) * h(-\tau) \end{aligned}$$

Riassumendo: se il processo di partenza  $X(t)$  è stazionario in senso lato (WSS) del second'ordine, e se il sistema è lineare tempo-invariante (LTI), allora l'uscita del sistema  $Y(t)$  è anch'essa WSS.



In frequenza, si ha

$$S_Y(f) = F[R_Y(\tau)] = S_X(f) \cdot H(f) \cdot H^*(f) = S_X(f) \cdot |H(f)|^2$$

I processi  $X(t)$  e  $Y(t)$  sono anche congiuntamente stazionari, infatti

$$\begin{aligned}
 R_{XY}(t, t + \tau) &= E[X(t)Y(t + \tau)] \\
 &= E \left[ X(t) \int_{-\infty}^{+\infty} h(t' + \tau - t') dt' \right] \\
 &= \int_{-\infty}^{+\infty} h(t') E[X(t)X(t + \tau - t')] dt' \\
 &= h(\tau) * R_X(\tau) = R_{XY}(\tau)
 \end{aligned}$$

In termini di densità spettrali, si ha

$$S_{XY}(f) = H(f) \cdot S_X(f)$$

### Covarianza di $Y(t)$

Per quanto riguarda  $X(t)$ , abbiamo che vale

$$C_X(t, t + \tau) = R_X(t, t + \tau) - \mu_X(t)\mu_X(t + \tau) = R_X(\tau) + \mu_X^2 = C_X(\tau)$$

Consideriamo il sistema  $h(t)$  che accetta in ingresso il processo  $\tilde{X}(t)$  e restituisce il processo  $\tilde{Y}(t)$ , con

- $\tilde{X}(t) = X(t) - \mu_X$
- $\tilde{Y}(t) = Y(t) - \mu_Y$

dove  $X(t)$  e  $Y(t)$  sono stazionari in senso lato (WSS). Si ha

- $R_{\tilde{X}}(t, t + \tau) = E[\tilde{X}(t) \cdot \tilde{X}(t + \tau)] = C_X(t, t + \tau) = R_{\tilde{X}}(\tau) = C_X(\tau)$
- $R_{\tilde{Y}}(\tau) = R_{\tilde{X}}(\tau) * h(\tau) * h(-\tau) = C_X(\tau) * h(\tau) * h(-\tau)$

Abbiamo che

$$\tilde{Y}(t) = h(t) * \tilde{X}(t) = h(t) * (X(t) - \mu_X) = Y(t) - \mu_Y$$

da cui si ottiene

$$R_{\tilde{Y}}(\tau) = C_Y(\tau)$$

ossia

$$C_Y(\tau) = C_X(\tau) * h(\tau) * h(-\tau)$$

che è la stessa relazione che esiste per l'autocorrelazione.

#### Esempio:

Sia

$$Y(t) = X(t) + X(t - t_0)$$

con  $\mu_X$ ,  $R_X(\tau)$  e WSS del second'ordine. Si ha

$$\mu_Y(t) = E[Y(t)] = E[X(t) + X(t - t_0)] = E[X(t)] + E[X(t - t_0)] = \mu_X(t) + \mu_X(t - t_0) = 2\mu_X$$

$$\begin{aligned}
 R_Y(t, t + \tau) &= E[Y(t)Y(t + \tau)] \\
 &= E[(X(t) + X(t - t_0))(X(t + \tau) + X(t - t_0 + \tau))] \\
 &= E[X(t)X(t + \tau) + X(t)X(t - t_0 + \tau) + X(t - t_0)X(t + \tau) + X(t - t_0)X(t - t_0 + \tau)] \\
 &= 2 \cdot R_X(\tau) + R_X(\tau - t_0) + R_X(\tau + t_0)
 \end{aligned}$$

Un modo alternativo per calcolare la funzione di autocorrelazione è usare la convoluzione,

$$R_Y(\tau) = R_X(\tau) * h(\tau) * h(-\tau)$$

dove si ha

$$h(t) = \delta(t) + \delta(t - t_0)$$

quindi si ha

$$\begin{aligned}
 R_Y(\tau) &= R_X(\tau) * (\delta(\tau) + \delta(\tau - t_0)) * (\delta(-\tau) * \delta(-\tau - t_0)) \\
 &= (R_X(\tau) + R_X(\tau - t_0)) * (\delta(\tau) + \delta(-\tau - t_0)) \\
 &= R_X(\tau) + R_X(\tau - t_0) + R_X(\tau + t_0) + R_X(\tau - t_0 + t_0) = 2 \cdot R_X(\tau) + R_X(\tau - t_0) + R_X(\tau + t_0)
 \end{aligned}$$

**Esercizio: Esercizio per casa**

Si ha il processo stocastico

$$Y(t) = X(t) + X(t - t_0)$$

con  $X(t)$  un processo gaussiano WSS. Calcolare:

- $\mu_Y$
- $R_Y(\tau)$
- $f_X(a, t)$
- $f_{X_1, X_2}(\alpha_1, \alpha_2; t_1, t_2)$

Nota: se  $X(t)$  è gaussiano e WSS, allora è anche stazionario in senso stretto (SSS). Se poi filtriamo tale processo SSS con un sistema lineare tempo-invariante (LTI), allora anche  $Y(t)$  sarà stazionario in senso lato (WSS).

**Processi bianchi****Definizione: Processo bianco**

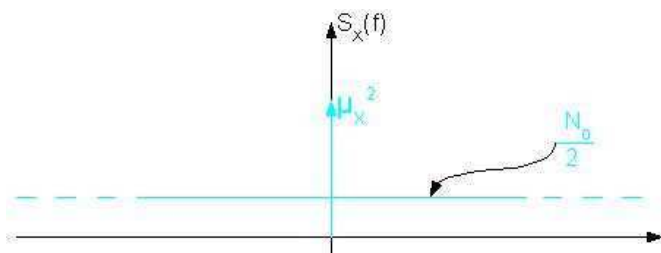
Un processo  $\{X(t), t \in T\}$  si dice bianco se:

1.  $X(t)$  è stazionario in senso lato (WSS) almeno del second'ordine;
2. la trasformata di Fourier dell'autocovarianza è costante

$$F[C_X(\tau)] = \text{cost}$$

il che vuol dire che la densità spettrale di potenza è

$$S_X(f) = F[C_X(\tau)] + \mu_X^2 \cdot \delta(f)$$



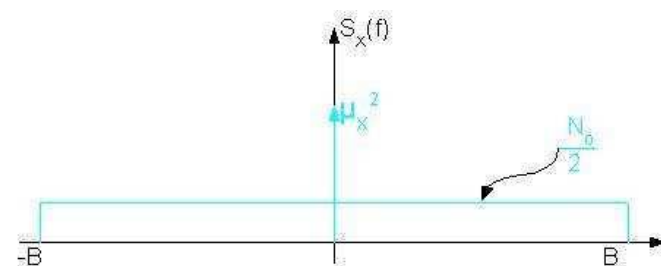
Nella realtà, i processi bianchi continui non esistono, perché la potenza sarebbe infinita con  $N_0 \neq 0$ .

$$P_X = \int_{-\infty}^{+\infty} S_X(f) df = +\infty$$

Quindi, dobbiamo restringere la trattazione ai processi bianchi in banda, cioè con densità spettrale di potenza costante su una banda limitata  $B > 0$ .

**Definizione: Processo bianco in banda**

Un processo bianco in banda ha la trasformata di Fourier della covarianza,  $C_X(f)$ , costante nell'intervallo  $[-B, B]$ .



Si ha

$$S_X(f) = \frac{N_0}{2} \cdot \text{rect}\left(\frac{f}{2B}\right) + \mu_X^2 \delta(f)$$

In questo caso, la potenza è

$$P_X = \int_{-\infty}^{+\infty} S_X(f) df = \int_{-B}^B \left(\frac{N_0}{2} + \mu_X^2 \delta(f)\right) df = \frac{N_0}{2} (2B) + \mu_X^2 = N_0 B + \mu_X^2$$

Si ha, quindi,



$$N_0 = \frac{P_X - \mu_X}{B} = \frac{\sigma_X}{B}$$

Nel caso di processi bianchi in banda limitata, la funzione di autocorrelazione è

$$C_X(\tau) = F^{-1} \left[ \frac{N_0}{2} \text{rect} \left( \frac{f}{2B} \right) \right] = \frac{N_0}{2} 2B \cdot \text{sinc}(2B\tau) = \sigma_X^2 \text{sinc}(2B\tau)$$

### Processi bianchi discreti

Se un processo è bianco e discreto (per esempio, può essere la versione campionata di un processo continuo), si ha sempre potenza finita nel periodo:

$$P_X = R_X(0) = \int_{-\frac{1}{2}}^{+\frac{1}{2}} S_X(f) df$$

Un processo bianco discreto, essendo la versione campionata di un processo bianco continuo, è sempre implicitamente considerato come in banda: per il teorema di Shannon, infatti, un segnale deve essere campionato ad una frequenza almeno doppia della banda del segnale,

$$f_C \geq 2B$$

quindi, deve esistere il valore

$$B < \infty$$

#### Esempio: Esempio di tema d'esame

Sia  $\{X(t), t \in \mathbb{R}\}$  un processo gaussiano, stazionario in senso lato (WSS) e bianco in banda  $B$ , con:

- $B = 10$
- $\mu_X = 2$
- $P_X = 24$

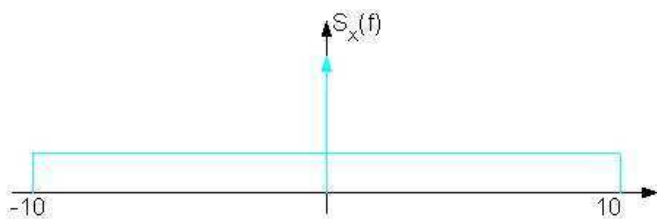
#### Soluzione:

Il processo  $X(t)$  è WSS e gaussiano, il che vuol dire che è anche stazionario in senso stretto (SSS). Per calcolare la funzione di autocovarianza, basta calcolare il valore di  $N_0$ :

$$N_0 = \frac{P_X - \mu_X}{B} = \frac{24 - 4}{10} = 2$$

da cui

$$\frac{N_0}{2} = 1$$



Si ottiene

$$C_X(\tau) = F^{-1} [S_X(f)] = F^{-1} \left[ \text{rect} \left( \frac{f}{20} \right) \right] = 20 \cdot \text{sinc}(20\tau)$$

Da notare che nel calcolo di  $F^{-1} [S_X(f)]$  non bisogna inserire la  $\delta$  del valore continuo, altrimenti si trova la funzione di autocorrelazione  $R_X(\tau)$ .

### Processi ciclostazionari

#### Definizione: Processo ciclostazionario

Un processo  $\{X(t), t \in T\}$  è ciclostazionario quando c'è invarianza alla traslazione periodica.

Un classico esempio di processo ciclostazionario è

$$X(t) = V \sin(t)$$

dove  $V$  è una variabile casuale.



**LAVORI IN CORSO! - WORK IN PROGRESS!**

**Davide Morellato** si sta occupando di questo testo; non apportare modifiche se l'ultima modifica ([http://it.wikiversity.org/w/index.php?title=Elaborazione\\_lineare\\_e\\_nonlineare\\_di\\_un\\_processo\\_stocastico&action=history](http://it.wikiversity.org/w/index.php?title=Elaborazione_lineare_e_nonlineare_di_un_processo_stocastico&action=history)) è recente.

Estratto da "[http://it.wikiversity.org/wiki/Elaborazione\\_lineare\\_e\\_nonlineare\\_di\\_un\\_processo\\_stocastico](http://it.wikiversity.org/wiki/Elaborazione_lineare_e_nonlineare_di_un_processo_stocastico)"

Categorie: [WIP](#) | [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

---

- Ultima modifica per la pagina: 14:28, 17 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Esercizio su Luca e Giorgio

Da Wikiversità, l'università aperta.

## Esercizio: Luca e Giorgio

Due amici, Luca e Giorgio, si danno appuntamento al bar tra le 8 e le 8:30; i due arrivano al bar in modo casuale ed indipendente.

1. Definire lo spazio di probabilità;
2. Calcolare la probabilità che Luca arrivi prima di Giorgio;
3. Sapendo che arrivati al bar si fermano per un tempo pari a  $\Delta T_L = 5$  minuti e  $\Delta T_G = 5$  minuti, calcolare la probabilità che si incontrino.

1. Si ha

$$\Omega = [0, T] \times [0, T]$$

$$F = \{A \subset \mathbb{R} \mid A = B \cap \Omega, B \in \mathbb{B}(\mathbb{R}^2)\}$$

$$f(L, G) = \begin{cases} \frac{1}{T^2} & (L, G) \in [0, T] \times [0, T] \\ 0 & \text{altrimenti} \end{cases}$$

2. Si ha, a causa dell'indipendenza,

$$P(L < G) = \frac{A(L)}{A_{tot}} = \frac{1}{2}$$

Una rappresentazione alternativa può essere data con gli integrali:

$$\int_0^T \int_0^x \frac{1}{T^2} d1 d2 = \frac{1}{2}$$

3. Lasciato per esercizio.

Estratto da "[http://it.wikiversity.org/wiki/Esercizio\\_su\\_Luca\\_e\\_Giorgio](http://it.wikiversity.org/wiki/Esercizio_su_Luca_e_Giorgio)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 23:18, 9 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Esempio del sistema di comunicazione binario simmetrico

Da Wikiversità, l'università aperta.

Questo è un esempio di applicazione della probabilità condizionata.

Si ha un sistema di comunicazione binario simmetrico



dove  $m$  sono i simboli emessi dalla sorgente,  $\hat{m}$  i simboli ricevuti dal destinatario. Nel caso ideale,  $m = \hat{m}$ ; nel caso reale, abbiamo che il messaggio originale  $m$  che transita sul canale è, in generale, alterato a causa di diversi fattori quali la presenza di rumore termico, interferenze con altri sistemi di trasmissione, non idealità<sup>[1]</sup> del canale. Consideriamo il caso in cui il canale può assumere solo due valori:  $m = 0$  ed  $m = 1$  (sorgente binaria).

Lo spazio di probabilità ha come insieme degli esiti

$$F = (0,0),(0,1),(1,0),(1,1)$$

Come al solito, la  $\sigma$ -algebra è l'insieme delle parti,  $|F| = 2^\Omega$ . Si definiscono gli eventi

- $A_0 = (0,0),(0,1) \Rightarrow m = 0$
- $B_0 = (1,0),(1,1) \Rightarrow m = 1$

Le probabilità a priori sono

- $P(m = 0) = q$
- $P(m = 1) = p$

dove  $q + p = 1$ .

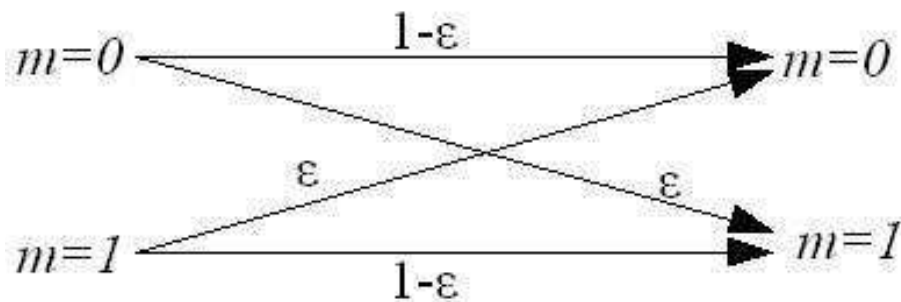
Ipotizziamo che il canale sia simmetrico, cioè la probabilità di crossover<sup>[2]</sup> è

$$P(\hat{m} = 1 | m = 0) = P(\hat{m} = 0 | m = 1) = \epsilon$$

Si ha

- $P((1,0) | A_0) = P((1,1) | A_0) = \epsilon$
- $P((0,1) | A_0) = \epsilon$
- $P((0,0) | A_0) = 1 - \epsilon$  è la probabilità che non vi siano errori, data la trasmissione  $m = 0$
- $P((0,1) | A_1) = P((0,0) | A_1) = \epsilon$
- $P((1,0) | A_1) = \epsilon$
- $P((1,1) | A_1) = 1 - \epsilon$

A questo punto, possiamo ridisegnare il canale usando il modello con le probabilità di transizione



Dal teorema della probabilità totale segue che

$$P((0,0)) = P((0,0)|A_0) \cdot P(A_0) + P((0,0)|A_1) \cdot P(A_1)$$

da cui

- $P((0,0)) = (1 - \epsilon) \cdot q$
- $P((0,1)) = \epsilon \cdot q$
- $P((1,0)) = \epsilon \cdot p$
- $P((1,1)) = (1 - \epsilon) \cdot p$

Calcoliamo la prob di ricevere “0” e “1”; in termini elementari

- $\hat{m} = 0 \rightarrow B_0 = (0,0), (1,0)$
- $\hat{m} = 1 \rightarrow B_1 = (0,1), (1,1)$

Allora, calcoliamo

- $P(B_0) = P((0,0)) + P((1,0)) = 1 - \epsilon \cdot q + \epsilon \cdot p$  (perché i due eventi sono disgiunti)
- $P(B_1) = P((0,1)) + P((1,1)) = \epsilon \cdot q + 1 - \epsilon \cdot p$

La probabilità di errore è data dall'evento errore  $E = (0,1), (1,0)$  è

$$P(E) = P(\{(0,1)\}) + P(\{(1,0)\}) = \epsilon \cdot (p + q) = \epsilon$$

Quindi, in termini di probabilità di errore rispetto alla probabilità dei singoli, per un canale binario simmetrico, l'errore è indipendente dalla distribuzione di probabilità degli ingressi.

$$P(E|\hat{m} = 1)$$

Definiamo l'evento

$$C = (0,1), (1,1) = \hat{m} = 1 = B_1$$

che indica che abbiamo ricevuto 1. Allora

$$P(E|C) = \frac{P(E \cap C)}{P(C)} = \frac{P((0,1))}{P(C)} = \frac{\epsilon \cdot q}{\epsilon \cdot q + (1 - \epsilon) \cdot p}$$

Casi limite:

- $q = 1$ , allora  $P(E|C) = 1$
- $q = 0$ , allora  $P(E|C) = 0$
- $p = q = \frac{1}{2}$ , allora  $P(E|C) = \epsilon$

## Probabilità a posteriori

Dato  $\hat{m} = 1$ , qual'è la probabilità di  $m = 0$  e  $m = 1$ ?

$$\begin{aligned}
 \blacksquare P_{01} &= P(m = 0 | \hat{m} = 1) = \frac{P(m = 0 \cap \hat{m} = 1)}{P(\hat{m} = 1)} = \frac{P((0, 1))}{P((0, 1), (1, 1))} = \frac{\epsilon \cdot q}{\epsilon \cdot q + (1 - \epsilon) \cdot p} \\
 \blacksquare P_{11} &= P(m = 1 | \hat{m} = 1) = \frac{P(m = 1 \cap \hat{m} = 1)}{P(\hat{m} = 1)} = \frac{(1 - \epsilon) \cdot q}{\epsilon \cdot q + (1 - \epsilon) \cdot p} \\
 \blacksquare P_{00} &= \dots \\
 \blacksquare P_{10} &= \dots
 \end{aligned}$$

Se  $\epsilon \ll 1/2$  (canale affidabile), allora

$$P_{11} \simeq \frac{p}{\epsilon \cdot q + p} \text{ e } P_{01} \simeq \frac{p}{\epsilon \cdot q}$$

Questo implica che  $P_{11} \gg P_{01}$ , cioè osservare  $\hat{m} = 1$  aumenta significativamente la probabilità che  $m = 1$ . Notate che per  $\epsilon$  molto piccolo,

$$P_{11} \rightarrow 1$$

dove

$$P_{11} > P(\text{probabilità a priori})$$

cioè la probabilità a posteriori è molto maggiore rispetto alla probabilità a priori. Questo è intuitivo: se così non fosse, allora non servirebbe avere il canale: potremmo dedurre tranquillamente il risultato della trasmissione, meglio che osservando la trasmissione stessa.

La condizione di massima incertezza si ha quando  $\epsilon = 1/2$ , cioè

$$\begin{cases}
 P_{01} = P_{11} \rightarrow \epsilon q = (1 - \epsilon)q \rightarrow \epsilon = p \\
 P_{10} = P_{00} \rightarrow \epsilon p = (1 - \epsilon)p \rightarrow \epsilon = q
 \end{cases}$$

In questo particolare caso è inutile l'osservazione della trasmissione sul canale: basta tirare a caso e si ottiene la stessa probabilità di sbagliare.

## Dipendenza sorgente e ricevitore

In termini di indipendenza statistica, in un canale ideale c'è dipendenza totale tra la sorgente ed il ricevitore. Ad una data realizzazione della sorgente corrisponde una sola realizzazione del ricevitore, che coincide con quella della sorgente. Nel caso di massima incertezza, al contrario, sorgente e ricevitore sono statisticamente indipendenti tra loro; infatti, guardando il risultato del ricevitore non posso dire assolutamente nulla sulla sorgente, e viceversa.

## Note

- ↑ La definizione di canale ideale è un sistema con guadagno 1 in banda passante e a fase lineare, quindi introduce un piccolo ritardo  $T$  ed eventualmente un'attenuazione.
- ↑ La **probabilità di crossover** è la probabilità di sbagliare nel canale binario simmetrico.

Estratto da "[http://it.wikiversity.org/wiki/Esempio\\_del\\_sistema\\_di\\_comunicazione\\_binario\\_simmetrico](http://it.wikiversity.org/wiki/Esempio_del_sistema_di_comunicazione_binario_simmetrico)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 13:11, 10 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

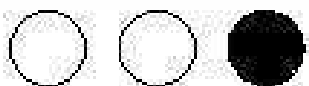
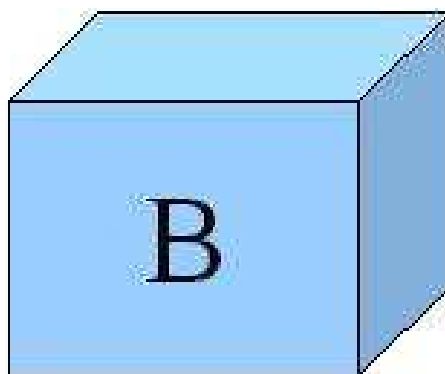
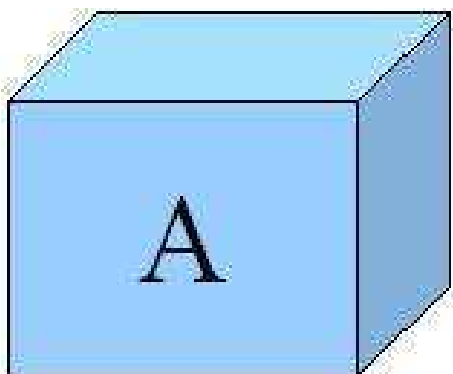
# Esercizio con palline bianche e nere nelle scatole

Da Wikiversità, l'università aperta.

*Esempio: Scatole e palline*

Ci sono due scatole A e B:

- in A ci sono due palline bianche ed una pallina nera;
- in B ci sono una pallina bianca ed una pallina nera.



*Studiare il sistema; qual è la probabilità di pescare una pallina bianca?*

## Soluzione

To do...

Estratto da "[http://it.wikiversity.org/wiki/Esercizio\\_con\\_palline\\_bianche\\_e\\_nere\\_nelle\\_scatole](http://it.wikiversity.org/wiki/Esercizio_con_palline_bianche_e_nere_nelle_scatole)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 12:19, 10 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Esercizio sul codice di Hamming 7 4

Da Wikiversità, l'università aperta.

*Esempio: Codice di Hamming*

Il codice di hamming(7,4) è un codice per la correzione degli errori sui canali binari che rappresenta una parola di 4 bit con una parola di 7 bit. Permette di correggere un errore e rilevare fino ad un massimo di due errori.

Vogliamo sapere la probabilità che la parola sia esatta, con probabilità di errore sul bit singolo  $P_{\varepsilon}(b) = 10^{-3}$ .

Si suppone un canale binario simmetrico indipendente.

Siccome il canale è indipendente, ogni singolo bit è indipendente dall'altro, cioè possiamo usare le prove bernoulliane.

Lo spazio campione è

$$\Omega = \{0,1\}$$

dove 0 indica che non c'è stato errore, 1 indica che c'è stato un errore. Si ha

$$F = \{\emptyset, \{1\}, \{0\}, \{0, 1\}\}$$

e si hanno:

- $P(0) = 1 - 10^{-3}$
- $P(1) = 10^{-3}$

Una parola è errata quando ci sono almeno due errori. L'evento errore è

$$\bullet C = \text{sbaglio almeno 2 bit sulla parola} = \{(c_1, c_2, \dots, c_7)\} \in \hat{F} \text{ t.c. } \sum_i c_i \geq 2$$

mentre l'evento successo è

$$\bullet \bar{C} = \{(c_1, c_2, \dots, c_7)\} \in \hat{F} \text{ t.c. } \sum_i c_i < 2$$

Sappiamo che

$$P_n(k) = \binom{n}{k} p^k q^{n-k}$$

quindi

$$P(C) = \sum_{k=2}^7 P_7(k) = \sum_{k=2}^7 \binom{7}{k} p^k q^{7-k}$$

il che vuol dire che la somma delle probabilità che ci siano  $k = 2,3,4,5,6,7$  errori è la probabilità totale. Vale



$$P(C) = 1 - P(\bar{C}) = 1 - \binom{7}{0} p^0 \cdot q^7 - \binom{7}{1} p^1 \cdot q^6$$

Se la parola non fosse codificata, allora l'unica probabilità di avere la parola giusta sarebbe la totale assenza di errori.

Estratto da "[http://it.wikiversity.org/wiki/Esercizio\\_sul\\_codice\\_di\\_Hamming\\_7\\_4](http://it.wikiversity.org/wiki/Esercizio_sul_codice_di_Hamming_7_4)"

Categorie: [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

---

- Ultima modifica per la pagina: 12:58, 16 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Esercizio del mazzo di carte con 3 carte estratte

Da Wikiversità, l'università aperta.

*Esercizio: Un mazzo di carte*

Prendete un mazzo di 52 carte ed estraete 3 carte in maniera indipendente, con reinserimento (ogni volta ci sono 52 carte).

1. Costruire il modello probabilistico.
2. Determinare la probabilità di pescare esattamente 2 cuori;
3. Determinare la probabilità che almeno una carta sia di cuori.

Per rispondere alla domande, conviene prendere:

$$\Omega = \{\text{Cuori}, \text{Picche}, \text{Fiori}, \text{Quadri}\}$$

Si ha

$$P(\text{Cuori}) = \frac{13}{52} = \frac{1}{4}$$

## Probabilità che due carta siano cuori

Si definisce un nuovo spazio di probabilità  $(\hat{\Omega}, \hat{F}, \hat{P})$  con

$$F = \{\emptyset, A, \bar{A}, \hat{\Omega}\}$$

dove

- $\hat{\Omega} = \Omega \times \Omega \times \Omega$
- $\hat{F} = F \otimes F \otimes F$

Nel caso  $n = 3$  e  $k = 2$ , cioè si hanno esattamente due cuori su tre carte pescate,  $C$  è l'evento:

$$C = \{c_1, c_2, c_3 \in \hat{F} \mid C_i = A \forall i \in I, |I| = 2\}$$

che ha probabilità

$$P(C) = \binom{3}{2} \cdot P^2(A)P(\bar{A}) = 3 \left(\frac{3}{4}\right)^2 \frac{3}{4} = \frac{9}{64}$$

## Probabilità che almeno una carta sia cuori

La probabilità che almeno una carta sia fiori è l'evento  $D$ :

$$D = \{c_1, c_2, c_3 \in \hat{F} \mid C_i = A \forall i \in I, |I| \geq 1\}$$

La probabilità di questo evento è:

$$P(D) = \sum_{i=1}^3 \binom{n}{k} P^k(A) P^{n-k}(\bar{A})$$

Si ha

$$P(\bar{D}) = \binom{n}{0} P^n(\bar{A}) = \binom{3}{0} \left(\frac{3}{4}\right)^3 = \frac{27}{64}$$

da cui si ottiene

$$P(D) = A - P(\bar{D}) = \frac{37}{64}$$

Estratto da "[http://it.wikiversity.org/wiki/Esercizio\\_del\\_mazzo\\_di\\_carte\\_con\\_3\\_carte\\_estratte](http://it.wikiversity.org/wiki/Esercizio_del_mazzo_di_carte_con_3_carte_estratte)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

---

- Ultima modifica per la pagina: 13:53, 16 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Serie di esercizi per il corso di Teoria dei segnali e dei fenomeni aleatori

Da Wikiversità, l'università aperta.

*Esercizio:*

Siano  $A, B, C$  tre eventi di uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ . Dimostrare che

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

*Esercizio:*

Giovanni e marco seguono il corso di teoria dei segnali e dei fenomeni aleatori, il cui esame finale prevede tre punteggi:  $A, B, C$ .

Sapendo che

1. la probabilità che Giovanni prenda  $B$  è  $0,3$
2. la probabilità che Marco prenda  $B$  è  $0,4$
3. la probabilità che nessuno dei due prenda  $A$  ma almeno uno dei due prenda  $B$ , vale  $0,1$

calcolare la probabilità che almeno uno dei due prenda  $B$  ma nessuno prenda  $C$ .

*Esercizio:*

Si lanciano due dadi indipendenti non truccati; calcolare la probabilità che:

1. la somma dei due dadi sia  $s \geq 8$ ;
2. la somma sia  $s = 8$ ;
3. si ottenga almeno un 6 nei due lanci.

*Esercizio:*

Si consideri la legge di probabilità di Bernoulli  $P_n(k)$ . Fissati  $n$  e  $p = P(A)$  la probabilità di successo, determinare il valore di  $k$  per cui si ha la massima probabilità  $P_n(k)$ .

1. calcolare analiticamente il massimo;
2. provare con Octave a verificare il risultato teorico.

*Soluzione:*

Per calcolare il valore giusto di  $k$ , si fa

$$\frac{P_n(k-1)}{P_n(k)} = \frac{\frac{n!}{(n-k+1)!(k-1)!} \cdot p^{k-1} q^{n-k+1}}{\frac{n!}{(n-k)!k!} \cdot p^k q^{n-k}} = \frac{n-q}{(n-k+1)p}$$

Nel caso in cui  $k < (n+1)p$ , si ha

$$\frac{P_n(k-1)}{P_n(k)} = \frac{(n+1)pq}{(n+1)(1-p)p} = 1 \Rightarrow \begin{cases} P_n(k+1) < P_n(k) & k < (n+1)p \\ P_n(k-1) > P_n(k) & k < (n+1)p \end{cases}$$

---

Estratto da "<http://it.wikiversity.org>

[/wiki/Serie\\_di\\_esercizi\\_per\\_il\\_corso\\_di\\_Teoria\\_dei\\_segnali\\_e\\_dei\\_fenomeni\\_aleatori](#)"

Categorie: [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

---

- Ultima modifica per la pagina: 14:21, 16 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Esercizio sulla memoria RAM

Da Wikiversità, l'università aperta.

## Esercizio: *Scatola di RAM*

Si ha una scatola con all'interno 16 banchi di RAM, di cui

- 12 banchi di RAM sono funzionanti;
- 4 banchi di RAM sono guasti.

Si estraggono 3 banchi di RAM a caso, contemporaneamente.

- Descrivere l'esperimento aleatorio associato;
- Determinare la densità di probabilità della variabile casuale  $X$  che rappresenta il numero di moduli funzionanti estratti;
- Implementare l'esperimento in GNU/Octave.

## Esperimento aleatorio

Si ha:

$$\Omega = \{\{G, G, G\}, \{F, F, F\}, \{G, F, F\}, G, G, F\}$$

Con probabilità associate:

- $P(\{G, G, G\}) = P(\{F, F, F\}) = \frac{1}{8}$
- $P(\{G, F, F\}) = P(\{G, G, F\}) = \frac{3}{8}$

dove  $F$  è un modulo funzionante,  $G$  è un modulo guasto. Si ha

- $P(G) = \frac{1}{4}$
- $P(F) = \frac{3}{4}$

## Variabile casuale $X$

Si devono utilizzare le disposizioni per  $n$  elementi in  $k$  gruppi  $3 - 1$ .

$$D_{n,k} = n(n-1) \cdots (n-k-1)$$

Casi possibili sono  $n = 16$  e  $k = 3$ .

Estratto da "[http://it.wikiversity.org/wiki/Esercizio\\_sulla\\_memoria\\_RAM](http://it.wikiversity.org/wiki/Esercizio_sulla_memoria_RAM)"

Categorie: Lezioni | Teoria dei segnali e dei fenomeni aleatori

- Ultima modifica per la pagina: 10:12, 22 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Esercizio al calcolatore sulla trasformazione delle VC

Da Wikiversità, l'università aperta.

Si scriva un generatore di numeri pseudo-casuali, sfruttando l'algoritmo lineare congruente:

$$x_{n+1} = (a \cdot x_n + c) \pmod{m}$$

con  $a, c, m \in \mathbb{N}$ . Si ha che:

1.  $c, a$  devono essere primi tra loro,  $MCD(a, c) = 1$ ;
2.  $a - 1$  deve essere multiplo di ogni fattore primo di  $m$ ;
3.  $a - 1$  deve essere multiplo di 4 se  $m$  è multiplo di 4.

Realizzare un generatore di numeri pseudo-casuali uniforme in  $(0,1)$ , utilizzando

- $a = 7^5$
- $m = 2^{31} - 1$
- $c = 0$

Estratto da "[http://it.wikiversity.org/wiki/Esercizio\\_al\\_calcolatore\\_sulla\\_trasformazione\\_delle\\_VC](http://it.wikiversity.org/wiki/Esercizio_al_calcolatore_sulla_trasformazione_delle_VC)"

---

- Ultima modifica per la pagina: 14:16, 22 ott 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.

# Esercizi sulla trasformazione di variabili casuali

Da Wikiversità, l'università aperta.

*Esercizio:*

Date due variabili casuali  $X$  e  $Y$  con densità spettrale  $U[0,1]$ , calcolare la funzione di densità di probabilità della variabile casuale

$$Z = \frac{X}{Y}$$

Sfruttiamo un metodo chiamato della variabile ausiliaria:

$$\begin{cases} Z = g_1(X, Y) = \frac{X}{Y} \\ W = g_2(X, Y) = X \end{cases} \Rightarrow \begin{cases} Y = g_1^{-1}(Z, W) = WZ \\ X = g_2^{-1}(Z, W) = \dots \end{cases}$$

Si ha

$$J_g = \begin{bmatrix} -\frac{1}{x^2} & \frac{1}{x} \\ 1 & 0 \end{bmatrix}$$

con determinante

$$\det(J_g) = -\frac{1}{x}$$

Si ha

$$f_{ZW} = \frac{f_{XY}(W, ZW)}{|\det(J_g)|} = \frac{2}{\frac{1}{|W|}} = 2 \cdot |W|$$

Il dominio è

- $W \in [-1, 0]$
- $Z \in [-\infty, 0]$

da cui si ottiene

$$Y = (X+1) \cdot \text{rect}\left(X + \frac{1}{2}\right) \Rightarrow Z = \frac{(X+1)}{X} \cdot \text{rect}\left(X + \frac{1}{2}\right) = \frac{W+1}{W} \cdot \text{rect}\left(W + \frac{1}{2}\right)$$

Immagine: TFA esempio X diviso Y variabile ausiliaria.png

Si ha

$$f_X(z) = \int_{-\infty}^{+\infty} f_{ZW}(z, w) dw = \int_{\frac{1}{z-1}}^0 2w \cdot dw$$

siccome

$$Z = \frac{W+1}{1} \Rightarrow w = -\frac{1}{(z-1)^2}$$

allora si ottiene

$$f_X(z) = \left[ \cancel{2} \cdot \frac{w^2}{\cancel{2}} \right]_{\frac{1}{(z-1)^2}}^0 = -\frac{1}{(z-1)^2}$$



Quindi, si ha

$$\left| f_Z(z) = \frac{1}{(z-1)^2} \right|$$

visto che prima si è lasciato indietro un modulo.

**Esercizio:**

Date due variabili casuali  $X$  e  $Y$  con funzioni di densità di probabilità di tipo  $U[0,1] \times U[0,1]$ , determinare la densità di probabilità della variabile casuale

$$Z = \frac{\max(X, Y)}{\min(X, Y)}$$

Per risolvere questo problema ci sono due metodi:

**Metodo 1:**

Si va a calcolare prima  $F_Z$ , per poi derivare  $f_Z$ . Si ha

$$F_Z(z) = P(Z \leq z) = P\left(\frac{\max(X, Y)}{\min(X, Y)} \leq z\right) =_{X, Y \geq 0} P(\max(X, Y) \leq z \cdot \min(X, Y))$$

- se vale che  $Y \geq X$ , allora si ha

$$P(Y \leq_z X) = \begin{cases} 0 & z < 1 \\ \int_0^1 \left( \int_{y/z}^y dx \right) dy = \frac{1}{2} \left( \frac{z-1}{z} \right) & z \geq 1 \end{cases}$$

- se vale, al contrario,  $Y < X$ , allora

$$P(X \leq_z Y) = \begin{cases} 0 & z < 1 \\ \int_0^1 \left( \int_{x/z}^x dy \right) dx = \frac{1}{2} \left( \frac{z-1}{z} \right) & z \geq 1 \end{cases}$$

Da qui si ottiene

$$F_Z(z) = \frac{z-1}{z} \Rightarrow f_Z(z) = \frac{1}{z^2} \quad z \geq 1$$

**Metodo 2:** *Lasciato per esercizio*

La definizione di  $Z$  può essere rivista come

$$\begin{cases} \frac{X}{Y} & X \geq Y \\ \frac{Y}{X} & X < Y \end{cases}$$

quindi si ha

$$f_Z(z) = f_{Z|X \geq Y}(z|x \geq y)P(X \geq Y) + f_{Z|X < Y}(z|x < y)P(X < Y)$$

**Esercizio:** *Condizionamento di variabili casuali*

Ciao!

**LAVORI IN CORSO! - WORK IN PROGRESS!**



**Davide Morellato** si sta occupando di questo testo; non apportare modifiche se l'ultima modifica

([http://it.wikiversity.org/w/index.php?title=Esercizi\\_sulla\\_trasformazione\\_di\\_variabili\\_casuali&action=history](http://it.wikiversity.org/w/index.php?title=Esercizi_sulla_trasformazione_di_variabili_casuali&action=history)) è recente.

Estratto da "[http://it.wikiversity.org/wiki/Esercizi\\_sulla\\_trasformazione\\_di\\_variabili\\_casuali](http://it.wikiversity.org/wiki/Esercizi_sulla_trasformazione_di_variabili_casuali)"

Categorie: [WIP](#) | [Lezioni](#) | [Teoria dei segnali e dei fenomeni aleatori](#)

---

- Ultima modifica per la pagina: 18:25, 13 nov 2008.
- Contenuti soggetti a licenza d'uso GNU Free Documentation License.