Dario Bannò
Informatica - Anno di corso 2004/2005

Appunti di Calcolo delle Probabilità

lezioni del prof. Bologna

(dalle distribuzioni in poi seguire dalle dispense)

CALCOLO COMBINATORIO

- PERMUTAZIONI

In quanti modi si possono ordinare n oggetti => numero di "ordinamenti" o "parole" che si possono formare con n oggetti (o "lettere")

Es/

quante parole si possono formare con le lettere a, b e c?

abc bac cab
acb bca cba => $6 = 3!$

$$n(n-1)(n-2)\ldots 3\cdot 2\cdot 1 = n!$$

- COMBINAZIONI

In quanti modi si possono combinare n elementi presi a k a k

$$\frac{n(n-1)\ldots(n-k+1)}{k!} = \frac{n!}{(n-k)!\,k!} = \underbrace{\binom{n}{k}}_{}$$

COEFFICIENTE BINOMIALE

Binomio di Newton

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Convenzioni e proprietà del coefficiente binomiale:

si pone: $0! = 1$

$$\binom{n}{0} = \binom{n}{n} = 1$$

$$\binom{n}{k} = 0 \quad \text{se} \quad k < 0 \quad \text{oppure} \quad k > n$$

proprietà:

$$\boxed{\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \qquad 1 \leq k \leq n}$$

IDENTITÀ DI TARTAGLIA

(v. triangolo di Tartaglia)

DEFINIZIONI DELLA PROBABILITÀ

- Spazio campionario ed eventi

• **Def** (SPAZIO CAMPIONARIO)

Si definisce SPAZIO CAMPIONARIO l'insieme di tutti i possibili esiti di un esperimento.

Es/ Lancio di due monete

$$S = \{(T,T), (T,C), (C,T), (C,C)\}$$

NB: - Indicheremo lo spazio campionario con S o Ω.
- Lo spazio campionario è detto anche "sample space" o "evento certo"

• **Def** (EVENTO)

Si definisce EVENTO qualsiasi sottoinsieme dello spazio campionario. Viene spesso indicato con una proposizione non ambigua (o un detto casuale).

Es/ Lancio di due monete

$$E = \{(T,T), (T,C)\}$$

è l'evento che si ottenga testa con la prima moneta.

NB: - indicheremo gli eventi con $E_1, E_2, \ldots$ o semplicemente con E.
- gli eventi possono anche chiamarsi eventi casuali o aleatori
- l'evento certo è l'evento che comprende tutti i casi possibili e coincide con lo spazio campionario S

$\emptyset$ è detto evento impossibile e indica un evento che non può mai verificarsi

quando un evento non è né certo, né impossibile è detto possibile

$E \cup F$ è l'evento unione dei 2 eventi E e F

$E \cap F$ è l'evento intersezione dei 2 eventi E e F

Def (EVENTI INCOMPATIBILI)

Dati 2 eventi E e F

se $E \cap F = \emptyset \Rightarrow E, F$ sono detti INCOMPATIBILI (o DISGIUNTI)

Def (EVENTO COMPLEMENTARE)

Si definisce EVENTO COMPLEMENTARE E^c dell'evento E l'insieme di tutti gli esiti dello spazio campionario che non sono contenuti in E.

$$\overline{E} = E^c = S \setminus E$$

Es/ Lancio di 2 monete

se $E = \{(T,C), (T,T)\}$

$\Rightarrow E^c = \{(C,T), (C,C)\}$

NB: $S^c = \emptyset$

- DEFINIZIONE FREQUENTISTICA (o STATISTICA)

La probabilità $P(E)$ di un evento E può essere definita come il rapporto (al limite) del numero di volte in cui si presenta l'evento E col numero di volte in cui l'esperimento viene ripetuto (n).

numero di volte in cui si presenta l'evento E

$$\boxed{P(E) = \lim_{n \to \infty} \frac{n_E}{n}}$$

(Legge empirica del caso)

numero di volte in cui l'esperimento viene ripetuto

$P(E)$ è quindi la frequenza limite di E.

<u>NB</u>: - Tale definizione, sebbene utile in pratica, presenta delle difficoltà concettuali dato che il limite matematico considerato può non esistere (non esiste una dimostrazione matematica).
- Si usa quando l'evento E è ripetibile

- DEFINIZIONE CLASSICA (Laplace (Pascal), 1812)

numero di <u>casi favorevoli</u> all'evento E

$$\boxed{P(E) = \frac{n_E}{N}}$$

numero di <u>casi possibili</u>

Es/ Lancio di un dado

$E =$ "numero pari" $= \{2, 4, 6\} \Rightarrow n_E = 3$

$S = \{1, 2, 3, 4, 5, 6\} \Rightarrow N = 6$

$\Rightarrow P(E) = \frac{n_E}{N} = \frac{3}{6} = \frac{1}{2}$

<u>NB</u>: Questa definizione può essere accettata solo per esperimenti che danno luogo a risultati <u>equiprobabili</u>

– DEFINIZIONE SOGGETTIVISTA

La probabilità soggettiva P di un evento E (quando E non è riconducibile allo schema classico né a quello frequentista) è la misura del grado di fiducia che un soggetto coerente ha nel verificarsi dell'evento E ovvero la quota unitaria di scommessa che un soggetto coerente è disposto a pagare.

Lo strumento che consente di dare una valutazione numerica di tale probabilità è lo SCHEMA DELLA SCOMMESSA

- Schema della scommessa:

  T : puntata dello scommettitore
  K : somma di vincita
  G : guadagno dello scommettitore
  |E| : indicatore (o valore di verità) dell'evento E

$$|E| = \begin{cases} 1 & , \ E \text{ si verifica} \\ 0 & , \ E \text{ non si verifica} \end{cases}$$

→ P : quota unitaria di scommessa (probabilità)

$$\boxed{P = \frac{T}{K}}$$

Si definisce così:

$$G = k|E| - T \quad \text{(guadagno)}$$

da cui: $\begin{cases} G_1 = k - T & , |E| = 1 \\ G_0 = -T & , |E| = 0 \end{cases}$

raccogliendo k a fattore comune:

$$G = k\left(|E| - \frac{T}{k}\right)$$

dato che $P = \frac{T}{k}$

$$\Rightarrow G = k(|E| - P)$$

da cui: $\begin{cases} G_1 = k(1-P) & , |E| = 1 \\ G_0 = -kP & , |E| = 0 \end{cases}$

NOTA:
- se $P < 0 \Rightarrow G_0 > 0, G_1 > 0$ (vincita certa)
- se $P > 1 \Rightarrow G_0 < 0, G_1 < 0$ (perdita certa)

Questi 2 casi non sono coerenti.

$$\Rightarrow \boxed{0 \leq P \leq 1}$$

Condizione di coerenza

⇒ Impostazione e DEFINIZIONE ASSIOMATICA

– Struttura algebrica degli eventi: $\mathcal{A}$ L'algebra di Boole

Dato uno spazio campionario Ω, si definisce evento un sottoinsieme A di Ω che appartiene all'insieme $\mathcal{A}$ di tutti i sottoinsiemi di Ω che soddisfano le seguenti condizioni:

1) $A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$ (complemento)

2) $A \in \mathcal{A};\ B \in \mathcal{A} \Rightarrow (A \cup B) \in \mathcal{A}$ (unione)

3) $\Omega \in \mathcal{A} \Rightarrow \bar{\Omega} = \emptyset \in \mathcal{A}$ (insieme vuoto)

$\mathcal{A}$ è quindi la struttura algebrica degli eventi e coincide con l'insieme delle parti di Ω.

$\mathcal{A}$ (detto "algebra di Boole") è un insieme chiuso rispetto alle operazioni di unione e negazione ed è una rappresentazione che permette di trattare gli eventi con lo strumento matematico dell'algebra degli insiemi.

Es/ "Lancio di una moneta"

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

$$\mathcal{A} = \mathcal{P}(\Omega) = \{\emptyset, \{T\}, \{C\}, \Omega\}$$

Se $|\Omega| = n$
$\Rightarrow |\mathcal{A}| = |\mathcal{P}(\Omega)| = 2^n$

- Proprietà e conseguenze di $\mathcal{A}$

- Chiusura rispetto all'unione finita

Dati Ω, $\mathcal{A}$

se $A_1, A_2, A_3 \in \mathcal{A}$

$\Rightarrow A = (A_1 \cup A_2) \in \mathcal{A}$

e $(A \cup A_3) \in \mathcal{A}$

$\Rightarrow (A_1 \cup A_2 \cup A_3) \in \mathcal{A}$

Per induzione:

$$\boxed{\begin{array}{ll} \text{Se } A_i \in \mathcal{A} & i = 1, 2, \ldots, n \\ \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A} & \end{array}}$$

Chiusura di $\mathcal{A}$ rispetto all'unione finita

- Proposizione

Se $\mathcal{A}$ è chiusa anche rispetto all'UNIONE INFINITA, ovvero se:

$A_i \in \mathcal{A} \qquad i = 1, 2, \ldots$

$\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

allora $\mathcal{A}$ è una σ-algebra (sigma algebra)
(o una classe σ-additiva o anche numerabilmente additiva)

NB: una σ-algebra è un'algebra ma non vale necessariamente il contrario

– Proprietà delle operazioni di unione ($\cup$), intersezione ($\cap$) e complemento ($\bar{\cdot}$) fra insiemi:

1. PROPRIETA' COMMUTATIVA

$\cup$) $A \cup B = B \cup A$

$\cap$) $A \cap B = B \cap A$

2. PROPRIETA' ASSOCIATIVA

$\cup$) $A \cup (B \cup C) = (A \cup B) \cup C$

$\cap$) $A \cap (B \cap C) = (A \cap B) \cap C$

3. PROPRIETA' DISTRIBUTIVA

$\cup$ rispetto a $\cap$) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

$\cap$ rispetto a $\cup$) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

4. Doppio complemento

$$\overline{(\bar{A})} = A$$

5. Operazioni con Ω

$\cup$) $A \cup \Omega = \Omega$

$\cap$) $A \cap \Omega = A$ (Ω è neutro per $\cap$)

6. Operazioni con ϕ

$\cup$) $A \cup \phi = A$ (ϕ è neutro per $\cup$)

$\cap$) $A \cap \phi = \phi$

7. Operazioni con il complementare

$\cup$) $A \cup \bar{A} = \Omega$

$\cap$) $A \cap \bar{A} = \phi$

8. Operazioni

$\cup$) $A \cup A = A$

$\cap$) $A \cap A = A$

- Proprietà e tipologia degli eventi

- Evento:

Ogni sottoinsieme $A \in \mathcal{A}$ di Ω è un evento.
Si distinguono fra questi:
- Ω : evento certo
- $\emptyset$: evento impossibile

- Evento POSSIBILE:

Per evento possibile si intende ogni sottoinsieme proprio di Ω non nullo, ovvero:

a) $A \subset \Omega$ (sottoinsieme proprio)
che implica $A \neq \Omega$
cioè A non è un evento certo

b) $A \neq \emptyset$
cioè A non è un evento impossibile

- Eventi INCOMPATIBILI:

Due eventi si dicono incompatibili se il loro prodotto (ovvero la loro intersezione) da luogo all'evento impossibile:

$$A \cap B = \emptyset$$

In generale

n eventi sono incompatibili se sono a 2 a 2 incompatibili, ovvero:

$$\boxed{A_i \cap A_j = \emptyset \,, \; i \neq j \,, \; \forall (i,j) \quad \begin{array}{l} i = 1,2,\ldots,n \\ j = 1,2,\ldots,n \end{array}}$$

- Eventi NECESSARI:

n eventi sono necessari se:

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

Nota: più eventi necessari e incompatibili costituiscono una partizione di Ω

Es/ la più semplice partizione di Ω è costituita da E, $\overline{E}$

$E \cup \overline{E} = \Omega \rightarrow$ eventi necessari

$E \cap \overline{E} = \emptyset \rightarrow$ eventi incompatibili

- Eventi ELEMENTARI:

Si dicono eventi elementari quei sottoinsiemi di Ω costituiti dai singoli elementi di Ω

Nota: essi costituiscono una partizione di Ω (sono necessari e incompatibili)

Es/ "Lancio di un dado"

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

Gli eventi elementari sono:

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

– CLASSE DI BOREL: $\mathcal{B}$

Dato uno spazio campionario Ω si definisce classe di Borel la più piccola σ-algebra definita su Ω.

- DEFINIZIONE ASSIOMATICA (Kolmogorov, 1933)

Dati Ω e $\mathcal{A}$ (σ-algebra),

si definisce funzione di probabilità $P(\cdot)$ una funzione definita sull'insieme $\mathcal{A}$ e a valori in $\mathbb{R}$

$$P(\cdot) : \mathcal{A} \to \mathbb{R}$$

che soddisfa le seguenti condizioni (o assiomi):

1. $P(A) \geq 0 \quad , \quad \forall A \in \mathcal{A} \quad (0 \leq P(A) \leq 1)$

2. $P(\Omega) = 1$

3. $$\left[\begin{array}{l} \text{se} \quad (A_h \cap A_j) = \emptyset \quad , \; h \neq j \qquad \forall (h, j) \\ \Rightarrow \quad P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P(A_i) \qquad \forall \{A_i\} \end{array}\right]$$

successione di eventi

condizione di additività completa
o di σ-additività
(additività infinita)

<u>Nota</u>: - $P(\cdot)$ è una misura, ovvero una funzione definita su una σ-algebra (in questo caso $\mathcal{A}$)

- $(\Omega, \mathcal{A})$ è definito spazio misurabile (o probabilizzabile)

- $(\Omega, \mathcal{A}, P(\cdot))$ è detto spazio di probabilità

- Alcune semplici proprietà

a) $P(\emptyset) = 0$

dim/

$$\Omega \cap \emptyset = \emptyset \quad , \quad \Omega \cup \emptyset = \Omega$$

$$P(\Omega) = P(\Omega) + P(\emptyset) \quad \text{ma} \quad P(\Omega) = 1$$

$$\Rightarrow 1 = 1 + P(\emptyset)$$

$$\Rightarrow P(\emptyset) = 0 \quad \text{cvd}$$

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

dim/

$$\Omega = A \cup \bar{A}$$

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

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

$$\Rightarrow P(\bar{A}) = 1 - P(A) \quad \text{cvd}$$

NB: $A \cap \bar{A} = \emptyset$

A e $\bar{A}$ sono incompatibili, per questo motivo $P(\Omega)$ si può esprimere come somma di $P(A)$ e $P(\bar{A})$

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

dim/

$$A \cup B = A \cup (\bar{A} \cap B) \quad \text{e} \quad A \cap (\bar{A} \cap B) = \emptyset$$

sono incompatibili A e $\bar{A} \cap B$

$$\Rightarrow P(A \cup B) = P(A) + P(\bar{A} \cap B) \quad *1$$

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

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

da cui:

$$P(\bar{A} \cap B) = P(B) - P(A \cap B) \quad *2$$

Unendo *1 e *2 otteniamo così:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{cvd}$$

Probabilità totale per <ins>3 eventi compatibili</ins>

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) +$$
$$- P(E_1 \cap E_2) - P(E_2 \cap E_3) - P(E_1 \cap E_3)$$
$$+ P(E_1 \cap E_2 \cap E_3)$$

dim/ dalla c.1 considerando $G = E_2 \cup E_3$

$$P(E_1 \cup E_2 \cup E_3) = P(E_1 \cup G) = P(E_1) + P(G) +$$
$$- P(E_1 \cap G)$$

$*1 \quad P(G) = P(E_2 \cup E_3) = P(E_2) + P(E_3) - P(E_2 \cap E_3)$

$*2 \quad P(E_1 \cap G) = P(E_1 \cap (E_2 \cup E_3)) =$
$$= P((E_1 \cap E_2) \cup (E_1 \cap E_3)) =$$
$$= P(E_1 \cap E_2) + P(E_1 \cap E_3) - P(E_1 \cap E_2 \cap E_3)$$

Sostituendo dalla $*1$ e $*2$ otteniamo:

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + \underbrace{P(E_2) + P(E_3) - P(E_2 \cap E_3)}_{P(G)} +$$

$$\underbrace{- P(E_1 \cap E_2) - P(E_1 \cap E_3) + P(E_1 \cap E_2 \cap E_3)}_{- P(E_1 \cap G)}$$

cvd

- Generalizzando ad n qualsiasi

$$P\left(\bigcup_{h=1}^{n} E_h\right) = \sum_{h=1}^{n} P(E_h) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - \dots + (-1)^{n-1} P(E_1 \cap E_2 \cap \dots \cap E_n)$$

PRINCIPIO DI INCLUSIONE/ESCLUSIONE

E' la probabilità che si verifichi almeno uno degli **n** eventi (n finito)

PROBABILITA' CONDIZIONATE E INDIPENDENZA

- PROBABILITA' CONDIZIONATA

Si definisce probabilita' condizionata dell'evento E dato l'evento F $(E|F)^*$ la quantita':

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \quad , P(F) \neq 0$$

<u>Nota</u>: $P(E|F)$ ha la stessa struttura delle normali probabilita', ovvero dato che si suppone F gia' verificato (dato):

=> $P(F)$ è il numero di <u>casi possibili</u> e $P(E \cap F)$ il numero di <u>casi favorevoli</u>

$$P(E|F) = \frac{P(E \cap F) \leftarrow \text{casi favorevoli}}{P(F) \leftarrow \text{casi possibili}}$$

Puo' risultare spesso molto utile calcolare la probabilita' condizionata riducendo lo spazio campionario <u>quando tutti gli esiti sono equiprobabili</u>

*: $(E|F)$: "E condizionato ad F" o "E dato F"

- LEGGE DELLE PROBABILITÀ COMPOSTE

- Dalla formula della probabilità condizionata:

$$\boxed{P(E\cap F) = P(F)\,P(E|F) = P(E)\cdot P(F|E)}$$

È la legge delle probabilità composte per 2 eventi.

- Per 3 eventi A_1, A_2, A_3 basta applicare la precedente per $A = A_1\cap A_2$ e A_3

$$P(A_1\cap A_2\cap A_3) = P(A\cap A_3) = P(A)\cdot P(A_3|A) =$$
$$= P(A_1\cap A_2)\cdot P(A_3|A_1\cap A_2) =$$
$$= P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2)$$

- Generalizzando a n eventi: $A_1\cap A_2\cap \dots \cap A_n$ (n finito)

$$\boxed{P(A_1\cap A_2\cap \dots\cap A_n) = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2)\cdot \dots \cdot P(A_n|A_1\cap A_2\cap\dots\cap A_{n-1})}$$

Legge delle probabilità composte per n eventi

- INDIPENDENZA STOCASTICA

Due eventi E e F si dicono indipendenti (o stocasticamente indipendenti) quando il verificarsi di uno non influisca sulla probabilità dell'altro, ovvero quando:

$$P(E|F) = P(E)$$

Dalla formula della probabilità condizionata:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \quad , \quad P(F) \neq 0$$

possiamo dire che E e F sono stocasticamente indipendenti se:

$$\boxed{P(E \cap F) = P(E) \cdot P(F)}$$

condizione di indipendenza stocastica

NB: - questa condizione è valida anche per $P(E) = 0$ e $P(F) = 0$

- quando non valgono le condizioni sopra elencate allora i due eventi E e F si dicono "dipendenti"

- Per 3 eventi) A_1, A_2, A_3 sono indipendenti se:

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

e

$$\begin{cases} P(A_1 \cap A_2) = P(A_1) \cdot P(A_2) \\ P(A_2 \cap A_3) = P(A_2) \cdot P(A_3) \\ P(A_1 \cap A_3) = P(A_1) \cdot P(A_3) \end{cases}$$

In generale:

n eventi $A_1, A_2, \_\_\_, A_n$ si dicono stocasticamente indipendenti se:

$$\boxed{\begin{gathered} P(A_1 \cap A_2 \cap \_\_\_ \cap A_k) = P(A_1) \cdot P(A_2) \cdot \_\_\_ \cdot P(A_k) \\ \forall k \mid 2 \leq k \leq n \\ (n \text{ finito o infinito}) \end{gathered}}$$

Osservazione: INCOMPATIBILITA' E INDIPENDENZA

Due o più eventi incompatibili non nulli non sono stocasticamente indipendenti.

dim/

Infatti se $A \cap B = \emptyset$ ← eventi incompatibili

$\Rightarrow P(A \cap B) = 0$

ma se $A, B \neq \emptyset \Rightarrow P(A) \neq 0, P(B) \neq 0$

$(\Rightarrow 0 < P(A) \leq 1$ e $0 < P(B) \leq 1)$

per cui: $P(A) \cdot P(B) \neq 0$

$\Rightarrow P(A \cap B) \neq P(A) \cdot P(B)$

Nota: L'incompatibilità è una proprietà intrinseca degli eventi, l'indipendenza invece no (dato che deve rispettare particolari condizioni).

- TEOREMA DELLE PROBABILITA' TOTALI

Dato uno spazio di probabilità $(\Omega, \mathcal{A}, P)$ se n eventi incompatibili formano una partizione di Ω, allora la probabilità di qualsiasi evento E si può calcolare come la sommatoria delle probabilità composte dell'evento E e degli eventi B_i, ovvero:

dati $B_1, B_2, \_\_\_, B_n \in \mathcal{A}$

se $\forall i,j \quad i \neq j \quad \underline{B_i \cap B_j = \emptyset}$ ← eventi incompatibili

e $\underline{\bigcup_{i=1}^{n} B_i = \Omega}$ ← partizione di Ω

$$\Rightarrow \boxed{P(A) = \sum_{i=1}^{n} P(B_i) \cdot P(A \mid B_i)}$$

dim/

$$A = A \cap \Omega \quad \text{e dato che} \quad \Omega = \bigcup_{i=1}^{n} B_i$$

$(B_i \cap A) \cap (B_j \cap A)$

$$\Rightarrow A = A \cap \bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} (A \cap B_i)$$

Dato che gli eventi B_i sono incompatibili, anche le singole intersezioni con A saranno incompatibili, ovvero:

$\forall i,j \;, i \neq j \quad$ se $B_i \cap B_j = \emptyset$

$$\Rightarrow \emptyset = A \cap \emptyset = A \cap (B_i \cap B_j) = (B_i \cap A) \cap (B_j \cap A)$$

$$\Rightarrow \underline{(B_i \cap A) \cap (B_j \cap A) = \emptyset}$$

tutti i $(B_i \cap A)$ eventi sono incompatibili

Quindi possiamo scrivere

$$P(A) = P\left[\bigcup_{i=1}^{n} (\underbrace{A \cap B_i}_{})\right] = P\left[\bigcup_{i=1}^{n} (\underbrace{B_i \cap A}_{})\right]$$

per le proprietà commutativa sono uguali

ma dato che i $(B_i \cap A)$ sono incompatibili

$$P(A) = P\left[\bigcup_{i=1}^{n} (B_i \cap A)\right] = \sum_{i=1}^{n} P(B_i \cap A) =$$

$$= \underline{\sum_{i=1}^{n} P(B_i) \cdot P(A|B_i)} \left(= \sum_{i=1}^{n} P(A) \cdot P(B_i/A) \right)$$

per la legge delle probabilità composte

- TEOREMA DI BAYES

Dalla legge della probabilità condizionata abbiamo visto che:

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

Se consideriamo n eventi $B_1, B_2, \_\_\_, B_n$ incompatibili e che costituiscono una partizione di Ω, per il teorema della probabilità totale avremo:

$$P(A) = \sum_{i=1}^{n} P(B_i) \cdot P(A|B_i)$$

$\Rightarrow$ per ogni evento B_k, $0 \leq k \leq n$

$$\boxed{P(B_k|A) = \frac{P(B_k) \cdot P(A|B_k)}{\sum_{i=1}^{n} P(B_i) \cdot P(A|B_i)}}$$

Nota: tale formula può applicarsi anche ad eventi B_i incompatibili e tali che $A \subset \bigcup_{i=1}^{n} B_i$ ma che non sono necessariamente una partizione di Ω

VARIABILI ALEATORIE

– VARIABILE ALEATORIA (V.A.) SEMPLICE

Dato un esperimento $\mathcal{E}$, si definisce variabile aleatoria o casuale (v.A. o v.c.) una funzione X del risultato ω dell'esperimento definita sullo spazio campionario Ω e a valori reali:

$$X(\omega): \Omega \to \mathbb{R}$$

[quindi X associa ad ogni punto $\omega \in \Omega$ un valore reale $X(\omega)$

X, Y, Z → variabili aleatorie

x, y, z → valori reali]

tale che

la sua immagine inversa $X^{-1}: \mathbb{R} \to \Omega$ dell'intervallo $(-\infty, x]$ (con x variabile) appartenga ad $\mathcal{A}$ (σ-algebra) per ogni x di $\mathbb{R}$, ovvero:

$$\boxed{X^{-1}(-\infty, x] \in \mathcal{A}, \quad \forall x \in \mathbb{R}}$$

condizione di misurabilità

nonché definendo $B_x = (-\infty, x]$

e $A_x = \{\omega : X(\omega) \leq x\}$

A_x sono tutti i risultati per cui $X(\omega) \leq x$

$$\Rightarrow \underline{X^{-1}(B_x) = A_x \in \mathcal{A}}$$

$X^{-1}(B_x)$ ovvero A_x è un evento

Nota: – $X^{-1}(-\infty, x]$ è un evento di $\mathcal{A}$

– (condizione di misurabilità di X)

una funzione $X: \Omega \to \mathbb{R}$ è detta misurabile se $\forall x \in \mathbb{R}$ $X^{-1}(-\infty, x] \in \mathcal{A}$ (σ-algebra)

— FUNZIONE DI RIPARTIZIONE (o CUMULATIVA)

Data una variabile aleatoria X, si voglia conoscere la probabilita' che X assuma valori minori o uguali di un numero reale x.

Considerato quindi l'evento $(X \le x)$ si definisce funzione di ripartizione di x la probabilita' dell'evento $(X \le x)$, ovvero:

$$F(x) = P(X \le x) = P(A_x)$$

$$\text{con } A_x = \{\omega \in \Omega : X(\omega) \le x\} \ (\text{evento})$$

• Proprieta' di $F(x)$

1a) $F(x)$ è NON DECRESCENTE

1b) $\Rightarrow P(x_1 < X \le x_2) = F(x_2) - F(x_1)$

2a) $\lim_{x \to -\infty} F(X) = 0$

2b) $\lim_{x \to +\infty} F(X) = 1$

3a) $F(x)$ è continua a destra di $\forall x$: $\lim_{h \to 0^+} F(x+h) = F(x)$

3b) $F(x)$ non è necessariamente continua a sinistra di $\forall x$: $\lim_{h \to 0^+} [F(x) - F(x-h)] \ge 0$

dim/

1a) $x_1 < x_2 \Rightarrow F(x_1) \le F(x_2)$

$$(-\infty, x_2] = (-\infty, x_1] \cup (x_1, x_2]$$

$$P(X \le x_2) = P(X \le x_1) + P(x_1 < X \le x_2)$$

$$F(x_2) = F(x_1) + P(x_1 < X \le x_2) \ge F(x_1)$$

(continua a pag. 38)

- VARIABILE ALEATORIA DISCRETA

Una variabile aleatoria si dice discreta se assume i valori di un insieme finito o al più infinito numerabile di valori reali:

$$X : x_1, x_2, \ldots, x_n, \ldots$$

ovvero:

se $\exists$ un insieme S finito o al più infinito numerabile con probabilità uguale a 1, cioè:

$$\exists\, S = \{ x_i \,;\, i = 1, 2, \ldots, n, \ldots \} : P(x \in S) = 1$$

(Es/ distribuzione uniforme, distribuzione di Bernoulli, ...)

- FUNZIONE DI PROBABILITA' (O DI MASSA)

Una funzione reale $p(x)$, $x \in \mathbb{R}$ è detta funzione di probabilità (o di massa) se:

$$p(x) \begin{cases} > 0 & , x = x_i \quad (i = 1, 2, \ldots, n, \ldots) \\ = 0 & , x \neq x_i \end{cases}$$

e $\sum_i p(x_i) = 1$

Nota:

se consideriamo 1 v.a. X discreta, la funzione di probabilità sarà:

$$p(x) = \begin{cases} P(X = x_i) & , x = x_i \\ 0 & , x \neq x_i \end{cases} \qquad (i = 1, 2, \ldots, n, \ldots)$$

– FUNZIONE INDICATRICE D'INSIEME

Dato un insieme $\Omega = \{\omega_1, \omega_2, \_\_, \omega_n\}$ ed un sottoinsieme A di Ω dicesi funzione indicatrice dell'insieme $A \subset \Omega$ la funzione:

$$\boxed{I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in \bar{A} \end{cases}} \quad , A \subset \Omega$$

Se $A \in \mathcal{A}$ (evento) $\Rightarrow I_A(\omega)$ è una V.A. che prende il nome di V.A. indicatrice di evento.

– V.A. INDICATRICE DI EVENTO

Sia $\mathcal{E}$ un esperimento casuale, Ω l'insieme dei risultati possibili, $\mathcal{A}$ una σ-algebra associata ad Ω, A un evento di Ω (cioè $A \subset \Omega$, $A \in \mathcal{A}$), allora si definisce V.A. indicatrice di evento la funzione:

$$\boxed{I_A(\omega) = x = \begin{cases} 1, \ \omega \in A & ; P(A) = P(x=1) = p \\ 0, \ \omega \in \bar{A} & ; P(\bar{A}) = P(x=0) = 1-p \end{cases}}$$

– ESPERIMENTO BERNOULLIANO

Dicesi esperimento bernoulliano un esperimento i cui risultati sono classificati in:

s: successo ; f: insuccesso

$$\Omega = \{s, f\} = \{\omega_1, \omega_2\}$$

Si definisca X t.c.: $X(\omega_1) = X(s) = 1$

$X(\omega_2) = X(f) = 0$

cioè $X = \begin{cases} 1, & \text{se è vero } S = \{s\} \\ 0, & \text{se è vero } \bar{S} = \{f\} \end{cases}$

Se S ha probabilità $P(S) = p$ (e quindi $P(\bar{S}) = 1 - p$) allora:

$$P(X=1) = p \quad ; \quad P(X=0) = 1 - p$$

$$\Rightarrow \boxed{X = \begin{cases} 1, & P(X=1) = P(S) = p \\ 0, & P(X=0) = P(\bar{S}) = 1 - p \end{cases}}$$

V. A. DI BERNOULLI

Nota: X assume due valori 1,0 quindi si dice V.C. binaria o dicotomica

— DISTRIBUZIONE DI BERNOULLI

Dato un esperimento bernoulliano E con $\Omega = \{s, f\}$ si dice che una v.c. X ha una distribuzione di bernoulli se:

$$X = \begin{cases} 1 & \text{se è vero } S = \{s\} \\ 0 & \text{se è vero } \bar{S} = \{f\} \end{cases}$$

$$\text{con} \quad P(S) = P(X=1) = p$$
$$\text{e} \quad P(\bar{S}) = P(X=0) = 1-p$$

X ha le seguenti proprietà:

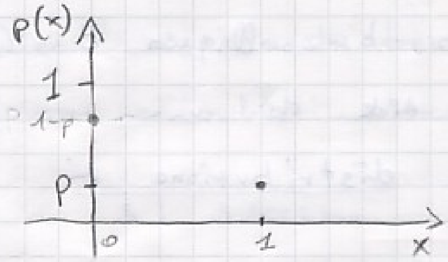
1) X assume i valori $x_1 = 1$ e $x_2 = 0$
2) gli eventi $x=1$ e $x=0$ sono necessari e incompatibili
3) $\sum_{i=1,2} p(x_i) = 1$

questo può essere espresso tramite la funzione di probabilità di X:

$$p(x) = p^x (1-p)^{1-x} \quad , \quad x = 0,1 \quad , \quad 0 \leq p \leq 1$$

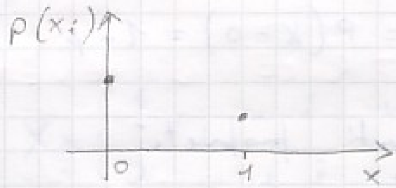
Quest'ultima rappresenta il modello probabilistico che descrive l'esperimento bernoulliano

• Rappresentazioni grafiche

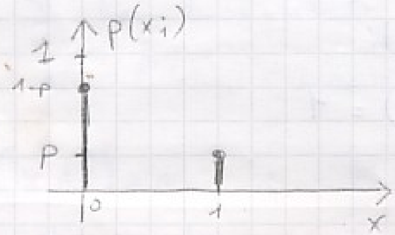


p(x) funzione di probabilità di X, v.A. avente una distribuzione di bernoulli

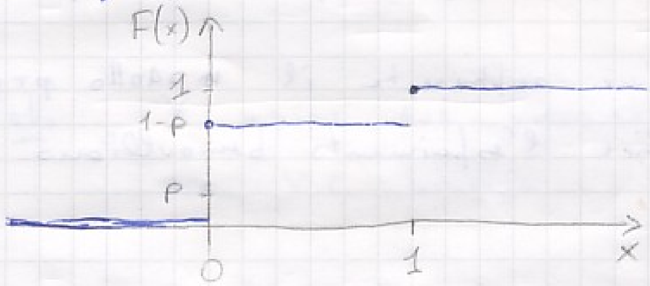
I punti $x_i = 0, 1$ con probabilità $p(x_i) \neq 0$ vengono chiamati punti di massa e spesso la rappresentazione grafica è relativa solo ad essi:



oppure mediante un diagramma a "bastoni"



La funzione di ripartizione di una v.A. di bernoulli ha grafico:

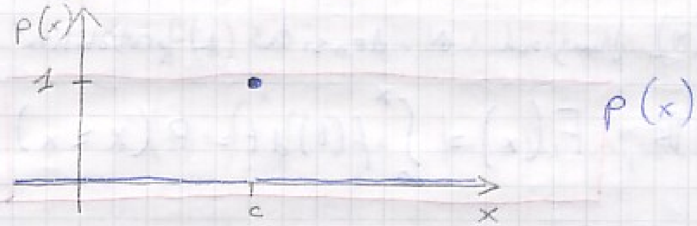


detto "grafico a gradini"

- VARIABILE ALEATORIA DEGENERE

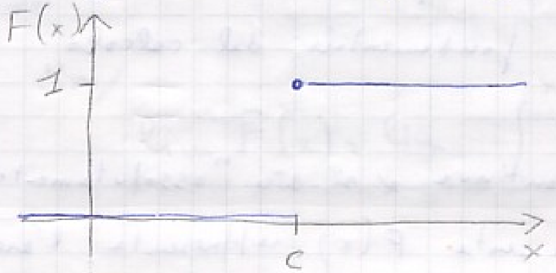
Una variabile aleatoria X viene detta degenere quando assume un solo valore c $(X=c)$ con $P(X=c)=1$.

- Funzione di probabilita'



$$p(x)=\begin{cases}1 & , \; x=c \\ 0 & , \; x\neq c\end{cases}$$

- Funzione di ripartizione



– VARIABILE ALEATORIA CONTINUA –

Una variabile aleatoria si dice continua quando può assumere tutti i valori di un intervallo di $\mathbb{R}$. (aperto, chiuso, blurintervallo, etc...), ovvero:

X V.A. CONTINUA se

$\exists f(\cdot)$ funzione di densità di probabilità t.c.

$$\forall x \in \mathbb{R}, \ \boxed{F(x) = \int_{-\infty}^{x} f(t)\,dt = P(X \leq x)}$$ integrale di probabilità di X

Nota: - $F(x)$ è la funzione di ripartizione di X e non potendo esprimersi mediante una somma, essa è invece una funzione integrale

$$F'(x) = f(x)$$

dal teorema fondamentale del calcolo integrale.

- $F(x)$ è continua e si dice "assolutamente continua"
- geometricamente $F(x)$ rappresenta l'area sottesa alla curva di $f(x)$ nell'intervallo $[-\infty, x]$
- la funzione $f(\cdot)$ è detta funzione di densità di probabilità e ha significato di probabilità per la V.A. X (ma non è comunque una probabilità)
- se X v.c. continua => ogni punto $X = x$ ha probabilità 0, cioè: $P(X = x) = 0$

dim/ $P(x < X \leq x+h) = F(x+h) - F(x) = \int_{x}^{x+h} f(t)\,dt = P(x \leq X \leq x+h)$

$\Rightarrow P(X = x) = 0$

- Proprietà di $f(\cdot)$

Una funzione di densità di probabilità $f(\cdot)$ soddisfa le seguenti condizioni:

a) $f(x) \geq 0$

dim/

se $f(x) < 0$ su un dato intervallo allora

$F(x) = \int_{-\infty}^{x} f(x)\,dx$ potrebbe risultare negativa

e ciò sarebbe impossibile.

In linea di principio potrebbe essere $f(x) < 0$ per un insieme di valori di x di misura nulla.

b) $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ (condizione di normalizzazione)

dim/

$$\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{+\infty} f(x)\,dx = 1$$

- NOTA: distinzione tra le V.A.

- V.A. X può assumere qualsiasi valore reale (infinito) (non ha restrizioni)
- V.A. DISCRETA invece può assumere solo un numero finito o infinito numerabile di valori
- V.A. CONTINUA può assumere ogni valore reale in un dato intervallo.

- VALORE MEDIO (o ATTESO)

Il valore medio o "speranza matematica" di una v.A. X è la media pesata di tutti i possibili valori che X può assumere, ognuno pesato con la probabilità che X lo assuma ed è definito da

$$\boxed{E(X) = \begin{cases} \sum x_i \, p(x_i) & \text{per V.A. DISCRETE} \\ \int_{-\infty}^{+\infty} x \cdot f(x) \, dx & \text{per V.A. CONTINUE} \end{cases}}$$

Nota: - X è nota se è nota la sua distribuzione di probabilità data da

$$\left(F(x), \; p(x), \; f(x) \right)$$

f.d. ripartizione — f.d. probabilità — f.d. densità

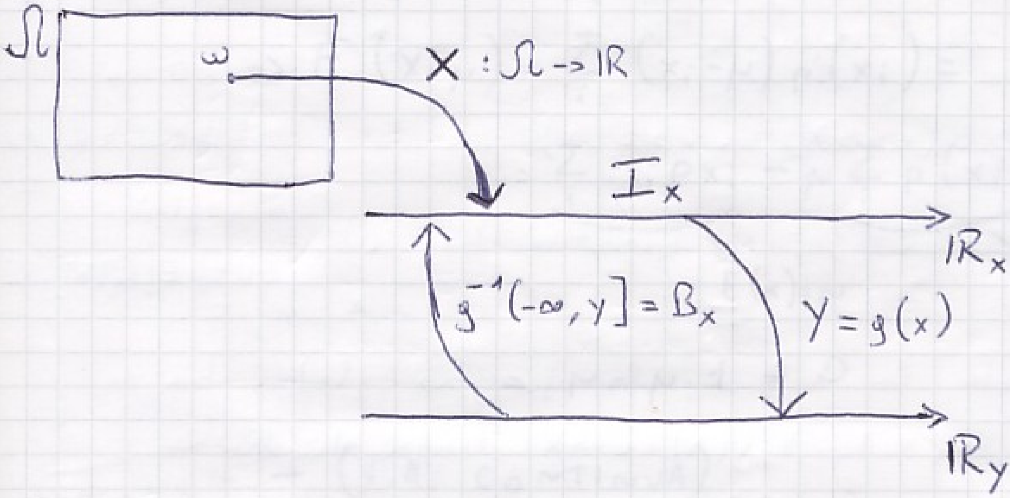
- $E[g(x)]$ rappresenta i valori di sintesi di una distribuzione

— V.A. FUNZIONE DI V.A.: $Y = g(x)$ —

Data una V.A. X $(x : \Omega \to \mathbb{R})$ e una funzione $g(x) : I_x \to \mathbb{R}$ con I_x insieme dei valori reali assunti da X, $Y = g(x)$ è una variabile aleatoria funzione di X (v.A.) se:

$$g^{-1}(-\infty, y] = B_x \in \mathcal{B} \text{ (boreliano)}$$

(cioè se soddisfa la condizione di misurabilità)



Nota: si ricorda che il boreliano B_x (vedi anche classe di Borel pag. 10) è quell'intervallo $(-\infty, x]$ t.c.

$$X^{-1}(-\infty, x] \in \mathcal{A}, \quad \forall x \in \mathbb{R}$$

che è la condizione di misurabilità della v.A. X

— VALORE MEDIO DI $Y = g(x)$

- se X è una V.A. DISCRETA con $p(x)$

$$E(Y) = \sum_j y_j \, p(y_j)$$

ma vale anche

$$E(Y) = E[g(x)] = \sum_i g(x_i) \, p(x_i)$$

- se X è una V.A. CONTINUA con $f(x)$

$$E(Y) = E[g(x)] = \int_{-\infty}^{+\infty} g(x) f(x) \, dx$$

- PROPRIETA' DEL VALORE MEDIO $E(X)$

a) Data la V.A. X con $E(X) = \mu$ e la funzione reale $g(x) = x - \mu$ che definisce la V.A. trasformata $Y = g(x) = x - \mu$, allora: $E(X - \mu) = 0$

ovvero:

il valore medio della V.A. SCARTO da μ è nullo.

dim/ - (V.A. DISCRETA)

$$E[g(x)] = \sum_i g(x_i)\,p(x_i)$$

$$\Rightarrow E(X-\mu) = \sum_i (x_i - \mu)\,p(x_i) =$$

$$= \underbrace{\sum_i x_i p x_i}_{E(x)=\mu} - \mu \underbrace{\sum_i p(x_i)}_{1} =$$

$$= \mu - \mu \cdot 1 = 0$$

- (V.A. CONTINUA)

$$E(x) = \int_{-\infty}^{+\infty} x f(x)\,dx$$

$$\Rightarrow E(X-\mu) = \int_{-\infty}^{+\infty} (x-\mu) f(x)\,dx =$$

$$= \int_{-\infty}^{+\infty} x f(x)\,dx - \int_{-\infty}^{+\infty} \mu f(x)\,dx =$$

$$= \underbrace{\int_{-\infty}^{+\infty} x f(x)\,dx}_{E(x)} - \mu \underbrace{\int_{-\infty}^{+\infty} f(x)\,dx}_{1} =$$

$$= \mu - \mu \cdot 1 = 0$$

b) Il valore medio di una costante (ovvero di una v.A. degenere) è la costante stessa:

$$E(c) = c$$

dim/

$X = c$ v.A. degenere $\quad P(X=c)=1$

$$p(x) = \begin{cases} 1, & x = c \\ 0, & x \neq c \end{cases} \quad (X \text{ v.A. binaria})$$

$$E(X) = E(c) = \sum x_i \, p(x_i) = c \cdot \underbrace{p(c)}_{=1} = c$$

c) Il valore medio di una trasformata lineare di X è uguale alla trasformata lineare di $E(X)$:

$$E(a+bX) = a + b\,E(X)$$

(l'operatore $E(\cdot)$ è un operatore lineare)

se $a = 0$
$E(bX) = b\,E(X)$
se $Y = g(X)$
$E(bY) = E[b\,g(X)] = b\,E[g(X)]$

dim/ Data la v.A. X, la funzione $g(x) = a + bx$ e la v.A. $Y = g(X) = a + bX$ trasformata lineare di X secondo $g(\cdot)$, allora:

$$E[g(X)] = \sum_i g(x_i)\,p(x_i)$$

$$\Rightarrow E(a+bX) = \sum_i (a + bx_i)\,p(x_i) = \sum_i \left[a\,p(x_i) + b x_i\,p(x_i)\right] =$$

$$= \sum_i a\,p(x_i) + \sum_i b x_i\,p(x_i) =$$

$$= a \underbrace{\sum_i p(x_i)}_{=1} + b \underbrace{\sum_i x_i\,p(x_i)}_{=E(x)} = a + b\,E(X)$$

cvd.

d) Il valore medio di una v.A. somma di più v.A. (v.A. continue) è la somma di tutti i valori medi delle singole v.A. che compongono la somma:

se $Y = X_1 + X_2 + \_\_ + X_n \Rightarrow E(X_1 + X_2 + \_ + X_n) = E(X_1) + E(X_2) + \_ + E(X_n)$

o anche $E[g_1(x) + \_ + g_n(x)] = E[g_1(x)] + \_\_ + E[g_n(x)]$

o anche $E[c_1 g_1(x) + \_ + c_n g_n(x)] = c_1 E[g_1(x)] + \_\_ + c_n E[g_n(x)]$

dim/ $$E[c_1 g_1(x) + c_2 g_2(x)] = c_1 E[g_1(x)] + c_2 E[g_2(x)] \ ?$$
$(n=2)$

$$E[c_1 g_1(x) + c_2 g_2(x)] = \int_{-\infty}^{+\infty} [c_1 g_1(x) + c_2 g_2(x)] f(x)\, dx =$$

$$= c_1 \underbrace{\int_{-\infty}^{+\infty} g_1(x) f(x)\, dx}_{\downarrow} + c_2 \underbrace{\int_{-\infty}^{+\infty} g_2(x) f(x)\, dx}_{\downarrow} =$$

$$= c_1 E[g_1(x)] + c_2 E[g_2(x)]$$

cvd

e) $E(X) = \mu$ è il valore che minimizza rispetto ad a l'espressione $E[(X-a)^2]$, ovvero:

$$E[(X-a)^2] = \varphi(a) = \min_a \quad \text{per } a = \mu$$

dim/ $\varphi(a) = E[(X-a)^2] = \sum_i (x_i - a)^2 p(x_i)$

$$\varphi'(a) = \sum_i 2(x_i - a)(-1) p(x_i) = -2\left[\sum_i x_i p(x_i) - a \sum_i p(x_i)\right] =$$

$$= -2(\mu - a)$$

$\Rightarrow$ per $a = \mu$ $\quad \varphi'(a) = -2(\mu - a) = 0$

mentre $\varphi''(a) = 2 > 0 \quad \forall a \quad \Rightarrow a = \mu$ minimo (da Pag. 23 / Proprietà di $F(x)$)

2a) $\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} P(X \leq x) = P\left[\lim_{x \to -\infty} (X \leq x)\right] = P(\emptyset) = 0$

Proprietà di continuità della $P(\cdot)$
(consente di scambiare l'operazione di limite con la funzione $P(\cdot)$)

2b) $\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X \leq x) = P\left(\lim_{x \to +\infty} (X \leq x)\right) = P(\mathbb{R}) =$

$= P(\Omega) = 1$

3a) $\lim_{h \to 0^+} \left[F(x+h) - F(x)\right] = \lim_{h \to 0^+} P\left[x < X \leq x+h\right] =$

$= P\left[\lim_{h \to 0^+} (x < X \leq x+h)\right] = P(\emptyset) = 0$

dato che $\lim_{h \to 0^+} (x, x+h] = \emptyset$

$\Rightarrow \lim_{h \to 0^+} F(x+h) = F(x)$

3b) $\lim_{h \to 0^+} \left[F(x) - F(x-h)\right] \neq \lim_{h \to 0^+} P\left(x-h < X \leq x\right) =$

$= P\left[\lim_{h \to 0^+} (x-h < X \leq x)\right] =$

$= P(X = x) \geq 0$

dato che $\lim_{h \to 0^+} (x-h, x] = x$

se $P(X = x) = 0 \Rightarrow F(x)$ è cont. anche a sinistra di x

se $P(X = x) > 0 \Rightarrow P(X = x) = F(x) - \lim_{h \to 0^+} F(x-h) =$

$= F(x) - F(x^-)$

- VARIANZA DI UNA V. A.

Data la v.A. X con $E(X) = \mu$ e la funzione $g(x) = (x-\mu)^2$ v.A. trasformata di X secondo $g(\cdot)$, si definisce varianza della v.A. X il valore medio della funzione $g(X)$

$$E[g(X)] = E[(X-\mu)^2] = V(X)$$

$$V(X) = \begin{cases} \sum_i (x_i - \mu)^2 p(x_i) & \text{per V.A. DISCRETE} \\ \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx & \text{per V.A. CONTINUE} \end{cases}$$

La quantità $\sigma_x = +\sqrt{V(X)}$ prende il nome di scarto quadratico medio o deviazione standard di X.

• Proposizione

$$V(X) = E(X^2) - [E(X)]^2$$

dim/

$$E(X) = \sum_i x_i p(x_i)$$

$$V(X) = \sum_i (x_i - \mu)^2 p(x_i) = \sum_i (x_i^2 - 2x_i\mu + \mu^2) p(x_i) =$$

$$= \underbrace{\sum_i x_i^2 p(x_i)}_{E(X^2)} - 2\mu \underbrace{\sum_i x_i p(x_i)}_{=E(X)=\mu} + \mu^2 \underbrace{\sum_i p(x_i)}_{=1} =$$

$$= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - E$$

$$= E(X^2) - [E(X)]^2 \qquad \text{cvd}$$

• Proprietà della varianza

a) $V(X) = 0 \iff X = c$ ovvero
la varianza di una v.A. degenere è nulla

dim/
$$V(x) = \sum_i (\underbrace{x_i}_{c} - \underbrace{\mu}_{E(c)=c})^2 \underbrace{p(x_i)}_{p(c)=1} = (c-c)^2 \cdot 1 = 0$$

b) $V(a+X) = V(X)$, $\forall a \in \mathbb{R}$

dim/ poniamo $a + X = Y$ v.A.
dato $\nu = E(Y)$ e $h(Y) = (Y-\nu)^2$
$$\Rightarrow V(Y) = E[(Y-\nu)^2] = E[(Y-E(Y))^2] =$$
$$= E([(a+X) - E(a+X)]^2) =$$

per la proprietà di E:
$E(a+bX) = a + bE(X)$
$$\rightarrow = E([\cancel{a} + X \cancel{-a} - E(x)]^2) =$$
$$= E([X - \underbrace{E(x)}_{\mu}]^2) = V(X) \quad \underline{cvd}$$

c) $V(bX) = b^2 V(X)$, $\forall b \in \mathbb{R}$

dim/
$$V(bX) = E([bX - E(bX)]^2) = E([bX - bE(x)]^2) =$$
$$= E(b^2[X - E(x)]^2) = b^2 E([X - E(x)]^2) =$$
$$= b^2 V(X) \quad \underline{cvd}$$

d) <u>corollario</u>: $V(a+bX) = b^2 V(X)$, $\forall a,b \in \mathbb{R}$

- MOMENTI

- **Momento ordinario di ordine r**

Data una V.A. X, una funzione $g(x) = x^r$ e una V.A. Y trasformata secondo la funzione $g(\cdot)$, $Y = g(X) = X^r$ con $r \in \mathbb{N}$ si definisce momento ordinario di ordine r della V.A. X la quantità

$$E[g(X)] = E[X^r]$$

Per $r = 1$ $\quad E[g(x)] = E(x) = \mu$ $\quad$ valore medio di X (momento primo)

$r = 2$ $\quad E[g(x)] = E(x^2)$ $\quad$ momento secondo

$r = 3$ $\quad E(x^3)$ $\quad$ momento terzo

$\vdots \quad \vdots \quad \vdots$

NB: come possiamo vedere il valore medio di una V.A. è il momento ordinario di X di ordine 1

- **Momento r-esimo rispetto ad a e momento centrale r-esimo**

È un caso particolare della prima*, se poniamo $g(x) = (x-a)^r$ $\quad a, r \in \mathbb{R}$ si definisce momento r-esimo rispetto ad a (con a, r opportuni) la quantità

$$E[g(x)] = E[(X-a)^r]$$

[*: basta considerare $y = x - a$, $g(y) = y^r$]

Se poniamo $a = \mu \Rightarrow E[g(X)] = E[(X-\mu)^r]$ è denominato momento centrale r-esimo

Per $r=1$ $\quad E[(X-\mu)] = 0$ (dalle proprietà di $E(X)$)
(momento centrale primo)

$r=2$ $\quad E[(X-\mu)^2] = V(X)$ varianza di X
(momento centrale secondo)

m

- FUNZIONE GENERATRICE DEI MOMENTI: $m(t)$

Data la v.A. X e la v.A. $g(X) = e^{tX}$, $t \in \mathbb{R}$, trasformata di X secondo la funzione $g(x) = e^{tx}$, si dice funzione generatrice dei momenti di X il valore medio di e^{tX}, ovvero:

$$m(t) = E(e^{tX}) = \begin{cases} \sum_{x_i} e^{tx_i} p(x_i) & , X \text{ V.A. DISCRETA} \\ \int_{-\infty}^{+\infty} e^{tx} f(x)\, dx & , X \text{ V.A. CONTINUA} \end{cases}$$

a condizione che esso esista almeno per valori di t di un intorno dello zero, ovvero

se $\exists\, E(e^{tX}) \quad \forall t : |t| \leq t_0$

$\Rightarrow E(e^{tX})$ è f. generatrice dei momenti di X

<u>Nota</u>: - $E(e^{tX})$ è funzione di t e viene indicata con $m(t)$

- il nome di funzione generatrice dei momenti deriva dal fatto che la derivata r-esima di $m(t)$ calcolata in $t=0$ fornisce il momento r-esimo ordinario di X:

$$m^{(r)}(0) = \sum_{x_i} x_i^r p(x_i) = E(X^r)$$

dato che $m^{(r)}(t) = \sum_{x_i} e^{tx_i} x_i^r p(x_i)$

• Teorema

- UNICITÀ DELLA F.G.M. (detta anche prima proprietà delle F.G.M.)

Date le V.A. X_1 e X_2 con:

X_1, $F_1(x)$, $m_1(t)$

X_2, $F_2(X)$, $m_2(t)$

allora: $F_1(x) = F_2(x) \iff m_1(t) = m_2(t)$

$\forall x$ e almeno $\forall t : |t| \leq t_0$

DISTRIBUZIONI DI PROBABILITA'

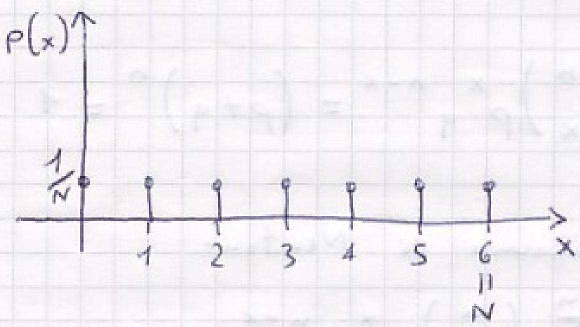
- DISTRIBUZIONE DI BERNOULLI

(v. pag. 27)

- DISTRIBUZIONE UNIFORME DISCRETA

$$X: \quad p(x) = \begin{cases} \frac{1}{N}, & x = 1, 2, \dots, N \\ 0, & x \neq 1, 2, \dots, N \end{cases}$$

$$= \frac{1}{N} \, I_{\{1,2,\dots,N\}}(x)$$



$p(x)$ è una funzione di probabilità, infatti:

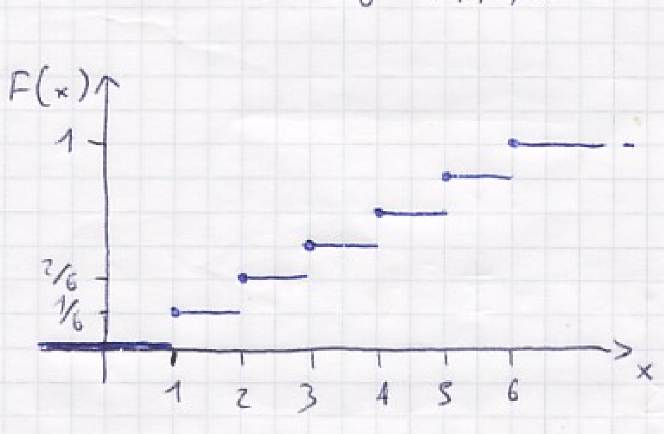
$$\sum_{x=1}^{N} p(x) = N \cdot \frac{1}{N} = 1$$

Es/

$\mathcal{E}$: lancio di un dado

$\Omega = \{1, 2, 3, 4, 5, 6\}$, $N = 6$, $X = X(\omega) = \omega$

$$X \sim p(x) = \frac{1}{6} I_{\{1,2,\dots,6\}}(x)$$



funzione di ripartizione

- DISTRIBUZIONE BINOMIALE

$$X: \quad p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & , \quad x = 0, 1, \ldots, n \\ 0 & , \quad x \neq 0, 1, \ldots, n \end{cases}$$

con parametri n e p tali che

$$n \in \mathbb{Z}^+ \quad , \quad 0 \leq p \leq 1$$

$$e \quad q = 1 - p$$

$p(x)$ è una funzione di probabilità, infatti:

$$\sum_{x=0}^{n} p(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1$$

dalla formula del binomio di Newton:

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

Nota: per $n=1$ si ritrova la distribuzione di bernoulli