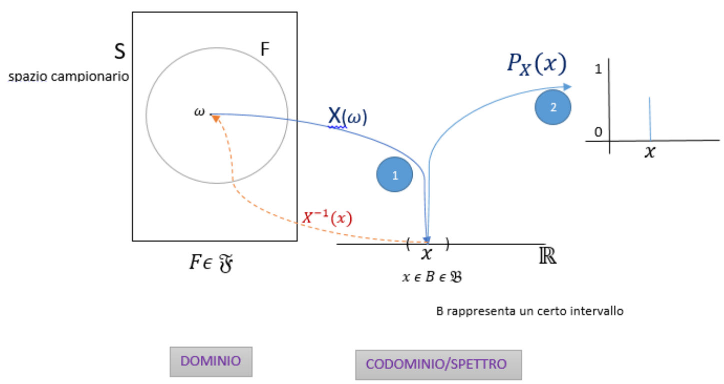
Variabili aleatorie

13 aprile 2017

Si introduce il concetto di variabile aleatoria discreta e continua e di legge di probabilitá. Si considera in seguito la funzione di ripartizione come caratterizzazione unificante delle distribuzioni di probabilitá continue e discrete. Infine, si definisce la funzione della densitá di probabiitá.

1 *Definizione di variabile aleatoria e misurabilitá*

Informalmente, una variabile aleatoria (V.A.) X o meglio una funzione aleatoria é una funzione $X(\omega)$ che mappa un esito dello spazio campionario $\omega \in S$ in un valore reale. A tale valore é possibile successivamente assegnare una legge di probabilitá $P_X(\cdot)$. L'azione di mapping della VA $X(\cdot)$ é schematizzata in Figura 1



Piú formalmente, sia $(S, \mathcal{F})$ lo spazio probabilizzabile associato allo spazio campionario. Sia $F \in \mathcal{F}$ un evento, e $B \in \mathcal{B}$, un evento ovvero un intervallo di $\mathcal{B}$, la minima σ-algebra contenente tutti gli intervalli di $\mathbb{R}$.

Definizione 1.1 (Funzione aleatoria) *Una funzione* $X : (S, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ *é una funzione aleatoria se*

$$X^{-1}(B) = \{\omega \in S | X(\omega) \in B\} = F$$

(ovvero é misurabile).

In questo caso F é il *dominio* della funzione, l'intervallo B il *codominio* e $X^{-1}(B)$ la sua *anti-immagine*

In buona sostanza la VA X *induce* in $\mathbb{R}$ le immagini degli eventi F, $X(F)$ definiti dalla relazione:

$$X(F) = \{x \in \mathbb{R} \mid \exists \omega \in F \text{ per cui } X(\omega) = x\}$$

Il seguente teorema ci dice che la misurabilitá di X, condizione fondamentale per poi poter definire una legge di probabilitá su X, si ha se e solo se l'anti-immagine ricade in un evento F.

Teorema 1.2 (Misurabilitá) *X é misurabile sse*

$$X^{-1}(-\infty, c) = \{\omega \in S \mid X(\omega) < c\} \forall c \in \mathbb{R} \tag{1}$$

Una rappresentazione del significato del teorema é delineata in figura 1

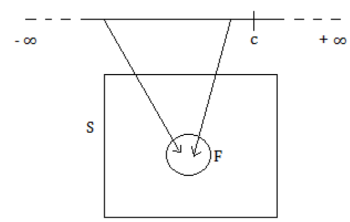


Figura 1: Condizione necessaria e sufficiente di misurabilitá: l'anti-immagine di B ricade in un evento F

Esempio 1.3 (VA Indicatrice) *Definiamo la seguente:*

$$X(\omega) = I_A(\omega) = \begin{cases} 1 & se\ \omega \in A \\ 0 & se \omega \notin A (\omega \in \sim A) \end{cases}. \tag{2}$$

Applichiamo la condizione di Equazione 1 per verificare la condizione di misurabilitá. Per definizione della funzione, i punti di interesse sono $(0, 1)$. *Facendo variare* c*, sono da considerare allora i seguenti casi:*

$$X^{-1}(-\infty, c) = \begin{cases} \emptyset & se\ c \leq 0 \\ \{\sim A\} & se\ 0 < c \leq 1 \\ \{\sim A, A\} = S & se\ c > 1 \end{cases}$$

Dunque $I_A(\omega)$ é misurabile

In condizione di misurabilitá, i seguenti corrispondono ad eventi della σ-algebra $\mathcal{F}$:

$$\{X(\omega) \leq x\}$$

$$\{x_1 < X(\omega) \leq x_2\}$$

$$\{X(\omega) > x\}$$

2 *Legge di probabilitá per una VA*

Si consideri lo spazio probabilizzabile $(\mathbb{R}, \mathcal{B})$. Una VA X *induce* su $(\mathbb{R}, \mathcal{B})$ la misura di probabilitá

$$B \in \mathcal{B} \mapsto P_X(B) = P(X^{-1}(B)) \in [0, 1] = prob(X \in B), \tag{3}$$

ovvero una funzione $P_X : \mathcal{B} \to [0,1]$ che viene detta *Legge di probabilitá* per la VA X

Si noti che:

- $P(X^{-1}(B))$: é la probabilitá definita nel dominio, ovvero nello spazio campionario S
- $P_X(B)$: é la legge di probabilitá definita nel codominio $\mathbb{R}$

La legge di probabilitá $P_X : \mathcal{B} \to [0,1]$ soddisfa gli assiomi di Kolmogorov:

A1 $PX(B) \geq 0$: infatti $P(X) = P(X^{-1}(B)) = P(F) \geq 0$

A2 $B, C \in \mathcal{B}$ e $B \cap C = \emptyset$:

$P_X(B \cup C) = P(X^{-1}(B \cup C)) =$

$= P(X^{-1}(B) \cup X^{-1}(C))$

$= P(X^{-1}(B)) + P(X^{-1}(C))$

$= P_X(B) + P_X(C)$

A3 $P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R}) = P(S) = 1$

3 *Variabili aleatorie continue e discrete*

Il codominio $B \in \mathcal{B}$ della variabile aleatoria X puó essere numerabile o non numerabile. Nel primo caso si dice che X ha uno *spettro discreto*, nel secondo caso uno *spettro continuo*. I due casi sono rappresentati in Figura 2.

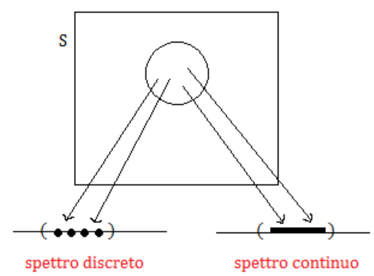


Figura 2: Spettro discreto e continuo di una VA

Definizione 3.1 *Una variabile aleatoria X con spettro discreto é una* variabile aleatoria discreta*. Una variabile aleatoria X con spettro continuo é una* variabile aleatoria continua*.*

3.1 *Variabili aleatorie discrete*

Una VA discreta puó essere rappresentata mediante la VA indicatrice introdotta precedentemente:

$$X(\omega) = \sum_k X_k I_{A_k}(\omega)$$

dove $A_k = X^{-1}(X_k) = \{\omega \in S : X(\omega) = x_k\}$

Nel caso discreto posso determinare esattamente qual é la probabilitá in un punto o evento elementare x dello spettro, che é possibile scrivere con una delle seguenti notazioni tra loro equivalenti

$$p_X(x) = P_X(\{X = x\}) = P_X(X = x))$$

Assegnata la legge di probabilitá $P_X(\cdot)$, se considero lo spettro discreto della VA posso rappresentare come la probabilitá si distribuisce sui vari punti dello spettro, come disegnato in Figura 5.

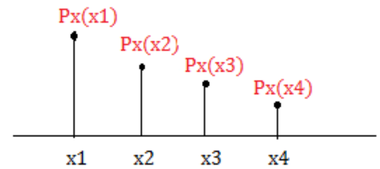


Figura 3: PMF (probability mass function): il grafico rappresenta come si distribuisce, nel caso discreto, la probabilitá sulle varie "masse"

Esempio 3.2 (Caso discreto) *Consideriamo l'esperimento del lancio di due monete, con spazio campionario:*

$$S = \{(T, T), (T, C), (C, T), (C, C)\} \tag{4}$$

e probabilitá $P(T, T) = P(T, C) = P(C, T) = P(C, C) = 1/4$

Definiamo la variabile aleatoria X = *"numero di teste T".*

Gli eventi di interesse, che costituiscono lo spettro discreto di X *sono dunque:*

$X(C, C) = 0, X(C, T) = X(T, C) = 1, X(T, T) = 2$

Il mapping effettuato da X *é rappresentato in Figura 4, e mette in evidenza il legame tra la probabilitá* P *nello spazio campionario* S *e la legge di probabilitá associata a* X*, cioé* P_X*. Ad esempio:* $P(C, T) = 1/4 = P_X(X = 1) = P(X^{-1}(1)) = P_X(X(\{C, T\}) = 1)$

Possiamo definire sui punti dello spettro discreto $X = 0, X = 1, X = 2$ *le corrispondenti misure di probabilitá:*

$p_X(0) = P_X(X = 0) = P(C, C) = 1/4,$

$p_X(1) = P_X(X = 1) = P(C, T) + P(T, C) = 1/2,$

$p_X(2) = P_X(X = 2) = P(T, T) = 1/4.$

La Figura mostra il grafico della PMF.

Notiamo come la PMF ricavata sia normalizzata a 1:

$\sum_{x=\{0,1,2\}} p_X(x) = p_X(0) + p_X(1) + p_X(2) = 1/4 + 1/2 + 1/4 = 1.$

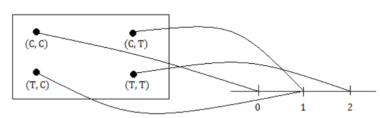


Figura 4: L'azione della VA X = "numero di teste T"

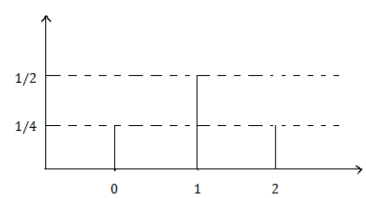


Figura 5: PMF della VA X = "numero di teste T"

3.2 *Variabili aleatorie continue*

Nel <u>caso continuo</u>, intuitivamente, tra due punti dello spettro esistono infiniti punti, dunque non é possibile misurare la probabilitá nel singolo punto. Ovvero, la misura di probabilitá in un punto é nulla.

Sempre intuitivamente, si puó considerare di misurare la probabilitá in un "intervalletto" dx (piccolo a piacere) :

$$f_X(x)dx = P_X(X \in dx)$$

La funzione $f_X(\cdot)$ viene detta *funzione densitá di probabilitá* (probability density function, PDF).

Intuitivamente, possiamo pensare alla PMF definita nel caso discreto come ad un caso particolare di densitá misurata solo in un insieme di punti di massa, ovvero una "densitá discreta" $p_X(x)$ che per ogni $x_i \in B$ vale:

$p_X(x_i) = P_X(\{X = x_i\}) = P_X(X = x_i)$

e vale $p_X(x) = 0$ al di fuori dello spettro discreto.

Per risolvere il dilemma discreto / continuo, anticipiamo che é possibile definire una funzione $F_X : \mathbb{R} \to [0,1]$ valida sia per X discrete che per X continue:

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

Essa viene detta *funzione di ripartizione* (f.d.r) o *funzione cumulativa* (CDF, Cumulative Density Function) o ancora semplicemente *funzione di distribuzione* (DF, Distribution Function). Da tale funzione sará possibile ricavare sia le funzioni di densitá continue $f_X(\cdot)$, sia le "densitá discrete" $p_X(\cdot)$.

4 *Definizione di funzione di ripartizione o funzione cumulativa (CDF)*

Definizione 4.1 (Funzione di ripartizione) *Definiamo* funzione di ripartizione *(o funzione cumulativa, CDF) della V.A. (continua o discreta) X, la funzione $F_X : \Re \to [0,1]$ per cui vale*

$$F_X(x) = P_X(\{X \leq x\}).$$

Esempio 4.2 (Caso discreto) *Lancio una moneta dove $P(T) = p, P(C) = q = 1 - p$, e definisco la VA $X(\cdot)$ come il mapping $X(T) = 1$ e $X(C) = 0$.*

Per costruire la CDF andiamo a variare il valore di x nel codominio (da $-\infty$ a $+\infty$) vedendo quali eventi elementari via via "cattura". Per l'A2 di Kolmogorov, essendo gli eventi disgiunti, la somma cumulativa *delle probabilitá di tali eventi é $P_X(X \leq x)$, cioé $F_X(x)$, il cui andamento, per l'appunto descrive come queste si accumulino per x crescenti.*

Nell'esempio considerato, gli eventi possibili sono $X(C) = 0$ e $X(T) = 1$, pertanto tale procedura si riduce ad analizzare i seguenti casi:

- *$x < 0$: $X(T) = 1 > x$ e $X(C) = 0 > x$*

  Funzione di ripartizione: $F_X(x) = P_X(X \leq x) = 0$.

- *$0 \leq x < 1$: $X(T) = 1 > x$ e $X(C) = 0 \leq x$*

  Funzione di ripartizione: $F_X(x) = P_X(X \leq x) = P(C) = q$

- *$x \geq 1$: $X(T) = 1 \leq x$ e $X(C) = 0 \leq x$.*

  Funzione di ripartizione: $F_X(x) = P_X(X \leq x) = P(T \cup C) = 1$

Abbiamo quindi una funzione di ripartizione che si disegna come in Figura 4.

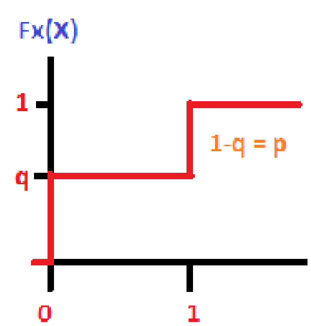


Figura 6: Grafico della funzione di ripartizione $F_X(x)$ con $X(T) = 1$ e $X(C) = 0$

Esempio 4.3 (Caso continuo) *Definiamo la VA continua:*

$$X(t) = a,$$

$\forall t \in S$

Questo é un caso degenere, perché X é una funzione che mappa qualsiasi esito t del dominio S in un unico punto del codominio B (quindi avremo che il punto $x = a$ ha probabilitá 1, ovvero è un evento certo, cfr. Figura 5 piú avanti).

In questo caso abbiamo due situazioni:

- *$x < a$: $X(t) = a > x$*

  Funzione di ripartizione: $F_X(x) = P_X(X \leq x) = P(\emptyset) = 0$

- *$x \geq a$: $X(t) = a \leq x$.*

  Funzione di ripartizione: $F_X(x) = P_X(X \leq x) = P(S) = 1$

Abbiamo quindi una funzione di ripartizione che si disegna come in Figura 4.

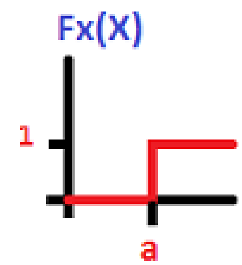


Figura 7: Grafico della funzione di ripartizione $F_X(x)$ con $X(t) = a$

5 *Proprietá della funzione di ripartizione*

Sia data una f.d.r $F_X(x) = P_X(\{X \leq x\})$. Allora valgono le seguenti proprietá

Prop. 5.1 *Se $x_1 < x_2$:*

$$F_X(x_1) \leq F_X(x_2) \tag{5}$$

Dimostrazione

Per ipotesi:
$\{X \leq x_1\} \subseteq \{X \leq x_2\}$
$P_X(\{X \leq x_1\}) \leq P_X(\{X \leq x_2\})$
Per definizione di CDF:
$F_X(x_1) \leq F_X(x_2)$

Prop. 5.2

$$F_X(+\infty) = 1 \tag{6}$$

$$F_X(-\infty) = 0 \tag{7}$$

Dimostrazione

$F_X(+\infty) = P_X(\{X \leq +\infty\}) = P(X^{-1}(\{X \leq +\infty\})) = P(S) = 1$
$F_X(-\infty) = P_X(\{X \leq -\infty\}) = P(X^{-1}\{X \leq -\infty\}) = P(\emptyset) = 0$

La $P_X(\{X > x\}) = S_X(x)$ definisce la funzione complementare della CDF, chiamata anche *funzione di sopravvivenza*

Prop. 5.3 .

$$P_X(X > x) = 1 - F_X(x) \tag{8}$$

Dimostrazione *Per costruzione, gli eventi sono disgiunti e partizionano $\Re$:*

$\{X \leq x\} \cap \{X > x\} = \emptyset$
$\Re = \{X \leq x\} \cup \{X > x\}$
Quindi, per A2 di Kolmogorov:
$P_X(\Re) = P_X(\{X \leq x\}) + P_X(\{X > x\})$
Ovvero:
$1 = P_X(\{X \leq x\}) + P_X(\{X > x\}) = F_X(x) + P_X(\{X > x\})$

Prop. 5.4 (non decrescenza) *Se $a < b$, allora*

$$P_X(a < X \leq b) = F_X(b) - F_X(a) \tag{9}$$

Dimostrazione *Per costruzione:*
$\{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\}$
Quindi, per A2 di Kolmogorov:
$P_X(\{X \leq b\}) - P_X(\{X \leq a\}) = P_X(\{a < X \leq b\})$
Per definizione di CDF:
$F_X(b) - F_X(a) = P_X(\{a < X \leq b\})$

Prop. 5.5 (Continuitá a destra)

$$F_X(x^+) = F_X(x) \tag{10}$$

Accenno di dimostrazione *Basta dimostrare che:*

$$\lim_{\varepsilon \to 0} P_X(X \leq x + \varepsilon) = F_X(x) \tag{11}$$

Per definizione di CDF:

$$P_X(X \leq x + \varepsilon) = F_X(x + \varepsilon)$$

Allora, l'Equazione 10 vale in virtú della seguente:

$$\lim_{\varepsilon \to 0} F_X(x + \varepsilon) = F_X(x^+) \tag{12}$$

L'Equazione 12 si dimostra utilizzando il limite di una successione di insiemi. Omettiamo.

Graficamente il risultato di Equazione 10 é rappresentato in Figura 5.

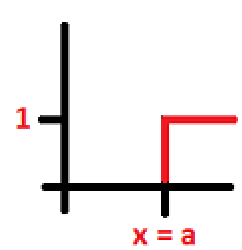


Figura 8: La CDF in $x = a$ assume il valore che si legge "avvicinandosi" da destra $F_X(a) = F_X(a^+) = 1$

Prop. 5.6

$P_X(X = x) = F_X(x) - F_X(x^-)$ (13)

Dimostrazione *Usiamo la Prop. 5.4,* $P(a < X \leq b) = F_X(b) - F_X(a)$

Poniamo: $a = x - \varepsilon$ *e* $b = x$

Allora:

$P_X(x - \varepsilon < X \leq x) = F_X(x) - F_X(x - \varepsilon)$

Passando al limite:

$\lim_{\varepsilon \to 0} P_X(x - \varepsilon < X \leq x) = P_X(X = x)$

$\lim_{\varepsilon \to 0} F_X(x - \varepsilon) = F_X(x^-)$

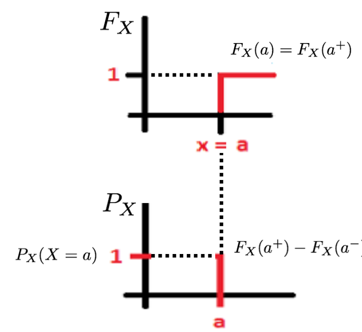


Figura 9: Calcolo della PMF (in basso) dalla CDF (in alto) usando le proprietá (5.5) e (5.6)

Le proprietá (5.5) e (5.6) insieme ci consentono di "disegnare" la PMF a partire dalla CDF. Riconsideriamo la CDF a gradino in Figura 5

Vediamo dalla CDF riportata in Figura 5 che:

$$P_X(X = a) = F_X(a) - F_X(a^-) = F_X(a^+) - F_X(a^-) = 1 - 0 = 1$$

cioé, l'evento $X = a$ assume probabilitá 1, ovvero é un evento certo.

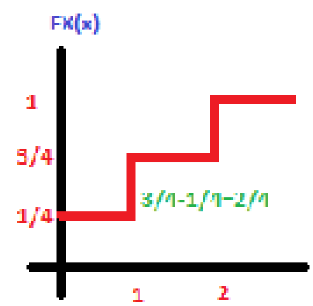


Figura 10: CDF ottenuta dal conteggio delle "teste" nell'esperimento del lancio di due monete

Esempio 5.7 (Lancio di due monete) *Definiamo la V.A.* $X(\cdot)$ *che conta il numero di esiti "testa" (T):*

$X(T, T) = 2$, $X(T, C) = X(C, T) = 1$ *e* $X(C, C) = 0$

Possiamo distinguere quattro casi:

- $x < 0$: $\{X \leq x\} = \emptyset$

  Funzione di ripartizione: $F_X(x) = 0$

- $0 \leq x < 1$: $\{X \leq x\} = \{C, C\}$

  Funzione di ripartizione: $F_X(x) = P_X(x = 1) = P(C, C) = \frac{1}{4}$

- $1 \leq x < 2$ $\{X \leq x\} = \{(C, C), (C, T), (T, C)\}$

  Funzione di ripartizione: $F_X(x) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$

- $x \geq 2$: $\{X \leq x = s\}$

  Funzione di ripartizione: $F_X(x) = 1$

  La CDF é dunque rappresentabile come in Figura 5.7

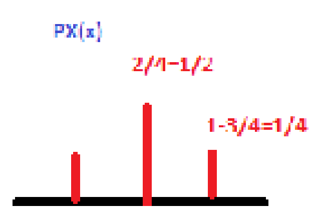


Figura 11: PMF ottenuta dal conteggio delle "teste" nell'esperimento del lancio di due monete

Dall'esempio precedente, usando come prima le proprietá (5.5) e (5.6), dalla CDF possiamo ottenere facilmente la PMF (Figura 5)

6 *Definizione di funzione di densitá di probabilitá*

Le proprietá (5.4), (5.5), (5.6), (**??**) ci dicono che la funzione $F_X : \Re \rightarrow [0, 1]$ é funzione assolutamente continua e derivabile in $\Re$ e, nel caso di X discreta é ancora derivabile con continuitá a meno di un numero finito di punti che costituiscono discontinuitá di prima specie

E' allora possibile enunciare la seguente

Definizione 6.1 (Densitá di probabilitá) *Una funzione $f_X : \Re \rightarrow \Re$ tale che:*

$$f_X(x) = \frac{dF_X(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{F_X(x + \Delta x) - F_X(x)}{\Delta x} \tag{14}$$

con $f_X(x) \geq 0$, per ogni $x \in \Re$ si dice funzione di densitá di probabilitá *associata a F_X*

Piú in dettaglio, nel caso continuo, f_X é sempre ben definita. Nel caso discreto, la densitá di probabilitá corrisponde alla PMF (Probability Mass Function):

$$P_X(X = x_i) = p_X(x_i) = p_i \tag{15}$$

dove $i = 1, 2, \cdots$ é un indice discreto. Nel caso discreto la CDF si scrive come la somma

$$F_X(k) = P_X(X \leq x_k) = \sum_{i \leq k} p_i. \tag{16}$$

tipicamente una funzione continua a tratti.

Resta tuttavia possibile, e talvolta utile, trattare la PMF definita in Equazione (15) come fosse continua,

$$f_X(x_i) = P_X(X = x_i). \tag{17}$$

La forma esatta della rappresentazione continua della PMF la si puó ottenere direttamente derivando la CDF come definito in Eq. (14). A tale scopo, si noti dagli esempi precedenti che per VA discrete la $F_X(k)$ é una funzione non decrescente a gradini. Come abbiamo visto, per le proprietá (5.5) e (5.6) ciascun gradino, nel generico punto $X = x_i$ ha un'ampiezza che vale

$$p_i = P_X(X = x_i) = F_X(x_i^+) - F_X(x_i^-)$$

Un modo equivalente di scrivere la precedente é di vedere ciascun gradino di altezza p_i come il risultato della moltiplicazione

$$p_i \times H(x - x_i) \tag{18}$$

dove $H(x - x_i)$ é una funzione gradino di altezza 1 o *funzione di Heaviside*

$$H(x - x_i) = \begin{cases} 0 & \text{if } x < x_i \\ 1 & \text{if } x > x_i \end{cases}. \tag{19}$$

Possiamo allora scrivere la CDF discreta (20) come

$$F_X(x) = \sum_{x_i \leq x} p_X(x_i) = \sum_{x_i \leq x} p_X(x_i) H(x - x_i). \tag{20}$$

Derivando:

$$f_X(x) = \frac{d}{dx} \sum_{x_i \leq x} p_X(x_i) H(x - x_i) = \sum_{x_i \leq x} p_X(x_i) \frac{d}{dx} H(x - x_i) = \sum_i p_i \delta(x - x_i) \tag{21}$$

dove $\delta(x - x_i)$ é la delta di Dirac.

$$\int_{-\infty}^{+\infty} g(x)\delta(x - x_0) = g(x_0), \int_{-\infty}^{+\infty} \delta(x) dx = 1 \tag{22}$$

Dalla Equazione (14), che definisce la pdf, e dal Teorema Fondamentale del Calcolo Integrale consegue che, per ogni $a < b \in \Re$,

$$P_X(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx \tag{23}$$

E' immediato notare che la PDF é normalizzata:

$$P_X(-\infty < X \leq +\infty) = F_X(+\infty) - Fx(-\infty) = \int_{-\infty}^{+\infty} fx(x) dx = 1 \tag{24}$$

Si utilizza talvolta la notazione abbreviata:

$$prob(X \in dx) = P_X(X \in dx) = f_X(x) dx. \tag{25}$$

Essa trova motivazione nella seguente approssimazione:

$$prob(X \in dx) = P_X(x < X \leq x + dx) = F_X(x + dx) - F_X(x) = \int_x^{x+dx} f_X(u) du \approx f_X(x) dx, \tag{26}$$

che vale per dx piccolo.

La medesima approssimazione mette in evidenza un importante fatto che vale per le distribuzioni continue:

$$P_X(X = x) = \lim_{dx \to 0} P_X(x < X \leq x + dx) = \lim_{dx \to 0} \int_x^{x+dx} f_X(u) du = 0 \tag{27}$$

ovvero, per una VA $X(\cdot)$ continua la probabilitá in un punto x arbitrario $P_X(X = x)$ ha *misura nulla*.

Ne consegue anche che per una variabile continua

$$P_X(a \leq X \leq b) = P_X(a < X \leq b) = P_X(a \leq X < b) = P_X(a < X < b), \tag{28}$$

potendosi scrivere, per esempio

$$P_X(a \leq X \leq b) = P_X(a \leq X < b) + P_X(X = b) = P_X(a \leq X < b) + 0. \tag{29}$$

Possiamo riassumere formalmente quanto discusso finora nella seguente

Definizione 6.2 *La densitá di una misura di probabilitá P_X, relativa alla VA X su $(\mathbb{R}, \mathcal{B})$ é una funzione misurabile (di Borel) f tale che, per qualunque $x \in \mathbb{R}$,*

$$P_X((-\infty, x]) = F_X(x) = \int_{-\infty}^{x} f(y)dy \tag{30}$$

6.1 *Semplici esempi di distribuzioni discrete e continue*

Esempio 6.3 (Una distribuzione discreta) *Sia X una VA discreta, di distribuzione*

$$P_X(\{X = x\}) = p_X(x) = \left(\frac{1}{2}\right)^x, \quad x > 0$$

Possiamo calcolare i valori di probabilitá che assume nei punti $x = 1, 2, 3, \cdots$

$$\begin{aligned} p_X(1) &= \left(\tfrac{1}{2}\right)^1 = \tfrac{1}{2} \\ p_X(2) &= \left(\tfrac{1}{2}\right)^2 = \tfrac{1}{4} \\ p_X(3) &= \left(\tfrac{1}{2}\right)^3 = \tfrac{1}{8} \\ &\cdots \end{aligned}$$

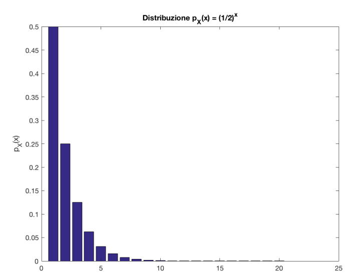


Figura 12: Andamento della distribuzione discreta $p_X(x) = \left(\frac{1}{2}\right)^x$

Se l'evento di interesse fosse $\{X \leq 3\}$

$$P_X(\{X \leq 3\}) = F_X(3) = \sum_{k=0}^{3} p(k) = p(1) + p(2) + p(3) = (\frac{7}{8}) \simeq 0.875$$

La probabilitá dell' evento $\{X > 3\}$ *sarebbe*

$$1 - P_X(\{X \leq 3\}) = P_X(\{X > 3\}) = 1 - 0.875 \simeq 0.125$$

Esempio 6.4 (Una distribuzione continua) *Sia X una VA continua, di densitá*

$$f(x) = e^{-x}, \quad x > 0$$

Come nel esempio discreto possiamo calcolare la probabilitá dell' evento, in un certo intervallo dello spettro continuo, per esempio $\{X \leq 3\}$

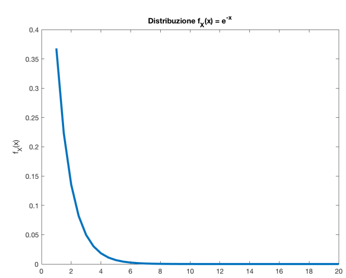


Figura 13: Andamento della distribuzione continua $f(x) = e^{-x}$

$$P_X(\{X \leq 3\}) = F_X(3) = \int_0^3 f(x)dx =$$
$$= \int_0^3 e^{-x}dx = [-e^{-x}]_0^3 = -e^{-3} + e^0 = 1 - e^3 = 0.95$$

Anche in questo caso la probabilitá dell' evento $\{X > 3\}$ *sarebbe*

$$1 - P_X(\{X \leq 3\}) = P_X(\{X > 3\}) = 1 - 0.95$$

Negli esempi precedenti possiamo verificare la proprietá di normalizzazione di cui deve godere una distribuzione (discreta o continua)

Nel primo caso deve valere:

$$\sum_{x=1}^{\infty} p_X(x) = 1 \tag{31}$$

Usiamo la somma della serie geometrica

$$\sum_{x=0}^{\infty} a^x = \frac{1}{(1-a)},$$

che riscriviamo come

$$a^0 + \sum_{k=1}^{\infty} a^k = \frac{1}{(1-a)},$$

ovvero:

$$\sum_{k=1}^{\infty} a^k = \frac{1}{(1-a)} - 1.$$

Da cui, per $a = \frac{1}{2}$ otteniamo la (31).

Per l'esempio continuo, la verifica é immediata:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} e^{-x}dx = \int_0^{\infty} e^{-x}dx = [-e^{-x}]_0^{\infty} = 1$$

Qualora il risultato é diverso da 1, posso sempre ottenere la normalizzazione dividendo la distribuzione per la costante numerica ottenuta. Per esempio nel caso continuo se

$$\int_{-\infty}^{+\infty} f(t)dt = N \neq 1,$$

allora la densitá normalizzata sarebbe:

$$\widetilde{f(t)} = \frac{f(t)}{N}.$$

7 *Prime distribuzioni notevoli*

7.1 *Distribuzione di Bernoulli*

La prima e importante distribuzione che introduciamo é la distribuzione discreta di **Bernoulli** di parametro p, che in letteratura viene indicata in varie notazioni

$$Bern(p) \equiv Bern(X \mid p) = Bern(X; P)$$

Nel seguito useremo spesso la notazione:

$$x \sim Bern(X \mid p)$$

che puó, a seconda dei casi assumere due significati:

1. X é distribuita secondo la legge di Bernoulli $Bern(X \mid p)$;
2. $X = x$ é campionata dalla bernoulliana $Bern(X \mid p)$.

La bernoulliana é sostanzialmente una distribuzione di una variabile aleatoria il cui spazio degli stati originario S ha le stesse caratteristiche - successo, insuccesso - dello spazio degli stati relativo all'esperimento del lancio di una moneta con risultati possibili T, C: $S = \{T, C\}$ con un mapping

$$X(T) = 1, \quad X(C) = 0.$$

In tal caso si definisce:

$$p_X(x) = \begin{cases} 1 - p = q, & x = 0 \\ p & x = 1 \end{cases}$$

Piú sinteticamente si usa la seguente definizione

$$P_X(\{X = x\}) \equiv Bern(X \mid p) = p^x q^{1-x}$$

La proprietá di normalizzazione é facilmente verificabile:

$$\sum_{x=0,1} p_X(x) = p_X(0) + p_X(1) = q + p = 1.$$

La CDF era giá stata calcolata precedentemente nell'esempio 4.2 e graficata in Figura 4. La possiamo sintetizzare come segue

$$F_X(x) = \begin{cases} 0 & x < 0 \\ q & 0 \le x < 1 \\ q + p = 1 & x \ge 1 \end{cases} \tag{32}$$

7.2 *Distribuzione uniforme*

Assumendo ora X VA continua, questa é distribuita con legge uniforme di parametri a, b denotata come

$$Unif(X; a, b) \equiv Unif(X \mid a, b),$$

la cui densitá si scrive:

$$Unif(X \mid a, b) = f_X(x) = \begin{cases} 0 & x < a \\ \frac{1}{b-a} & a \leq x \leq b \\ 0 & x > b \end{cases} \tag{33}$$

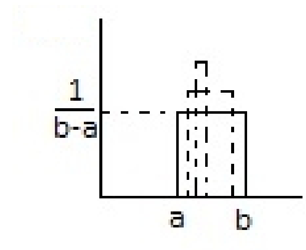


Figura 14: Densitá della distribuzione uniforme

L'andamento della densità é mostrato in Figura 14: si noti come, al restringersi dell'intervallo di supporto $[a, b]$, la distribuzione si "alzi" per mantenere costante l'area sottesa (che per normalizzazione deve essere pari a 1); infatti:

$$P_X(-\infty \leq X \leq +\infty) = \int_{-\infty}^{\infty} f_X(x)dx = \int_a^b \frac{1}{b-a}dx = \\ = \frac{1}{b-a}[x]_a^b = \frac{b-a}{b-a} = 1 \tag{34}$$

Il calcolo della CDF,

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du = \int_a^x \frac{1}{b-a}du = \\ = \frac{1}{b-a}[u]_a^x = \frac{x-a}{b-a}, \tag{35}$$

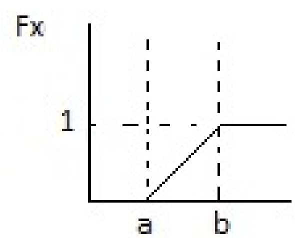


Figura 15: CDF della distribuzione uniforme

ci dice che $F_X(x)$ cresce linearmente al crescere di x nell' intervallo $[a, b]$, come visualizzato in Figura 15.