Probabilità 1

esperimento → è un processo il cui risultato è soggetto ad incertezza

S → **Spazio dei campioni**
S → insieme di tutti i possibili risultati di un esperimento

$\#(A) = n$ → **CARDINALITÀ di un insieme** → numero di elementi dell'insieme

- Lo spazio dei campioni varia in caso di **reinserimento** e in caso di **ordinamento** (diverso ordine diverso risultato es. $(1,2) \neq (2,1)$)

Evento → E
→ è un insieme di risultati, cioè un sottospazio dello spazio dei campioni

$$A \subseteq S$$

sottoinsieme

Unione → $C = A \cup B$. Un elemento $v \in C$ se $v \in A$ oppure $v \in B$

Intersezione → $C = A \cap B$. Un elemento $v \in C$ se $v \in A$ e $v \in B$

Differenza → $C = A - B$. Un elemento $v \in C$ se $v \in A$ e $v \notin B$

Probabilità → L'obbiettivo della probabilità è assegnare ad ogni evento A un numero P(A), detto probabilità dell'evento A che dia misura della possibilità che si verifichi l'evento A.

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

1) $P(A) \geq 0 \ \forall A$
2) $P(S) = 1$
3) Se $A_1, A_2, \dots, A_n$ sono eventi disgiunti → $A_i \cup A_j = \emptyset$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^{n} P(A_i), \forall n$$

A^c → insieme complementare di A nello spazio dei campioni S.

1) $P(A^c) = 1 - P(A)$

2) $P(\emptyset) = 0$

3) Se $A \subset B$ allora $P(A) < P(B)$

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

modello ELM → Equaly Likely Model

↳ tutti gli eventi hanno uguale probabilità di accadere

↳ es. lancio di moneta, dado

$$P(A) = \frac{\text{numero di casi favorevoli}}{\text{numero di casi possibili}}$$

Importante (numero casi favorevoli)

$n \to$ numero palline $\quad K \to$ estrazioni

Insiemi ordinati con reimmissione $\quad n^K \quad$ 6 6 6

Insiemi ordinati senza reimmissione $\quad n(n-1)(n-2)\dots(n-K+1) \quad$ 6 5 4

Insiemi non ordinati con reimmissione

$$\frac{(n-1+K)!}{(n-1)!\,K!}$$

Insiemi non ordinati senza reimmissione

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

nCK

Probabilità condizionata ed eventi indipendenti

La seconda probabilità è quindi condizionata dal primo evento

Dati due eventi A e B, si definisce probabilità condizionata di B dato A e si scrive $P(B|A)$:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \text{se } P(A) > 0$$

Eventi indipendenti

A e B sono detti indipendenti se:

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

- Quando il verificarsi di uno non influenza la probabilità di verificarsi dell'altro.
- A e B^c sono indipendenti
- A^c e B sono indipendenti
- A^c e B^c sono indipendenti
- Analogo in caso di più eventi

- In caso contrario, sono detti **dipendenti**.

Eventi disgiunti

→ Si riferiscono alla relazione tra eventi che non possono verificarsi contemporaneamente.

Probabilita' II

Variabile Aleatoria

Una variabile aleatoria X e' una funzione:

$$X : S \to R$$

che associa ad ogni elemento $\omega \in S$ un numero $X(\omega) = x$

Variabili aleatorie → Lettere Maiuscole e il loro valore con lettere minuscole

Supporto di una variabile aleatoria → indica l'insieme di tutti i possibili valori casuali di un dato esperimento e si indica con S_x.

Esempi:

Variabili Aleatorie Discrete:

↳ Supporto composto da Insieme numerabile (contabile) di elementi

Casi: 1) L'insieme e' costituito da un numero finito di elementi
es. Esperimento di lanciare una moneta due volte $S_x = \{0, 1, 2\}$

2) L'insieme e' costituito da un numero infinito di elementi
es. Lancio ripetuto di un dado finche' non si ottiene un 6. $S_y = \{0, 1 \ldots\}$

Variabili aleatoria continua:

↳ Supporto composto da insieme non numerabile di elementi

- L'insieme e' un intervallo continuo.
  es. Esperimento l'arrivo di un cliente in banca. $S_z = (0, \infty)$

Funzione di densita' di probabilita'

S_x di una variabile aleatoria Discreta → insieme numerabile

PMF → funzione di massa di x

La PMF (Probability mass function) e' una funzione matematica che descrive la distribuzione probabilita' di una variabile aleatoria discreta. Assegna a ogni possibile valore della variabile la probabilita' che essa assuma esattamente quel valore.

La PMF di X denotata $f_X(x)$ é definita come:
→ é maiuscola

$$f_X : S_X \to [0,1], \quad f_X(x) = P(X = x), \quad x \in S_X$$

proprietà:

$f_X(x) \geq 0 \ \forall x \in S_X$

$\sum_{x \in S_X} f_X(x) = 1$ → Somma totale 1 per rispetto a S_X

Per una PMF

media → $\mu = EX = \sum_{x \in S_X} x f_X(x)$

varianza → $\sigma^2 = \sum_{x \in S_X} (x-\mu)^2 f_X(x)$

dev. std → $\sigma = \sqrt{\sigma^2}$

CDF (cumulative distribution function)

La CDF descrive la probabilità che una variabile aleatoria (discreta o continua) assuma un certo valore minore o uguale a un certo x. É definita come:

$$F_X(t) = P(X \leq t), \quad -\infty < t < +\infty$$

Distribuzione Uniforme

Distribuzione discreta più comune → distribuzione discreta uniforme

- es. Dado

PMF → $f_X(x) = \frac{1}{n}$

media $\mu = \sum_{x=1}^{n} x f_X(x) = \frac{n+1}{2}$

varianza $\sigma^2 = \sum_{x=1}^{n} x^2 f_X(x) - \mu^2 = \frac{n^2-1}{12}$

Distribuzione Binomiale si basa sul:

- Processo di Bernoulli → un esperimento casuale che ha solo due possibili risultati

p → prob succ. singolo processo

PMF Bernoulli → $f_X(x) = p^x (1-p)^{1-x}$

media $\mu = p$

varianza $\sigma^2 = np(1-p)$

Distribuzione Binomiale

3 Principi:
- E' costituito da n processi di Bernoulli
- i processi sono fra loro indipendenti
- ogni processo ha la stessa probabilità p di successo

X e' la variabile che conta i successi degli n processi

PMF Binomiale $f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$

media $\mu = np$

varianza $\sigma^2 = n(n-1)p^2$

es. famiglia di 5 figli, sia $X = \{$numero di femmine$\}$ ...

Distribuzione di Poisson (abbiamo un intervallo di tempo)

- Variabile aleatoria continua
- Eventi rari
- λ → media dell'evento nell'intervallo di tempo $[0,1]$

$X \rightarrow \{$numero di eventi che accadono nell'intervallo $[0,t]\}$

PMF distr. di Poisson $f_X(x) = e^{-\lambda t} \dfrac{(\lambda t)^x}{x!}$ e → eulero 2,78

media e varianza coincidono $\mu = \sigma^2 = \lambda$

Distribuzioni continue → V. aleatorie continue (es. altezza)

Probability Density function (PDF)

Si definisce funzione di densità di probabilità (PDF), f_X associata a X la funzione

$$f_X : S_X \to \mathbb{R}$$

t.c.

$$P(X \in A) = \int_A f_X(x)\, dx$$

A come $[a,b]$

esempio

$$P(c \leq X \leq d) = \int_c^d f_X(x)\, dx$$

→ area sottesa dalla funzione f_X nell'intervallo $[c,d]$

Definizione

La PDF descrive la distribuzione di probabilità di una variabile aleatoria continua. A differenza della PMF (che assegna probabilità a valori discreti), la PDF rappresenta la densità di un punto ma non la probabilità diretta.

Proprietà PDF:

$f_X(x) \geq 0 \quad \forall x \in S_X$

$\int_{x \in S_X} f_X(x) = 1$

CDF → Cumulative Distribution Function → $F_X(t)$

La funzione di ripartizione o di distribuzione di X $F_X(t)$, descrive la probabilità che una variabile aleatoria (discreta o continua) assuma un valore minore o uguale a un certo punto t. E' definita come:

$$F_X(t) = P(X \leq t) = -\infty < t < +\infty$$

$F_X(t)$ e' non decrescente

Formule Generali

media $\mu = E X = \int_{x \in S_x} x \, f_x(x)$

varianza $\sigma^2 = \int_{x \in S_x} (x-\mu)^2 f_x(x)$

Distribuzione Continua uniforme

sull' intervallo $[a,b]$

$$f_x(x) = \frac{1}{b-a} \quad , \quad a < x < b$$

Sono eventi che seguono ELM

$$\mu = \int_a^b x \frac{1}{b-a} dx = \frac{b+a}{2}$$

Distribuzione normale

La variabile aleatoria casuale ha distribuzione normale

PDF → di una distribuzione normale

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ \frac{-(x-\mu)^2}{2\sigma^2} \right\} \quad , \quad -\infty < x < \infty$$

La distribuzione normale viene identificata da due parametri

↳ media μ e dev. standard σ

PDF della distr. norm. standard $(0,1)$

$$f_x(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{-x^2}{2} \right\}$$

CDF distr. norm.

$$F_x(t) = \int_{-\infty}^{t} f_x(x) \, dx \quad , \quad -\infty < t < \infty$$

Distribuzioni di variabili aleatorie di misura di un tempo di attesa

Variabile aleatoria che rappresenta un tempo prima che accada un evento (RARO)

Distribuzioni per variabili aleatorie di una distribuzione di un evento raro

Distribuzione esponenziale

$S_x\ (0,\infty)$

PDF distr. esponenziale $\qquad f_x(x) = \lambda e^{-\lambda} \qquad x>0,\ \lambda>0$

CDF distr. esponenziale $\qquad F_x(t) = 1 - e^{-\lambda t}$

media $\quad \mu = \dfrac{1}{\lambda}$

varianza $\quad \sigma^2 = \dfrac{1}{\lambda^2}$

Distribuzione gamma

↳ Generalizzazione di una distribuzione esponenziale.

PDF distrib. Gamma $\qquad f_x(x) = \dfrac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \quad , x>0,\ \lambda>0,\ a>0$

media $\quad \mu = \dfrac{a}{\lambda}$

varianza $\quad \sigma^2 = \dfrac{a}{\lambda^2}$

Distribuzione Chi quadro

↳ è una distribuzione di probabilità continua utilizzata in statistica per verificare ipotesi, stimare varianze e confrontare le frequenze osservate con quelle stesse.

PDF $\qquad f_x(x) = \dfrac{1}{\Gamma(p/2)\,2^{p/2}}\, x^{\frac{p}{2}-1}\, e^{-x/2} \quad , x>0,\ p>0$

$$\Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x}\, dx$$

Distribuzione di t-Student

↳ è una distribuzione di probabilità continua che governa il rapporto tra due variabili aleatorie, la prima con distribuzione normale standard e la seconda al quadrato, segue una distribuzione chi quadrato.

PDF distr. t-student $\qquad f_x(x) = \dfrac{\Gamma[(r+1)/2]}{\sqrt{r\pi}\,\Gamma(r/2)} \left(1 + \dfrac{x^2}{r}\right)^{-(r+1)/2} \quad , x \in \mathbb{R}$

$T = \dfrac{Z}{\sqrt{V/r}}$

$Z \sim N(0,1)$ (normale std.)

$V \sim \chi^2(r)$ (chi-quadrato con r gradi di libertà indipendente da Z)

Riassunto

Variabile aleatoria discreta

Caso generale

$$\mu = \sum_{x \in S_X} x f_X(x)$$

$$\sigma^2 = \sum_{x \in S_X} (x-\mu)^2 f_X(x)$$

Distribuzione uniforme $f_X(x) = \frac{1}{n}$

$$\mu = \sum_{x \in S_X} x f_X(x) \to \sum_{x \in S_X} x \frac{1}{n} = \dots = \frac{n+1}{2}$$

$$\sigma^2 = \sum_{x \in S_X} (x-\mu)^2 f_X(x) \to \sum_{x \in S_X} (x-\mu)^2 \frac{1}{n} = \dots = \frac{n^2-1}{12}$$

Distribuzione di Bernoulli ($p \to$ probabilità successo singolo evento)

$$f_X(x) = p^x (1-p)^{1-x}$$

$$\mu = p$$

$$\sigma^2 = np(1-p)$$

Distribuzione binomiale (n processi di Bernoulli)

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\mu = np$$

$$\sigma^2 = n(n-1)p^2$$

stats. binom. (pmf / mean / var)

Distribuzione di Poisson

intervallo di tempo $[0, t]$

$\lambda \to$ media dell'evento nell'intervallo di tempo

$$f_X(x) = e^{-\lambda t} \frac{(\lambda t)^x}{x!}$$

$$\mu = \sigma^2 = \lambda$$

Distribuzioni continue