SISTEMA ASSIOMATICO DI KOLMOGOROV

$\mathcal{F}$ σ-algebra è un insieme delle parti di Ω tale che:

- $\emptyset, \Omega \in \mathcal{F}$
- se $A \in \mathcal{F}$, $A^c \in \mathcal{F}$
- se $(A_n)_{n \in \mathbb{N}} \subset \mathcal{F}$, $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

PROBABILITÀ

Data la coppia $(\Omega, \mathcal{F})$, la probabilità è una funzione $\mathbb{P}: \mathcal{F} \longrightarrow [0,1]$ tale che

$\mathbb{P}(\Omega) = 1$

se $(A_n)_{n \in \mathbb{N}} \subset \mathcal{F}$ a due a due disgiunti

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$$

La terna $(\Omega, \mathbb{P}, \mathcal{F})$ è lo spazio di prob.

PROPRIETÀ

- $\mathbb{P}(\emptyset) = 0$ e $\mathbb{P}(\Omega) = 1$
- $\mathbb{P}\left(\bigcup_{i=1}^{n} A_n\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$
- $0 \leq \mathbb{P}(A) \leq 1$
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- se $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
- se $A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
- inclusione - esclusione

CASO DI UNO SPAZIO FINITO

↓

se gli eventi sono *equiprobabili* allora si parla di

↓

PROBABILITÀ UNIFORME

↓

$$\mathbb{P}(A) = \frac{\#A}{\#\Omega} = \frac{\text{casi favorevoli}}{\text{casi possibili}}$$

↓

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

Formula di disintegrazione:

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$$

INSIEMI

$B = (B \setminus A) \cup (A \cap B)$

$A = (A \setminus B) \cup (A \cap B)$

$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$

Leggi di De Morgan :

- $(F \cup G)^c = F^c \cap G^c$
- $(F \cap G)^c = F^c \cup G^c$

Leggi distributive :

- $(F \cap G) \cup H = (F \cup H) \cap (G \cup H)$
- $(F \cup G) \cap H = (F \cap H) \cup (G \cap H)$

2 EVENTI SI DICONO:

INDIPENDENTI

se il fatto che si verifichi o meno il primo non altera la probabilità che si verifichi il secondo. In questo caso si ha

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

esempio:

estrazioni con reinserimento

Oss ① L'indipendenza a 2 a 2 non garantisce l'indipendenza

② L'indipendenza dell'intersezione di tutti gli eventi non garantisce l'indipendenza

③ $A_1, \dots, A_n$ indipendenti allora $A_1^c, \dots, A_n$ indipendenti
(vale per tutte le combinazioni di A_i o A_i^c es: $A_1, A_2^c, A_3, A_4^c, A_5^c, \dots, A_n$)

DIPENDENTI

se il fatto che si verifichi o meno il primo altera la probabilità che si verifichi il secondo. In questo caso si ha:

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

"probabilità di A sapendo B"

esempio:

estrazioni senza reinserimento

Oss ① $\mathbb{P}(B|B^c) = \emptyset$

② $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

③ Se $\mathbb{P}(A|B) = \mathbb{P}(A)$ allora $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ cioè A,B indip.

④ $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$
$= \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)$
formula di disintegrazione

⑤ $\mathbb{P}[A|B] = \mathbb{P}[B|A]\dfrac{\mathbb{P}[A]}{\mathbb{P}[B]}$ formula di Bayes

ESPERIMENTO A PROVE RIPETUTE INDIPENDENTI

o schema di Bernoulli

si ripetono delle prove o esperimenti n volte

ogni prova può avere due esiti: successo / insuccesso

es: lancio di una moneta:
testa = successo, croce = insuccesso

Il successo avviene con probabilità $p \in [0,1]$

es: lancio di una moneta, $p = \frac{1}{2}$

Il numero di successi $S_n \in \{0, \dots, n\}$

$$\boxed{P[k \text{ successi su } n \text{ prove}]} = \sum_{\substack{J \subset \{1,\dots,n\} \\ \#J = k}} P[\text{successo in } J, \text{insuccesso in } J^c]$$

$$= \sum_{\substack{J \subset \{1,\dots,n\} \\ \#J = k}} p^k (1-p)^{n-k}$$

$$\boxed{= \binom{n}{k} p^k (1-p)^{n-k}} = \mathcal{B}_p(k,n)$$

distribuzione binomiale

TEOREMA DI DE MOIVRE-LAPLACE

↓

si ripetono delle prove o esperimenti n volte

↓

Il successo avviene con probabilità $p \in [0,1]$

↓

Il numero di successi è definito $S_n \in \{0, \dots, n\}$

↓

si ha $\forall a, b \in \mathbb{R}$, con $a < b$

$$\boxed{P\left[np + a\sqrt{np(1-p)} \leq S_n \leq np + b\sqrt{np(1-p)} \right] \xrightarrow{n \to \infty} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy}$$

esempio: Se volessimo sapere $P[470 \leq S_{1000} \leq 530]$ scriveremmo:

$$470 = 1000 \cdot \frac{1}{2} + a\frac{\sqrt{1000}}{2} = 500 + 5a\sqrt{10} \longrightarrow 5a\sqrt{10} = -30 \longrightarrow a = -\frac{6}{\sqrt{10}} = -1,9$$

$$530 = 1000 \cdot \frac{1}{2} + b\frac{\sqrt{1000}}{2} \longrightarrow 5b\sqrt{10} = 30 \longrightarrow b = 1.9$$

In questo modo ho ottenuto a e b e posso usare il Teorema di DML

LEGGE DEI GRANDI NUMERI

↓

n prove ripetute indipendenti

$p \in (0,1)$ prob. di successo, S_n il numero di successi

↓

la frequenza dei successi tende a p per $n \to \infty$ per ogni $\varepsilon > 0$

$$\boxed{\mathbb{P}\left[\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right] \longrightarrow 0 \quad \text{per } n \to \infty}$$

la probabilità di non avere neanche un successo dopo n prove è $(1-p)^n$

VARIABILI ALEATORIE

Una variabile aleatoria è una funzione $X : \Omega \to S$ dove S è un insieme arbitrario

esempio:

lancio di un dado : $\Omega = \{1, \dots, 6\}$ $\mathbb{P}$ uniforme $S = \mathbb{R}$ $X : \Omega \to \mathbb{R}$ $X(\omega) = \omega$

Alla variabile aleatoria è legata la

LEGGE

La legge è la coppia $(S_X, \mathbb{P}_X)$ con, $S_X = X(\Omega)$ e $\mathbb{P}_X = \mathcal{P}(S_X) \to \mathbb{R}$

N.B. $\mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A)) = \mathbb{P}(X \in A)$. S_X può essere implicito.

ad ogni variabile aletoria è associata una

DENSITÀ DISCRETA

$p_X : S_X \to \mathbb{R}$ $p_X(x) = \mathbb{P}[X = x]$

esempio: lancio di un dado

$X : \Omega \to S = \{1, \dots, 6\}$ $x \in S$ $p(x) = \frac{1}{6}$

COMPOSIZIONE DI V.A.

$X : \Omega \to S$ $f : \Omega \to S'$ $f(X) : \Omega \to S'$

UGUAGLIANZA DI V.A.

$X : \Omega \to S$ $Y : \Omega' \to S'$ sono uguali se $\Omega = \Omega'$, $\mathbb{P} = \mathbb{P}'$
$(\Omega, \mathbb{P})$ $(\Omega', \mathbb{P}')$ e se $\mathbb{P}[X = Y] = 1$

l'uguaglianza implica uguaglianza in legge ma non vale il viceversa

$\Longrightarrow$ X e Y v.a. sono uguali in legge se $S_X = S_Y$ $p_X = p_Y$ (oppure $\mathbb{P}_X = \mathbb{P}_Y$)

ESEMPI DI V.A.

Bernoulli

$X \in \{0,1\}$ $\quad p_X(1) = p \quad p_X(0) = 1-p$

Quando si usa? Quando si ha un esperimento con due possibili esiti: successo (=1) / insuccesso (=0)

Binomiale

$S_X = \{0,1,\dots,n\}$, $K \in \{0,\dots,n\}$

$p_X(K) = \binom{n}{K} p^K (1-p)^K$

Quando si usa? Quando si hanno n prove ripetute indipendenti con due esiti (es K successi su n prove)

$X \overset{leg.}{=} X_1 + \dots + X_n$

Bernoulli di parametro p

Geometrica

$p \in [0,1]$ $\quad S_X = \mathbb{N} \setminus \{0\}$

$p_X(K) = p(p-1)^{K-1}$

Quando si usa?

Per calcolare il primo successo in un esperim. a prove ripetute indip.

Poisson

$\lambda > 0 \quad S_X = \mathbb{N}$

$p_X(K) = e^{-\lambda} \frac{\lambda^K}{K!}$

Quando si usa?

Per gli eventi rari

Binomiale negativa

$K \geq 1$, $p \in [0,1]$, $n \geq 1$

$p_X(n) = \binom{n+K-1}{n} p^K (1-p)^n$

Quando si usa?

Per calcolare il numero di prove necessarie per ottenere K successi

DISTRIBUZIONI CONGIUNTE

$X_1 : \Omega \to S_1, \dots, X_n : \Omega \to S_n$, $X = (X_1, \dots, X_n) : \Omega \longrightarrow S_1 \times \dots \times S_n$

Chiameremo <u>distribuzione congiunta</u> la legge di $(X_1, \dots, X_n)$

$\downarrow$

DISTRIBUZIONI MARGINALI

$X = (X_1, \dots, X_n)$ su $(\Omega, \mathbb{P})$ con distribuzione congiunta $\mathbb{P}_X$

Chiameremo distribuzioni marginali le leggi delle sottofamiglie di $(X_1, \dots, X_n)$, in particolare le leggi dei singoli $X_1, \dots, X_n$

- $p_{X_1}, \dots, p_{X_n}$ densità discrete marginali

$$p_{X_1}(x) = \sum_{(x_2, \dots, x_n)} p_X(x, x_1, \dots, x_n)$$

$\vdots$

$$p_{X_n}(x) = \sum_{(x_1, \dots, x_{n-1})} p_X(x_1, \dots, x_{d-1}, x)$$

INDIPENDENZA DI V.A.

X, Y v.a. su $(\Omega, \mathbb{P})$ sono indipendenti se $\forall A_1 \subset S_1, A_2 \subset S_2$ $\mathbb{P}[X \in A_1, Y \in A_2] = \mathbb{P}[X \in A_1]\mathbb{P}[Y \in A_2]$

si estende a famiglie finite: $\mathbb{P}[X_1 \in A_1, \dots, X_n \in A_n] = \mathbb{P}[X_1 \in A] \cdots \mathbb{P}[X_n \in A_n]$ se $X_1, \dots, X_n$ indip.

Sono fatti equivalenti:
- $X_1, \dots, X_n$ v.a. indipendenti
- $p_{(X_1,\dots,X_n)}(x_1, \dots, x_n) = p_{X_1}(x_1) \cdots p_{X_n}(x_n)$ per ogni $x_1 \in S_1, \dots, x_n \in S_n$

Esempio n gettoni numerati da 1 a n estratti

con reinserimento $\downarrow$ INDIPENDENTI

senza reinserimento $\downarrow$ NON INDIPENDENTI

VALORE ATTESO

Data una v.a. reale $X : \Omega \to \mathbb{R}$ su $(\Omega, \mathbb{P})$ e una densità discreta p associata a $\mathbb{P}$, se $\sum_{\omega \in \Omega} |X(\omega)|\, p(\omega) < \infty$ (X integrabile) oppure $X > 0$ possiamo definire il valore atteso come

$$\boxed{\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\, p(\omega)}$$

insieme alla **MEDIANA** (m_x tc $\mathbb{P}[X \leq m_x] \geq \frac{1}{2}$ e $\mathbb{P}[X \geq m_x] \leq \frac{1}{2}$)

e alla **MODA** (m_{o_x} tc $p_x(x) \leq p_x(m_{o_x})\ \forall x$)

è un **INDICATORE DI CENTRALITÀ**

PROPRIETÀ

- $a \in \mathbb{R} \Rightarrow \mathbb{E}[a] = a$
- $a \in \mathbb{R}$, X v.a. $\Rightarrow \mathbb{E}[aX] = a\mathbb{E}[X]$
- $\mathbb{P}[X=0] = 1 \Rightarrow \mathbb{E}[X] = 0$
- $\mathbb{P}[X \geq 0] = 1$ e $\mathbb{E}[X] = 0 \Rightarrow \mathbb{P}[X = 0] = 1$
- $\mathbb{P}[X = Y] = 1$, X int. $\Rightarrow$ Y int. e $\mathbb{E}[X] = \mathbb{E}[Y]$
- X, Y int/pos e $\mathbb{P}[X \leq Y] = 1 \Rightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$
- X, Y int $\Rightarrow$ X+Y int e $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

X, Y indipendenti, $\mathbb{E}[X], \mathbb{E}[Y] < \infty$

Allora $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$

Valore medio binomiale

Possiamo scrivere una v.a. binomiale di parametri np come $X_1 + \dots + X_n$ con $X_1 \dots X_n$ indip. Poiché

$X_i = \begin{cases} 0 & 1-p \\ 1 & p \end{cases}$ si ha $\mathbb{E}[X_i] = 1 \cdot p + 0(1-p) = p$

$\Rightarrow \mathbb{E}[X_1 + \dots + X_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n] = np$

VARIANZA

Momenti : Se X è una v.a. reale

se $p \in \mathbb{R}$ il momento assoluto di ordine p di X è $\mathbb{E}[|X|^p]$

se $p \in \mathbb{Z}$ e se $\mathbb{E}[|X|^p] < \infty$ allora X ha momento di ordine p uguale a $\mathbb{E}[X^p]$

Sia X una v.a. dotata di momento secondo $(\mathbb{E}[X^2] < \infty) \rightarrow$ $\boxed{Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]}$

$\downarrow$

binomiale: $np(1-p)$

$$\overset{\shortparallel}{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}$$

DEVIAZIONE STANDARD

$$\boxed{sd(X) = \sqrt{Var(X)}}$$

COVARIANZA

- bilineare
- simmetrica

X, Y v.a. reali su $(\Omega, \mathbb{P})$ dotate di momento secondo

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

$$\downarrow$$

$$Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\rho(X, Y) = \frac{Cov(X, Y)}{sd(X)\, sd(Y)}$$ coeffic. di correlazione

VARIABILI ALEATORIE A VALORI INTERI

$X : \Omega \to \mathbb{N}$ o $\mathbb{Z}$

formula di convoluzione

X, Y indipendenti e intere allora $p_{X+Y}(n) = \sum_{j \in \mathbb{Z}} p_X(j) p_Y(n-j) \quad n \in \mathbb{Z}$

funzione generatrice delle probabilità

X v. a. a valori in $\mathbb{N}$ $\quad g_X(t) = \mathbb{E}[t^X] = \sum_{n=0}^{\infty} p_X(n) t^n$

PROPRIETÀ

$g_X = g_Y$ in un intorno di 0 $\Leftrightarrow$ sono uguali in legge

$\mathbb{E}[X(X-1)] = \lim_{t \to 1^-} g_X''(t)$

X, Y indipendenti $\Rightarrow g_{X+Y} = g_X g_Y$

$\mathbb{E}[X] = \lim_{t \to 1^-} g_X'(t)$

ESEMPI DI FUNZIONI

Bernoulli

X v.a. di Bernoulli di param. p

$$g_X(t) = (1-p) + pt$$

Binomiale

X binomiale (n,p)

$$\boxed{g_X(t)} = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} t^k = \sum_{k=0}^{n} \binom{n}{k} (pt)^k (1-p)^{n-k} =$$

$$\boxed{= (tp + 1-p)^n}$$

Poisson

$X \sim \text{Poisson}(\lambda)$

$$p_X(n) = e^{-\lambda} \frac{\lambda^n}{n!} \qquad n \geq 0$$

$$\boxed{g_X(t)} = \mathbb{E}[t^X] = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} t^n$$

$$\boxed{= e^{\lambda(t-1)}}$$

$$\boxed{\mathbb{E}[X]} = g_X'(1) = \dots \boxed{= \lambda}$$

$$\boxed{\mathrm{Var}(X)} = \dots \boxed{= \lambda}$$

Geometrica

X geometrica di parametro p

$$\boxed{g_X(t)} = \sum_{n=1}^{\infty} t^n (1-p)^{n-1} = tp \sum_{n=1}^{\infty} (t(1-p))^{n-1} = tp \sum_{n=0}^{\infty} (t(1-p))^n$$

$$\boxed{= \frac{tp}{1-t(1-p)}}$$

$$\boxed{\mathbb{E}[X]} = g_X'(1) = \dots \boxed{= \frac{1}{p}}$$

$$\boxed{\mathrm{Var}(X)} = \dots \boxed{= \frac{1-p}{p^2}}$$

FUNZIONE GENERATRICE DEI MOMENTI

X v.a. reale:

$\phi_X : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ $\quad \phi_X(t) = \mathbb{E}[e^{tX}]$

$\phi_X(0) = 1$ ma può accadere $\phi_X(t) = +\infty$ $\forall t \neq 0$

- ϕ_X analitica (in ogni aperto in cui assume valori finiti)

"•" valgono se ϕ_X (ed ev. ϕ_Y) $< \infty$ in un intorno di 0

- $\phi_X^{(n)}(0) = \mathbb{E}[X^n]$ $\forall n \geq 0$
- X, Y indipendenti allora $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$
- $\phi_X = \phi_Y$ in un intorno di 0 $\iff X \overset{\text{legge}}{=} Y$

TEOREMI LIMITE

__disuguaglianza di Markov__

$$\mathbb{P}[X \geq a] \leq \frac{1}{a} \mathbb{E}[X]$$

__disuguaglianza di Chebychev__

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq a] \leq \frac{1}{a^2} \mathrm{Var}(X)$$

CONVERGENZA

Si dice che $(X_n)_{n \geq 0}$ converge in probabilità a X

se $\forall \varepsilon > 0$ $\boxed{\lim_{n \to \infty} \mathbb{P}[|X_n - X| \geq \varepsilon] = 0}$

LEGGE DEBOLE DEI GRANDI NUMERI

$X_1, \dots, X_n$ indip. e con uguale distribuzione

$m = \mathbb{E}[X_1]$, $S_n = X_1 + \dots + X_n$ si ha

$$\boxed{\frac{S_n}{n} \longrightarrow m \quad \text{in probabilità}}$$

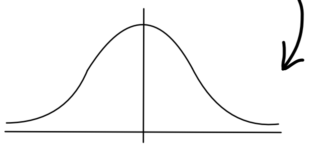
TEOREMA DEL LIMITE CENTRALE

$X_1, \dots, X_n$ v.a. indipendenti, ugual distribuzione, media comune e varianza comune

$S_n = X_1, \dots, X_n$, $\forall a < b$ si ha

$$\boxed{\mathbb{P}\left[a \leq \frac{S_n - \mathbb{E}[S_n]}{\mathrm{sd}(S_n)} \leq b\right] \xrightarrow{n \to \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}y^2} \, dy}$$

"Converge" ad una gaussiana



Il TLC è un'estensione del Teorema di De Moivre - Laplace

Si riferisce ad una LEGGE

Il TLC ha valore UNIVERSALE

DISUGUAGLIANZA DI CONCENTRAZIONE

Se $X_1, \dots, X_n$ sono indipendenti e con legge di Bernoulli(p) $p \in (0,1)$ posto $S_n = X_1 + \dots + X_n$ allora per ogni $\varepsilon > 0$ esiste $H(p,\varepsilon) > 0$ tale che $\mathbb{P}\left[\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right] \leq 2e^{-nH(p,\varepsilon)}$

MODELLI CONTINUI

Perché ne abbiamo bisogno?

Poiché in uno spazio di probabilità discreto Ω non possiamo definire una v.a. X per cui $X(\Omega) = I$ con $I \in \mathbb{R}$, poiché $X(\Omega)$ è sempre <u>al più</u> numerabile ($\subseteq \mathbb{N}$). Potremmo voler definire una v.a. il cui insieme dei valori sia più che numerabile (es. tempo $\to \mathbb{R}$)

Cosa cambia?

La nozione di probabilità su un insieme generale Ω non necessariamente numerabile è la stessa definita in precedenza <u>MA</u> è definita su una σ-algebra di $\mathcal{F} \subsetneq \mathcal{P}(\Omega)$

NON vale la densità finita discreta!

Notazioni

Ω sarà chiamato <u>spazio campionario</u> e gli elementi di $\mathcal{F}$ saranno chiamati <u>eventi</u>. Le proprietà della σ-algebra (le "stesse" viste per uno spazio di probabilità discreto) servono a garantire che tramite unione, intersezione e complementare si abbiano ancora eventi.

Variabili aleatorie

Per definire una v.a. non basta più una generica funzione $X : \Omega \to S$ con S insieme qualunque. Si richiede che anche l'insieme di arrivo S si munito di una σ-algebra $\mathcal{S}$

Definizione Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e $(S, \mathcal{S})$ uno spazio misurabile. Una funzione $X: \Omega \to S$ è una variabile aleatoria se per ogni $C \in \mathcal{S}$ si ha $\{X \in C\} := X^{-1}(C) \in \mathcal{F}$

σ-algebra di Borel

Se $S = \mathbb{R}^d$ sceglieremo sempre la σ-algebra di Borel o dei Boreliani $\mathcal{S} = \mathcal{B}(\mathbb{R}^d)$, la più piccola σ-algebra che contiene gli aperti.

Se $S = \mathbb{R}$ e $\mathcal{S} = \mathcal{B}(\mathbb{R})$ allora sono equivalenti:

- X è una variabile aleatoria
- $X^{-1}((-\infty, x]) \in \mathcal{F} \quad \forall x \in \mathbb{R}$

Legge di una v.a.

$(\Omega, \mathcal{F}, \mathbb{P})$ spazio di probabilità, $(S, \mathcal{S})$ spazio misurabile, $X: \Omega \to S$ v.a.
la legge $\mathbb{P}_X$ di X è la misura di probabilità $\mathbb{P}_X : S \to \mathbb{R}$ definita da:

$$\boxed{\mathbb{P}_X[A] = \mathbb{P}[X \in A] = \mathbb{P}[\{\omega \in \Omega \mid X(\omega) \in A\}] \quad \text{con } A \in \mathcal{S}}$$

Osservazioni sulla legge

1. NON può essere caratterizzata dalla densità discreta
2. costruzione canonica: $\tilde{\mathbb{P}}$ prob. su $(S, \mathcal{S})$, allora $\Omega = S$, $\mathcal{F} = \mathcal{S}$, $\mathbb{P} = \tilde{\mathbb{P}}$, $X: \Omega \to S$ def. da $X(\omega) = \omega$ è una v.a. $(X^{-1}(A) = A)$ con legge $\mathbb{P}_X = \tilde{\mathbb{P}}$
3. composizione di una v.a. con una funz. misurabile rimane una v.a.

↳ X v.a. su $(\Omega, \mathcal{F}, \mathbb{P})$ a valori $(S, \mathcal{S})$ e $\varphi: S \to S'$ $(S', \mathcal{S}')$ misurabile ($\Rightarrow \varphi^{-1}(B) \in \mathcal{S}$ $\forall B \in \mathcal{S}'$). Se $B \in \mathcal{S}' \Rightarrow \varphi^{-1}(X)(B) = X^{-1}(\varphi^{-1}(B)) \in \mathcal{F}$ $\Rightarrow \varphi(X)$ v.a.

4. Uguaglianza in legge: $X \overset{\text{legge}}{=} Y$ se $\mathbb{P}_X = \mathbb{P}_Y$. La legge può essere caratterizzata attraverso la funzione generatrice dei momenti

5. Uguaglianza di v.a.: $X = Y$ con prob. 1 $\mathbb{P}[X = Y] = 1$

6. leggi congiunte marginali: $X: \Omega \to S_1 \times S_2$ $(S_1, \mathcal{S}_1)$ $(S_2, \mathcal{S}_2)$ $X = (X_1, X_2)$ $\mathbb{P}_X$ legge cong.
$\mathbb{P}_{X_1}[A] = \mathbb{P}_X[A \times S_2]$ $A \in \mathcal{S}_1$

7. indipendenza. $(\Omega, \mathcal{F}, \mathbb{P})$ $(X_i)_{i \in I}$ $X_i: \Omega \to S_i$ $(S_i, \mathcal{S}_i)$ $(X_i)_{i \in I}$ indipendenti se per ogni $J \subset I$ finiti, per ogni $A_i \in \mathcal{S}_i$ $i \in J$ $\boxed{\mathbb{P}\left[\bigcap_{i \in J} \{X_i \in A_i\}\right] = \prod_{i \in J} \mathbb{P}[X_i \in A_i]}$

FUNZIONE DI RIPARTIZIONE

X v.a. su $(\Omega, \mathcal{F}, \mathbb{P})$ la funzione di ripartizione F_X di X è $F_X(x) = \mathbb{P}[X \leq x]$ $x \in \mathbb{R}$

proprietà

- F_X è monotona crescente
- F_X è continua a destra
- $\lim_{x \to \infty} F_X(x) = 1$; $\lim_{x \to -\infty} F_X(x) = 0$

Se $F: \mathbb{R} \to \mathbb{R}$ possiede una di queste 3 prop. allora esiste uno spazio $(\Omega, \mathcal{F}, \mathbb{P})$ e su di esso una v.a. X tc $F_X = F$. La f. di ripartizione caratterizza la distribuzione

Osservazioni

① F_X funzione di ripartizione, $a<b$, $F_X(b)-F_X(a) = \mathbb{P}[X\in(-\infty,b]] - \mathbb{P}[X\in(-\infty,a]] =$

$= \mathbb{P}[\{X\in(-\infty,b]\}\setminus\{X\in(-\infty,a]\}] = \mathbb{P}[X\in(a,b]]$

② Discontinuità della funzione di ripartizione:

X v.a., F_X funzione di ripartizione, allora se $x\in\mathbb{R}$ $\lim_{y\to x^-} F_X(y) = \mathbb{P}[X<x]$

le discontinuità di F_X sono i valori $x\in\mathbb{R}$ tc $\mathbb{P}[X=x]>0$

③ X, Y v.a. reali $\Rightarrow$ X, Y indip. $\Leftrightarrow$ $F_{(X,Y)}(x,y) = F_X(x)\,F_Y(y)$

CASO MULTIDIMENSIONALE

X v.a. su $\mathbb{R}^d$ $X=(X_1,\dots,X_d)$ $F_X:\mathbb{R}^d\to\mathbb{R}$ è definita $F_X(x)=\mathbb{P}[X_1\le x_1,\dots,X_d\le x_d]$ $\forall x=(x_1,\dots,x_d)$

proprietà

- F_X monotona crescente rispetto a ogni variabile
- F_X è continua a destra rispetto a ogni variabile
- per ogni $i=1,\dots,d$ $\lim_{x_i\to-\infty} F_X(x_1,\dots,x_d)=0$
- per ogni $i=1,\dots,d$ $\lim_{x\to+\infty} F_X(x_1,\dots,x_d) = F_{X^{(i)}}(x^{(i)})$
  dove $X^{(i)} = (X_1, X_2,\dots,X_{i-1}, X_{i+1},\dots,X_d)$ e
  $x^{(i)} = (x_1,\dots,x_{i-1},x_{i+1},\dots,x_d)$
- F_X identifica la distribuzione, nel senso che se X, Y v.a.
  $\mathbb{P}_X = \mathbb{P}_Y \Leftrightarrow F_X = F_Y$

VALORE ATTESO

(modello continuo)

$(\Omega, \mathcal{F}, \mathbb{P})$, X v.a. reale

- se $X \geq 0 \Rightarrow \mathbb{E}[X] = \sup \{ \mathbb{E}[Y]$ tc Y v.a. reale discreta tc $0 \leq Y \leq X \}$
- se poi $\mathbb{E}[|X|] < \infty$, posto $X_+ = \max(X, 0)$ e $X_- = \min(-X, 0) \Rightarrow \mathbb{E}[X] = \mathbb{E}[X_+] - \mathbb{E}[X_-] \rightarrow \boxed{\begin{array}{l} |X| = X_+ + X_- \\ X = X_+ - X_- \end{array}}$

proprietà

omogeneità

$\mathbb{E}[aX] = a\mathbb{E}[X]$

linearità

$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

positività

- $X \leq Y$ con prob. 1 $\Rightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$
- $X \geq 0$ con prob. 1 e $\mathbb{E}[X] = 0 \Rightarrow X = 0$ con prob. 1

X, Y indipendenti e integrabili/positive $\Rightarrow X \cdot Y$ integrabile/positiva e $\mathbb{E}[X \cdot Y] = \mathbb{E}[X]\mathbb{E}[Y]$

varianza, covarianza e correlazione

Cose che funzionano:

Markov, $a\mathbb{1}_{\{X \geq a\}} \leq X$,

$\mathbb{E}[\mathbb{1}_{\{X \geq a\}}] = \mathbb{P}[X \geq a]$, TLC

Non valgono le proprietà associate alla densità discreta

VARIABILI ALEATORIE (ASSOLUTAMENTE) CONTINUE

densità continua:

Una funzione $f : \mathbb{R} \to \mathbb{R}$ tale che:

- $f \geq 0$ $\quad$ ($f_X(x) \geq 0$ per q.o. $x \in \mathbb{R}$)
- $\int_{\mathbb{R}} f(x)\,dx = 1$

Una v.a. è assolutamente continua se esiste una densità continua f_X tale che:

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

In tal caso se $A = (-\infty, x]$

$$F_X(x) = P[X \in (-\infty, x]] = \int_{-\infty}^{x} f_X(x)\,dx$$

se poi $a < b$, $A = [a, b]$ (o (a,b), $[a,b)$, $(a,b]$)

$$F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx$$

Osservazioni:

X ass. continua,

$P[X = x] = 0 \quad \forall x \in \mathbb{R}$

Sotto opportune condizioni $F_X' = f_X$

CASO MULTIDIMENSIONALE

Una densità è una funzione $f : \mathbb{R}^d \to \mathbb{R}$ tale che

- $f(x) \geq 0$ per q.o. $x \in \mathbb{R}^d$
- $\int_{\mathbb{R}^d} f(x)\,dx = 1$

Una v.a. $X : \Omega \to \mathbb{R}^d$ è ass. continua se esiste una densità f_X t.c. $P[X \in A] = \int_A f_X(x)\,dx \quad A \in \mathcal{B}(\mathbb{R}^d)$

X, Y con densità $f_X, f_Y \Rightarrow X, Y$ indip se

$$f_{(X,Y)}(x, y) = f_X(x)\,f_Y(y) \quad \text{per q.o. } (x, y)$$

In particolare se $x = (x_1, \dots, x_d) \in \mathbb{R}^d$

$$F_X(x) = P[X_1 \leq x_1, \dots, X_d \leq x_d]$$
$$= P[X \in (-\infty, X_1) \times (-\infty, X_2) \times \dots \times (-\infty, X_d)]$$
$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_d}$$

Si ottiene $\partial x_1 \dots \partial x_d F(x_1, \dots, x_d) = f_X(x_1, \dots, x_d)$

DENSITÀ MARGINALI

$X : \Omega \to \mathbb{R}^d$ v.a. $X = (X_1, \dots, X_d)$ con densità f_X allora per ogni $J \subset \{1, \dots, d\}$ la v.a. $(X_i)_{i \in J} : \Omega \to \mathbb{R}^{\#J})$ ha densità. In particolare $\forall i = 1, \dots, d$ la v.a. X_i ha densità:

$$f_{X_i}(x) = \int_{\mathbb{R}^{d-1}} f_X(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_d)\, dx_1 \dots dx_{i-1}\, dx_{i+1} \dots dx_d$$

DENSITÀ CONDIZIONALE

(X, Y) ha densità $f_{(X,Y)}$; la densità condizionale di Y dato X è:

$$f_{Y|X}(y|x) = \begin{cases} \dfrac{f_{(X,Y)}(x,y)}{f_X(x)} & \text{se } f_X(x) \neq 0 \\ 0 & \text{altrimenti} \end{cases}$$

tipo Bayes

Osserviamo che $\underbrace{f_X(x)}_{\text{se } = 0} = \underbrace{\int_{\mathbb{R}} f_{(X,Y)}(x,y)\, dy}_{\Rightarrow \quad = 0}$

VALORE ATTESO

Se X v.a. reale con densità f_X, allora $\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x)\, dx$ se $\int_{\mathbb{R}} |x| f_X(x)\, dx < \infty$

nel discreto: $\sum x\, p_X(x)$ se $\sum |x|\, p_X(x)$

In generale, se X è v.a. su $\mathbb{R}^d$ con densità f_X e se $\varphi : \mathbb{R}^d \to \mathbb{R}$ misurabile allora $\varphi(X)$ è una v.a. e $\mathbb{E}[\varphi(X)] = \int_{\mathbb{R}^d} \varphi(x) f_X(x)\, dx$ se $\int_{\mathbb{R}} |\varphi(x)|\, f_X(x)\, dx < \infty$

FORMULA DEL CAMBIO DI VARIABILE

X v.a. su $\mathbb{R}^d$ con densità f_X, supponiamo $f_X = 0$ fuori da A

Sia $\varphi : A \to B$ diffeomorfismo, allora la v.a. $Y = \varphi(X)$ ha densità f_Y

$$f_Y(y) = f_X(\varphi^{-1}(y)) \, |\det D\varphi^{-1}(y)| \qquad y \in B$$

determinante della matrice D

Osservazioni

1. se $d = 1$ → se siamo in $\mathbb{R}$ $\quad f_Y(y) = f(\varphi^{-1}(y)) \left| \frac{d\varphi^{-1}}{dy}(y) \right|$ $\quad \frac{1}{\varphi'(\varphi^{-1}(y))}$

2. Se X è una v.a. con densità f_X e se $f_X = 0$ su un aperto B, allora $\mathbb{P}[X \in B] = 0$

FORMULA DI CONVOLUZIONE

Se X, Y sono v.a. indipendenti con densità f_X, f_Y, allora $X+Y$ ha densità

$$f_{X+Y}(z) = f_X * f_Y(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x)\, dx$$

SOMMA DI GAUSSIANE

Se X è $N(m_1, \sigma_1^2)$ e $Y = N(m_2, \sigma_2^2)$ e se X, Y indipendenti allora $X + Y = N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

$N(m_1, \sigma_1^2)$ = v.a. Gaussiana di media m_1 e varianza σ_1^2

↳ TLC: se $X_1 = \dots = X_n = N(m, \sigma^2) \to S_n = X_1 + \dots + X_n = N(nm, n\sigma^2)$ e $\frac{S_n - \mathbb{E}[S_n]}{sd(S_n)} = N(0,1)$

ESEMPI

Distribuzione uniforme

densità: $f_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$

valore atteso: $\exists$ se $\int_a^b \frac{1}{b-a} |x| \, dx < \infty$

$$\mathbb{E}[X] = \int_a^b \frac{1}{b-a} x \, dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

momento secondo:

$$\mathbb{E}[X^2] = \int_a^b \frac{1}{b-a} x^2 dx = \frac{a^2+ab+b^2}{3}$$

varianza: $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{(b-a)^2}{12}$

F. generatrice dei momenti: $\phi_X(t) = \mathbb{E}[e^{tx}] = \int_a^b \frac{1}{b-a} e^{tx} dx = \frac{e^{tb} - e^{ta}}{t(b-a)}$

Funzione di ripartizione: $F_X(x) = \begin{cases} 0 & \text{se } x \le a \\ \frac{x-a}{b-a} & \text{se } x \in (a,b) \\ 1 & \text{se } x \geq b \end{cases}$ se $x \in \mathbb{R}$

Esponenziale

densità: $f_X(x) = \lambda e^{-\lambda} \mathbb{1}_{[0,\infty)}$

funzione di ripartizione: $F_X(x) = \begin{cases} x \le 0 & 0 \\ x > 0 & 1-e^{-\lambda x} \end{cases}$

Valore atteso: $\mathbb{E}[X] = \frac{1}{\lambda}$

varianza: $Var(X) = \frac{1}{\lambda^2}$

proprietà di mancanza di memoria: $\mathbb{P}[X > x+y \mid X > x] = e^{-\lambda y} = \mathbb{P}[X > y]$

$$\mathbb{P}[X > x+y] = \mathbb{P}[X > x] \, \mathbb{P}[X > y]$$

Funzione generatrice dei momenti:

$$\phi_X(t) = \mathbb{E}[e^{tx}] = \int_0^{+\infty} \lambda e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda - t}$$

Gaussiana

densità: $f_X(x) = \frac{e^{-\frac{1}{2\sigma^2}(x-m)^2}}{\sqrt{2\pi\sigma^2}}$

valore atteso: $\mathbb{E}[X] = m$

varianza: σ^2

Cauchy

densità: $f_X(x) = \frac{1}{\pi(1+x^2)}$

non ha valore atteso finito

Gamma

densità: $f_X(x) = \frac{1}{\Gamma(r)} \lambda^r x^{r-1} e^{-\lambda x} \mathbb{1}_{[0,\infty)}$

• $r=1 \to$ esponenziale

valore atteso: $\mathbb{E}[X] = \frac{r}{\lambda}$

varianza: $Var(X) = \frac{r}{\lambda^2}$

Casi particolari:

• $r \in \mathbb{Z} \Rightarrow G(r,\lambda) =$ {somma di r v.a. esp. indip. di param λ}

• Caso χ^2 = Gamma$\left(\frac{n}{2}, \frac{1}{2}\right) \to \mathbb{E}[X] = n$, $Var(X) = 2n$

• Student (n gradi di libertà): $f_X(x) = c_n\left(1 + \frac{x^2}{n}\right)^{-\frac{1}{2}(n+1)}$, $\mathbb{E}[X] = 0$, $Var(X) = \begin{cases} \frac{n}{n-2} & \text{se } n > 2 \\ +\infty & \text{se } n = 1,2 \end{cases}$

STATISTICA

È una disciplina di analisi dei dati

Statistica descrittiva

Si occupa della sintesi dei dati

media campionaria $\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n)$

varianza campionaria $\bar{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2$

covarianza campionaria $\text{cov} = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})(y_k - \bar{y})$

Statistica inferenziale

→ Quella che studieremo

Analisi dei dati effettuata attraverso un modello

↓

A partire dai dati traiamo delle conclusioni

↓

Si studia un fenomeno attraverso un campione rappresentativo e dei modelli

esempio: sondaggi, opinioni di voto

→ è sconsigliabile sondare ogni volta 60 mln di persone

a volte è impossibile ripetere un esperimento un grande numero di volte

esempio significativo: lancio di una moneta n volte

$x_1, \ldots, x_n \quad x_i \in \{0,1\}$

fenomeno rilevante: prob p di ottenere 1

Terminologia:

- popolazione: interezza del fenomeno che stiamo studiando

  ↓

  insieme di unità statistiche che si vogliono esaminare

  es: i 60 mln di italiani

- Campione: la parte di popolazione che andremo a studiare

  es: 1200 soggetti maggiorenni

quantile

Un quantile di ordine $\alpha \in [0,1]$ è il valore q_α tc

$$\phi(q_\alpha) = \alpha$$

dove $\phi(q_\alpha)$ è la f. di ripartizione di una Gaussiana standard

$$\boxed{\phi(q_\alpha) = \int_{-\infty}^{q_\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx}$$

problema 1

qual è il valore di p

1. la LGN è un teorema limite
2. Il valore stimato ha delle fluttuazioni

si trova una "soluzione" attraverso la LGN

MA

NON è un valore esatto

problema 2

trovare un range di valori per p

problema 3

possiamo "mettere alla prova" i valori che abbiamo trovato di p?

MODELLI STATISTICI

Θ insieme di valori dei parametri. Chiameremo modello statistico la terna $(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta \in \Theta}$

- Campione: Un campione di n individui estratto da una popolazione di legge m_θ è una famiglia $X_1, \dots, X_n$ di v.a. indipendenti e con legge m_θ
- Statistica: Una statistica, all'interno di un modello statistico, è una v.a. che non dipende dal parametro
- Stimatore: Una statistica che è in funzione del campione → Vogliamo sapere quando uno stimatore è corretto e quando no.

TEORIA DEGLI STIMATORI

Stimatore corretto
(non distorto)

$(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta \in \Theta}$ modello statistico, $X_1, \dots, X_n$ campione, U è uno stimatore corretto se

$$\boxed{\mathbb{E}_\theta[U] = \theta \quad \forall \theta \in \Theta}$$

Stimatore consistente
(asintoticamente non distorto)

$(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta \in \Theta}$ modello statistico, $\forall n$ sia dato un campione di numerosità n estratto da una popolazione di legge m_θ, dove $(m_\theta)_{\theta \in \Theta}$ una famiglia di prob. La successione di stimatori $(U_n)_{n \geq 1}$ è consistente se per ogni $\theta \in \Theta$ e per ogni $\varepsilon > 0$

$$\boxed{\lim_{n \to \infty} \mathbb{P}_\theta[|U_n - \theta| \geq \varepsilon] = 0}$$

RISCHIO QUADRATICO

U stimatore di θ

$$R_\theta(U) = \mathbb{E}_\theta[(U-\theta)^2]$$

Introduce un ordinamento parziale tra gli stimatori

↓

U e V stimatori, U è preferibile a V se $R_\theta(U) \leq R_\theta(V) \quad \forall \theta \in \Theta$

VEROSIMIGLIANZA

= probabilità di ottenere esattamente ciò che si è stimato/dedotto dall'esperimento

$$L_\theta(x_1, \dots, x_n) = \begin{cases} p_\theta(x_1) \cdot \ldots \cdot p_\theta(x_n) & \text{caso discreto} \\ f_\theta(x_1) \cdot \ldots \cdot f_\theta(x_n) & \text{caso continuo} \end{cases}$$

STIMA DI MASSIMA VEROSIMIGLIANZA

Dato un campione di numerosità n di legge m_θ, sia L_θ la funzione di verosimiglianza.

Uno stimatore U è <u>uno</u> stimatore di massima verosimiglianza se

$$\boxed{L_U(X_1, \dots, X_n) = \sup_{\theta \in \Theta} L_\theta(X_1, X_2, \dots, X_n)}$$

→ vogliamo massimizzare L_θ in funzione del parametro

Osservazione: Per la LGN, una media empirica che è corretta è anche consistente

MODELLI ESPONENZIALI

Supponiamo $\Theta \subset \mathbb{R}$

caso discreto

Sia m_θ probabilità sugli interi con densità p_θ tale che esistano $c_\theta > 0$ e due funzioni

$T: \mathbb{N} \to \mathbb{R}$ e $g: \mathbb{N} \to \mathbb{R}$

tali che

$$\boxed{p_\theta(k) = c_\theta\, g(k)\, e^{\theta T(k)}} \quad k \in \mathbb{N}$$

caso continuo

Sia m_θ probabilità su $\mathbb{R}$ con densità f_θ tale che esistano $c_\theta > 0$ e due funzioni misurabili

$T: \mathbb{R} \to \mathbb{R}$ e $g: \mathbb{R} \to \mathbb{R}$

tali che

$$\boxed{f_\theta(x) = c_\theta\, g(x)\, e^{\theta T(x)}} \quad x \in \mathbb{R}$$

ESEMPI

popolazione geometrica

$$p_\theta(k) = p(1-p)^{k-1} = \underbrace{\frac{p}{1-p}}_{c_p}\, e^{\overbrace{k}^{T(k)} \log(1-p)} \qquad g = 1$$

non è importante che ci sia $e^{\theta T(k)}$, ci può essere funzione iniettiva/invertibile (come log)

popolazione esponenziale

$$f_\theta(x) = \lambda e^{-\lambda x} \mathbb{1}_{[0,+\infty)}(x) = \underbrace{\lambda}_{c_\lambda}\, \underbrace{\mathbb{1}_{[0,+\infty)}(x)}_{g(x)}\, e^{\lambda \overbrace{(-x)}^{T(x)}}$$

Teorema $\Theta \subset \mathbb{R}$ aperto, $(m_\theta)_{\theta\in\Theta}$ modello esponenziale, consideriamo $\forall n$ un campione di numerosità n estratto da una popolazione di legge m_θ e supponiamo

- lo stimatore θ_n di massima verosimiglianza esiste in Θ per ogni n

Allora $(\theta_n)_{n\geq 1}$ è consistente → si può "riadattare" al caso discreto

Lemma $(X_n)_{n\geq 1}$ v.a, $X_n \xrightarrow{\text{in prob.}} \ell$, $\ell\in\mathbb{R}$ $\Rightarrow$ $\varphi:\mathbb{R}\to\mathbb{R}$ continua, $\varphi(X_n) \xrightarrow{\text{in prob.}} \varphi(\ell)$

REGIONI DI FIDUCIA

Dati un modello statistico $(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta\in\Theta}$ e un numero $\alpha\in(0,1)$, un sottoinsieme aleatorio $D\subset\Theta$ è una regione di fiducia per il parametro θ al livello $1-\alpha$ se per ogni θ

$$\boxed{\mathbb{P}_\theta[\theta\in D] = \mathbb{P}_\theta[\omega : \theta\in D(\omega)] \geq 1-\alpha}$$

- Quantità pivotale: Una quantità pivotale è una funzione del campione e del parametro tale che
  1. È invertibile rispetto a θ (dato il campione)
  2. la sua legge non dipende dal parametro

- quantile: Sia F una funzione di ripartizione. Se $\alpha\in(0,1)$, il quantile di ordine α per F è ogni numero q tale che
  1. $F(q)\geq\alpha$
  2. $F(q^-)\leq\alpha$ dove $F(q^-) = \lim_{y\to q^-} F(y)$

OSS: Ogni mediana è un quantile di ordine $1/2$

ESEMPI

popolazione di Bernoulli

$$\mathbb{P}_p\left[\frac{|X_1+\dots+X_n-np|}{\sqrt{np(1-p)}}\leq q\right]\approx 2\phi(q)-1$$

$$\phi(q)=\int_{-\infty}^{q}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx$$ f. di ripartizione di $N(0,1)$

Dato α : $2\phi(q)-1=1-\alpha \rightarrow \phi(q)=1-\frac{\alpha}{2}$

ovvero $q=q_{1-\frac{\alpha}{2}}$ quantile di una Gaussiana standard

$$\sqrt{np(1-p)}\leq\frac{\sqrt{n}}{2}\longrightarrow p\in\left(\overline{X}-\frac{\sqrt{n}}{2}q_{1-\alpha}\right)$$

con probabilità $\approx 2\Phi(q_{1-\frac{\alpha}{2}})-1=1-\alpha$

popolazione Gaussiana

$N(m,\sigma^2)$, supponiamo di conoscere σ

$$Z=\frac{X_1+\dots+X_n-nm}{\sigma\sqrt{n}}$$ $\quad m=\overline{X}-\frac{\sigma}{\sqrt{n}}Z$

↳ ha legge $N(0,1)$ (Gaussiana standard)

↳ $m\in\left(\overline{X}-\frac{\sigma}{\sqrt{n}}q_{1-\frac{\alpha}{2}},\ \overline{X}+\frac{\sigma}{\sqrt{n}}q_{1-\frac{\alpha}{2}}\right)$

con probabilità $1-\alpha$

TEST STATISTICI

A cosa servono?

Vogliamo mettere alla prova il modello che abbiamo formulato per il nostro esperimento

$(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta \in \Theta}$ modello statistico, Θ insieme dei valori dei parametri, $X_1, \dots, X_n$ campione

1. Formulare l'ipotesi: partizioniamo Θ

valori di θ compatibili con l'ipotesi $= \Theta_0 \quad \cup \quad \Theta_1 =$ valori di θ incompatibili con l'ipotesi

chiaramente $\Theta_0 \cap \Theta_1 = \emptyset$

2. Pianificazione: stabilisce quali risultati portano a rigettare o meno l'ipotesi nulla

→ Identificare un evento: regione critica o regione di rifiuto

ERRORI

- 1^a specie: rigettare un'ipotesi nulla corretta
  ↳ insiste sui parametri dell'ipotesi nulla } si può controllare
- 2^a specie: non rigettare un'ipotesi nulla non corretta
  ↳ insiste sui parametri dell'ipotesi alternativa } non si può controllare → ma possiamo cercare di individuarlo

LIVELLO DI UN TEST

Dati C una regione critica e $\alpha \in [0,1]$

il test di regione critica C ha livello α se

$$\underbrace{\sup_{\theta \in \Theta} \mathbb{P}_\theta[C]}_{\text{il più grande errore di } 1^a \text{ specie}} = \alpha$$

OK anche $\leq$

→ se α è molto piccolo è improbabile compiere un errore di 1^a specie

POTENZA DI UN TEST

La potenza di un test di regione critica C è

la funzione $\pi_C : \Theta_1 \to [0,1]$ definita come

$\pi_C(\theta) = \mathbb{P}_\theta[C]$ → capacità di accorgersi che l'ipotesi nulla non è corretta noto il valore del param. θ

3. Stabilire il livello α del test e, tra i test di livello α, scegliere quello di potenza massima

LEGAME TRA REGIONI DI FIDUCIA E REGIONI CRITICHE

di test con ipotesi semplici

$\Theta_0 = \{\theta_0\}$, $\Theta_1 = \Theta \setminus \{\theta_0\}$

CASO 1 — Viceversa — CASO 2

CASO 1

D regione di fiducia al livello $1-\alpha$

$\mathbb{P}_\theta[\theta_0 \in D] = \mathbb{P}_\theta[\omega : \theta_0 \in D(\omega)] \geq 1-\alpha$

allora $C = \{\omega : \theta_0 \notin D(\omega)\}$ è una regione critica di livello α

$\mathbb{P}_\theta[C] = 1 - \mathbb{P}_\theta[\theta_0 \in D] = 1-(1-\alpha) = \alpha$

CASO 2

$(C_{\theta_0})_{\theta_0 \in \Theta}$ regioni critiche, cioè $\forall \theta_0$ C_{θ_0} è regione critica di livello α per il test $\begin{cases} \mathcal{H}_0 : \theta = \theta_0 \\ \mathcal{H}_1 : \theta \neq \theta_0 \end{cases}$

Allora $D(\omega)$ è regione di fiducia al livello $1-\alpha$

$D(\omega) = \{\theta \in \Theta : \omega \notin C_\theta\}$

MODELLI A RAPPORTO DI VEROSIMIGLIANZA CRESCENTE

$\underbrace{(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta \in \Theta}}_{\text{modello statistico}}$, $\underbrace{(X_1, \dots, X_n)}_{\text{campione}}$, supponiamo $\Theta \subset \mathbb{R}$ intervallo. Il modello è a rapporto di verosimiglianza crescente rispetto ad una v.a. T se per ogni $\theta_1 < \theta_2$ $\dfrac{L_{\theta_2}(X_1, \dots, X_n)}{L_{\theta_1}(X_1, \dots, X_n)}$ è una funzione str. crescente di T

ESEMPIO popolazione di Bernoulli

$L_\theta(x_1, \dots, x_n) = \theta^{x_1 + \dots + x_n}(1-\theta)^{n-(x_1+\dots+x_n)}$ $\quad x_i \in \{0,1\}$, sia $\theta_1 < \theta_2$

Allora $\dfrac{L_{\theta_2}}{L_{\theta_1}} = \left(\dfrac{\theta_2}{\theta_1}\right)^{n\overline{X}} \left(\dfrac{1-\theta_2}{1-\theta_1}\right)^{n(1-\overline{X})}$ è str. crescente rispetto a $T = \overline{X}$

TEST UNILATERALE

$(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta \in \Theta}$ modello statistico, $(X_1, \dots, X_n)$ campione. Supponiamo $\Theta \subset \mathbb{R}$ modello a verosimiglianza crescente rispetto ad una v.a. T; consideriamo il test: $\mathcal{H}_0 : \theta \leq \theta_0$ $\quad \mathcal{H}_1 : \theta > \theta_0$ e $C = \{T > S\}$.

Allora
- $\sup_{\theta \leq \theta_0} \mathbb{P}_\theta(C) = \mathbb{P}_{\theta_0}(C)$
- il test con regione critica C è più potente di ogni altro test a livello $\mathbb{P}_{\theta_0}(C)$

$\longrightarrow$ Un analogo risultato vale per il test: $\mathcal{H}_0 : \theta \geq \theta_0$ $\quad \mathcal{H}_1 : \theta < \theta_0$

$\longrightarrow$ $(\Omega, \mathcal{F}, \mathbb{P}_\theta)_{\theta \in \Theta}$ $\quad (X_1, \dots, X_n)$, e $\Theta = \{\theta_0, \theta_1\}$, L funzione di verosimiglianza

Per $c > 0$, $C = \{L_{\theta_0} \leq c L_{\theta_1}\}$, ed il test $\mathcal{H}_0 : \theta = \theta_0$, $\mathcal{H}_1 : \theta \neq \theta_0$. Allora:

- $\mathbb{P}_{\theta_0}(C) \leq \mathbb{P}_{\theta_1}(C)$
- C è la regione critica più potente di ogni altro test di livello $\mathbb{P}_{\theta_0}(C)$

P-VALUE

=

soglia di rigetto

Esiste una famiglia $(C_\alpha)_{\alpha \in (0,1)}$ tale che: ① $\bigcup C_\alpha = \Omega$, ② $\bigcap C_\alpha = \emptyset$, ③ $\alpha < \alpha' \Rightarrow C_\alpha \subset C_{\alpha'}$ ④ C_α è una regione critica di livello α

Allora per ogni $\omega \in \Omega$ esiste un unico p t.c.:

- Se $\alpha < p$ allora $\omega \notin C_\alpha$ $\rightarrow$ ipotesi nulla non rigettata
- Se $\alpha > p$ allora $\omega \in C_\alpha$ $\rightarrow$ ipotesi nulla rigettata

ES. $\longrightarrow$ lancio di monete, $\mathcal{H}_0 : p = \frac{1}{2}$

$$\underline{\text{p-value}} = \mathbb{P}_{\frac{1}{2}}\left[\sqrt{n}\,\frac{|\overline{X} - \frac{1}{2}|}{1/2} > \sqrt{n}\,\frac{|\overline{x} - \frac{1}{2}|}{1/2}\right]$$

$\overline{x} \rightarrow$ dati dell'esperimento, $\overline{X} \rightarrow$ campione

POPOLAZIONI GAUSSIANE

$X_1, \dots, X_n$ hanno legge $N(m, \sigma^2)$

stimatori di massima verosimiglianza:

$$L_{(m,\sigma^2)}(x_1, \dots, x_n) = (2\pi)^{-\frac{n}{2}} \sigma^{-n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - m)^2} \longrightarrow \log L_{(m,\sigma^2)} = -\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - m)^2$$

Attraverso il calcolo differenziale otteniamo: $\hat{m} = \bar{X}$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$

$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - m)^2$ se la media è nota

Modello Gaussiano è esponenziale

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-m)^2} = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{m^2}{2\sigma^2}}}_{C_\theta} e^{\overbrace{(x, -x^2)}^{T(x)} \overbrace{\left(\frac{m}{\sigma^2}, \frac{1}{2\sigma^2}\right)}^{\ell(m,\sigma^2)}}$$

$\ell(m, \sigma^2)$ è invertibile e continua

- $X_1, \dots, X_n$ indipendenti e con legge $N(0,1)$, A matrice ortogonale. Allora $A(X_1, \dots, X_n)$ è ancora un vettore di v.a. indipendenti di legge $N(0,1)$
- $X_1, \dots, X_n$ indipendenti e con legge $N(m, \sigma^2)$, $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$, $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$

Allora:
- $\bar{X}$, S^2 sono indipendenti
- $\bar{X}$ ha distribuzione $N(m, \frac{\sigma^2}{n})$
- $(n-1)\frac{S^2}{\sigma^2}$ ha distribuzione $\chi^2(n-1)$
- $\frac{\bar{X} - m}{S}\sqrt{n}$ ha distribuzione di Student $t(n-1)$

Oss X, Y Gaussiane indip. $\Rightarrow$ $X+Y$ Gaussiana

POPOLAZIONI GAUSSIANE

VARIANZA NOTA

$(X_1, \dots, X_n)$, $N(m, \sigma^2)$ σ^2 noto

- intervalli di fiducia

- quantità pivotale : $Z = \sqrt{n}\,\dfrac{\bar{X} - m}{\sigma}$ legge $N(0,1)$

- regione di fiducia al livello $1-\alpha$:

$$\left\{ |Z| \leq q_{1-\frac{\alpha}{2}} \right\} \rightarrow m \in \left[\bar{X} - \frac{\sigma}{m} q_{1-\frac{\alpha}{2}} , \bar{X} + \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \right]$$

Z - test : $\mathcal{H}_0 : m = m_0$

test bilatero: $\mathcal{H}_1 : m \neq m_0$

regione critica a liv. α $\left\{ |Z| > q_{1-\frac{\alpha}{2}} \right\} \rightarrow \left\{ |\bar{X} - m| > \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \right\}$

p-value : $p = \mathbb{P}_{m_0}\left[\sqrt{n}\,\dfrac{(\bar{X} - m)}{\sigma} > \sqrt{n}\,\dfrac{|\bar{x} - m|}{\sigma} \right]$

test unilatero: modello a rapp. di verosimiglianza crescente rispetto a $\bar{X}$

$$\frac{L_{m_2}(x_1, \dots, x_n)}{L_{m_1}(x_1, \dots, x_n)} = \frac{e^{-\frac{1}{\sigma^2}\sum (x_i - m_2)^2}}{e^{-\frac{1}{\sigma^2}\sum (x_i - m_1)^2}} = e^{\frac{n}{\sigma^2}(m_2 - m_1)\bar{X} - \frac{n}{2\sigma^2}(m_2^2 - m_1^2)}$$

dunque una regione critica al livello α $\left\{ \bar{X} > m_0 + \frac{\sigma}{\sqrt{n}} q_{1-\alpha} \right\}$

VARIANZA NON NOTA

$(X_1, \dots, X_n)$, $N(m, \sigma^2)$, parametri (m, σ^2)

- intervalli di fiducia:

quantità pivotale $T = \sqrt{n}\,\dfrac{\bar{X} - m}{S} \rightarrow t(n-1)$

regione di fiducia: $\left\{ |T| \leq t_{n-1,\, 1-\frac{\alpha}{2}} \right\}$

densità Student : $c_n \left(1 + \dfrac{x^2}{n} \right)^{-\frac{n+1}{2}}$

TEST DI STUDENT:

test bilatero sulla media:

$\mathcal{H}_0 : m = m_0$, σ^2 qualsiasi

$\mathcal{H}_1 : m \neq m_0$, σ^2 qualsiasi

una regione critica di livello α

$\left\{ |T| > t_{n-1,\, 1-\frac{\alpha}{2}} \right\}$ $T = \sqrt{n}\,\dfrac{\bar{X} - m_0}{S}$

ha distr. di Student "decentrata"

test unilatero

$\mathcal{H}_0 : m \leq m_0$, σ^2 qualsiasi

$\mathcal{H}_1 : m > m_0$, σ^2 qualsiasi

regione critica liv. α

$\left\{ T > t_{n-1,\, 1-\alpha} \right\}$

VARIANZA

intervalli di fiducia per la media.

Quantità pivotali:

$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - m)^2$ ha distribuzione $\chi^2(n)$ (quantità pivotale se m ha valore noto)

$(n-1)\frac{S^2}{\sigma^2}$ ha legge $\chi^2(n-1)$

regione di fiducia $\left\{ \chi^2_{n-1,\frac{\alpha}{2}} \leq (n-1)\frac{S^2}{\sigma^2} \leq \chi^2_{n-1,1-\frac{\alpha}{2}} \right\}$

meglio: $\left\{ (n-1)\frac{S^2}{\sigma^2} \geq \chi^2_{n-1,\alpha} \right\}$ ovvero $\left\{ \sigma^2 \leq (n-1)\frac{S^2}{\chi^2_{n-1,\alpha}} \right\}$

test bilatero

$\mathcal{H}_0 : \sigma^2 = \sigma_0^2$, m qualsiasi

$\mathcal{H}_1 : \sigma^2 \neq \sigma_0$, m qualsiasi

(poco interessante)

test unilatero

$\mathcal{H}_0 : \sigma^2 \leq \sigma_0^2$, m qualsiasi

$\mathcal{H}_1 : \sigma^2 > \sigma_0^2$, m qualsiasi

$f(x) = c_m \, \sigma^{-(m+1)} \, x^{\frac{m-3}{2}} \, e^{-\frac{x^2}{2\sigma^2}}$

costante di normalizzazione

densità di una somma di m Gaussiane indipend. di media 0 e varianza σ^2

regione critica a livello α:

$$\left\{ (n-1)\frac{S^2}{\sigma_0^2} > \chi^2_{n-1,1-\alpha} \right\}$$