1 Introduzione

probabilità $\rightarrow$ misurare l'incertezza
statistica:

- descrittiva
- differenziale $\rightarrow$ campione casuale per stimare un esito

---

probabilità:

$$\frac{casi favorevoli}{casi totali}\ \underline{\textbf{SE}}\ \text{equiprobabili}$$

per contare i casi ci si appoggia alla combinatoria

---

partizione: separazione di A in sottoinsiemi senza elementi comuni

---

NB:

- $\wedge$ - and $\rightarrow$ $A \cap B = \{x | x \in A \wedge x \in B\}$
- $\vee$ - or $\rightarrow$ $A \cup B = \{x | x \in A \vee x \in B\}$

Principi della combinatoria:

1. A insieme, $\{E_i\}_{i=1}^{n}$ partizione di A $\rightarrow$ $\#A = \sum_{i=1}^{n} \#E_i$
   - A,B insiemi, AxB è l'insieme di coppie ordinate (a,b)
2. $\#(AxB) = \#A \cdot \#B \rightarrow \{A_i\}_{i=1}^{n} = \bigotimes_{i=1}^{n} A_i$
3. A,B, $\#(A \cup B) = \underline{\#A + \#B - \#(A \cap B)}$ (non perfetto) $\rightarrow$

$$\# \cup_{i=1}^{n} A_i = \sum_{i=1}^{n} \#A_i - \sum_{i<j} \#(A_i \cap A_j) + \sum_{i<j<k} \#(A_i \cap A_j \cap A_k) + \dots$$

$$\downarrow$$

$$\mathbf{+(-1)^{n+1}\ \# \cap_{i=1}^{n} A_i}$$

2 Permutazioni e anagrammi

Fattoriale $\rightarrow$ x! $=$ 9! $=$ 9·8·7·6· · · · ·2·1

NB: 0! $=$ 1

- "prendiamo" ha 9! anagrammi
- "anagramma" ha tre ripetizioni di a e due ripetizioni di m, quindi per calcolare i casi unici:

$$\frac{9!}{3!\cdot 2!}$$
$$\downarrow$$

per calcolare la probabilitàdegli elementi n, ma mi interessano solo k elementi allora:

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

se non sono interessato all'ordine, allora:

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

chiamato anche **coefficente binominiale**

Proprietà:

- $\binom{n}{k} = \binom{n}{n-k}$
- $\binom{n}{0} = \binom{n}{n} = 1$
- $\sum_{k=0}^{n} \binom{n}{k} = 2^{\mathrm{n}}$
- $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$

3 Esperimenti aliatori

Un esperimento si definisce **aliatorio** o casuale se con i dati iniziali il risultato è incerto. I risultati a 2 a2 incompatibili di un esperimento aliatorio sono chiamati **esiti**. Ω denota lo **spazio degli esiti**. Un **evento** è un osservabile di un esperimento aliatorio.

Una parte di Ω può essere considerata come famiglia:

$$\mathcal{F} \subseteq \mathrm{P}(\Omega)$$

Questa è definita come **algebra** se:

- $\Omega \in \mathcal{F}$
- se A$\in \mathcal{F}$ allora A^c$\in \mathcal{F}$
- se A,B$\in \mathcal{F}$, allora A$\cup$B$\in \mathcal{F}$

  potremmo scrivere anche $\{A_i\}_{i=1}^{n} \subseteq \mathcal{F}$ allora $\cup_{i=1}^{n} A_i \in \mathcal{F}$

Proprietà

- $\emptyset \in \mathcal{F}$
- se A,B$\in \mathcal{F}$ allora A$\cap$B$\in \mathcal{F}$
- se $\{A_i\}_{i=1}^{n} \subseteq \mathcal{F}$ allora $\cap_{i=1}^{n} A_i \in \mathcal{F}$
- se A,B $\in \mathcal{F}$, allora A$\cdot$B $\in \mathcal{F}$
- se A,B $\in \mathcal{F}$, allora A$\triangle$B $\in \mathcal{F}$

$\mathcal{F} \subseteq \mathrm{P}(\Omega)$ è una **tribù** se:

- $\Omega \in \mathcal{F}$
- A$\in \mathcal{F} \Rightarrow$ A^C$\in \mathcal{F}$
- per ogni famiglia <u>numerabile</u> $\{A_i\}_{i=1}^{+\infty} \subseteq \mathrm{P}(\Omega)$,
  allora $\cup_{i=1}^{+\infty} A_i \in \mathcal{F}$

NB: generalmente una tribù è un'alge se hanno elementi finiti

$\mathcal{F}$ tribù su Ω. Ogni E$\in \mathcal{F}$ (E è sottoinsieme di Ω) si dice **Evento**. I singoletti si chiamano **eventi elementari**. E si verifica se il risultato dell'esperimento appartiene ad E $\mathcal{F}$ tribù su Ω $(\Omega,\mathcal{F})$

Dati Ω, $\mathcal{F}$ tribù su Ω $(\Omega,\mathcal{F})$ si chiama **spazio probabilizzabile**.

$(\Omega,\mathcal{F})$, una funzione P:$\mathcal{F} \to \mathcal{R}$ si dice **funzione di probabilità** se:

- per ogni evento E P(E)$\geq$0
- P(Ω)=1
- data una famiglia numerabile $\{E_i\}_{i=1}^{+\infty}$ di eventi a 2 a 2 disgiunti:

$$P(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i) \text{ (additività)}$$

Proprietà delle probabilità

- $P(\emptyset) = 0$
- $E \in \mathcal{F}$ allora $P(E^c) = 1 - P(E)$
- $E, F \in \mathcal{F}$, $E \subseteq F \Rightarrow P(E) \leq P(F)$
  $E \in \mathcal{F}$ $P(E) \leq 1$
- $E, F \in \mathcal{F}$ $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
  $P(E \cup F) \leq P(E) + P(F)$
  $(E_i)_{i=1}^{n}$¡$\mathcal{F}$, $P(\cup_{i=1}^{n} E_i) = \sum_{k \in \mathcal{P}(\{1-n\})} (-1)^{\#k+1}\ P(\cap_{j \in k} E_j)$
  $(E_i)_{i=1}^{+\infty} \subset \mathcal{F}$, $P(\cup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} P(E_i)$
- (disuguaglianza di bonferrow)
  $\sum_{i=1}^{+\infty} P(E_i) - \sum_{i<j} P(E_i \cap E_j) \leq P(\cup_{i=1}^{+\infty} E_i) \leq \sum_{i=1}^{\infty} P(E_i)$

4 Probabilità condizionata

$(\Omega, \mathcal{F}, P)$, $E, F \in \mathcal{F}$ con $P(F) \neq 0$, allora la probabilità di E condizionale a F è:

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

Dato $(\Omega, \mathcal{F}, P)$ e due sotto tribù $\mathcal{F}_1, \mathcal{F}_2$ di $\mathcal{F}$ allora $\mathcal{F}_1 e \mathcal{F}_2$ sono indipendenti se se ogni elemento di $\mathcal{F}_1$ è indipendente da ogni elemento di $\mathcal{F}_2$

$$P(E_1 \cap E_2\ |\mathcal{F}) = P(E_1|\mathcal{F}) \cdot P(E_2|\mathcal{F})$$

5 funzione di probabilità $(\Omega, \mathcal{F}, P)$

1. Ω finito o numerabile

   Ω è dato

   $\mathcal{F} = \mathcal{P}(\Omega)$

   P: assegnamo ad ogni singoletto $(\omega \in \Omega)$ un probabilità tale che:

   P$(\omega)\geq$0

   $\sum$P(ω)=1

   A questo punto $\forall$ E $\in \mathcal{F}$ P(E):=$\sum_{\omega \in E}$P(ω)

2. Spazi prodotto

   considerando più ripetizioni di un esperimento o l'unione di più esperimenti: data una famiglia di sottoinsiemi di Ω dette $\mathcal{A}$. La tribù di $\mathcal{F}_{\mathcal{A}}$ generata da $\mathcal{A}$ come la più piccola tribù contenente $\mathcal{A}$

   $$\mathcal{F}_{\mathcal{A}} = \sigma(\mathcal{A}) = \cap\{\mathcal{G} : \mathcal{G} \text{ è tribù in } \Omega \text{ e } \mathcal{A} \subseteq \mathcal{G}\}$$

   quindi il prodotto $\Omega_1 \times \Omega_2$, la tribù sarà:

   $$\mathcal{F} = \mathcal{F}_{\mathrm{E}} \bigotimes \mathcal{F}_{\mathrm{E}} = \sigma(\{E_1 \times E_2 : E_1 \in \mathcal{F}_{\mathrm{E}},\, E_2 \in \mathcal{F}_{\mathrm{E}}\})$$
   $$(\Omega_1, \mathcal{F}_1, P_1),\ (\Omega_2, \mathcal{F}_2, P_2)$$
   $$\Downarrow$$
   $$\Omega = \Omega_1 \times \Omega_2$$
   $$\mathcal{F} = \mathcal{F}_{\mathrm{E}} \bigotimes \mathcal{F}_{\mathrm{E}} = \sigma(\{E_1 \times E_2 : E_1 \in \mathcal{F}_{\mathrm{E}},\, E_2 \in \mathcal{F}_{\mathrm{E}}\})$$
   $$\mathrm{P} : \mathrm{P}(\mathrm{E}_1 \times \mathrm{E}_2) = \mathrm{P}_1(\mathrm{E}_1) \cdot \mathrm{P}_2(\mathrm{E}_2)$$

   Quindi:
   con un numero finito di esperimenti $\{(\Omega_{\mathrm{i}}, \mathcal{F}_{\mathrm{i}}, P_{\mathrm{i}})\}_{\mathrm{i}\in \mathrm{I}}$ allora lo spazio prodotto ha forma:
   $\Omega = \bigotimes_{i\in I} \Omega_i$
   $\mathcal{F} = \bigotimes_{i\in I} \mathcal{F}_i = \sigma(\Pi_{i\in I}\ E_i : E_i \in \mathcal{F}_i$ e $\exists$n tc $\forall j \geq n$ E$_{\mathrm{j}}$=$\Omega_i)$
   P = $\bigotimes_{i\in I}$ P$_{\mathrm{i}}$ cioé P$(\Pi_{i\in I}$ E$_{\mathrm{i}})$ = $\Pi_{i\in I}$P$_{\mathrm{i}}$(E$_{\mathrm{i}}$)

6 Trasformazioni lineari di variabili aleatorie

07/04/

X variabile aleatoria con legge F_X. Se X è variabile aleatoria discreta:

$$\varphi_Y(y) = \sum_{x \in g^{-1}(\{y\})} \varphi_X(x)$$

Se X è variabile aleatoria assolutamente continua abbiamo 2 strategia:

1. Ricaviamo la legge di Y usando la forma di X e della funzione g
2. usiamo il teorema generale

Teorema del cambio di variabile

Sia X variabile aleatoria continua di densità f_X, sia $Y = g(x)$ con $g : \mathbb{R} \to \mathbb{R}$ continua a tratti t.c. $P(g(x) = 0) = 0$. Allora:

$$f_Y(y) = \sum_{x \in g^{-1}(\{y\})} \frac{f_X(x)}{|g^1(x)|}$$

7 Vettori aleatori

Dato uno spazio probabilizzabile $(\Omega, \mathcal{F}, P)$ consideriamo 2 variabili aliatorie $X, Y : \Omega \to \mathbb{R}^2$

Def: dati $(\Omega, \mathcal{F}, P)$ e X, Y variabili aleatorie su di esso si chiama **coppia di variabili aleatorie** o **variabile aleatoria doppia** o **2-vettore aleatorio**. La funzione $V : \Omega \to \mathbb{R}^2 : V(\omega) = (X(\omega/Y(\omega)))$. Il supporto del vettore aleatorio V:

$$\mathcal{R}_V = \mathcal{R}_{X,Y} = \mathcal{R}_X \times \mathcal{R}_Y = \{(x, y) \in \mathbb{R}^2 : x \in \mathcal{R}_X, y \in \mathcal{R}_Y\}$$

Def: Data (X, Y) coppia di variabili aleatorie, la sua funzione di ripartizione è:

$$F_{X,Y}((x, y)) = F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

$F_{X,Y}$ si chiama anche funzione di ripartizione **congiunta** di X e Y

Def: Data (X, Y) coppia di variabili aleatorie, chiameremo **funzione di ripartizione di X condizionata a Y** la funzione:

$$F_{X|Y}(x|y) := \frac{F_{X,Y}(x,y)}{F_Y(y)}$$

Def: Dato $(\Omega, \mathcal{F}, P)$ e due tribù $\mathcal{F}_1, \mathcal{F}_2 \subset \mathcal{F}, \mathcal{F}_1, \mathcal{F}_2$ sono indipendenti se lo sono le tribù $\sigma(x)$ *e* $\sigma(y)$ da esse generate

Prop: X, Y sono indipendenti se e solo se:

$$\forall (x, y) \in \mathbb{R}^2 \; F_{X|Y}(x|y) = F_X(x) F_Y(y)$$

Prop: X, Y sono indipendenti se e solo se:

$$\forall (x,y) \in \mathbb{R}^2 \ F_{X|Y}(x|y) = F_X(x) \ e \ F_{Y|X}(y|x) = F_Y(y)$$

8 Vettori aleatori discreti

Def: Siano X, Y variabili aleatorie discrete su $(\Omega, \mathcal{F}, P)$ chiamiamo **densità discrete congiunte** la funzione $\varphi_{X,Y} : \mathbb{R}^2 \to [0,1]$ definita:

$$\varphi_{X,Y}(x,y) = P(X = x, Y = y)$$

la **densità discreta di** X **condizionata a** Y è $\varphi_{X,Y}$ definità:

9 Vettori aleatori misti

caso speciale di va discreta e va continua

Modelli di variabili aleatorie discrete

10 Bernoulliane

Def:
Una variabile aleatoria di dice bernoulliana di parametro $p \in [0, 1]$ se la densità discreta è:

$$X \sim bin(1, p)$$

$$\varphi_x(x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \\ 0 & altrimenti \end{cases}$$

la cdf di una bernoulliana (p) è $F_x(x) = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$

11 Binominale

Def:
una variabile aleatoria discreta X è binominale di parametri n e p se è la somma di n bernulliane(p) indipendenti

$$X \sim bin(n, p)$$

$$\varphi_x(k) = \begin{cases} \binom{n}{k} p^k (1 - p)^{n-k} & k \in \{0, ..., n\} \\ 0 & altrimenti \end{cases}$$

12 Schema o processo di Bernoulli

21/04/

Dati infiniti esperimenti indipendenti e identicamente distribuiti

$$(X_i)_i \in \mathbb{N} \; iid \; X_i \sim bin(1,p)$$

$\Omega = \{0,1\}^{\mathbb{N}/\{0\}}$
Tribù $\mathcal{F}$ generata dai cilindri
P uguale al prodotto delle probabilità delle componenti

12.1 Cilindri

I cilindri sono sottoinsiemi $c \subseteq \Omega$ tali che esiste un $n \in \mathbb{N}/\{0\}$ e un vettore $v \in \{0,1\}^n$:

$$C = \{\omega \in \Omega : \omega_i = v_i \; 1 \leq i \leq n\}$$

Es:

- un successo seguito da due insuccessi:
  Cilindro: $n = 3 \; v = (1,0,0) \Rightarrow prob = p(1-p)^2$
- primo successo al k-esimo lancio:
  Cilindro: $(0,0,...,0_{k-1},1_k) \Rightarrow prob = (1-p)^{k-1}p$
- prob 3° lancio sia un successo:
  $(\cdot\cdot\cdot 1*) = (001) \cup (101) \cup (011) \cup (111)$
  $P(\cdot\cdot\cdot 1*) = \sum P(...) = (1-p)^2 p + 2(1-p)p^2 + p^3 = P(p + (1-p))^2$

13 Geometriche

Una varibile aleatoria ($T_1 := inf\{i \geq 1 : \omega_i = 1\}$) è una geometrica di parametro p $X \sim geom(p)$ se è l'istante precedente al primo successo in un processo di Bernoulli di parametro p

cdf di una geometria:

$$F_X(x) \begin{cases} 0 \quad x < 0 \\ \sum_{k=0}^{x} \varphi_x(k) = 1 - (1-p)^x \quad x \geq 0 \end{cases} \tag{1}$$

Assenza di memoria: $\forall n, k \in \mathbb{N} \quad P(x \geq n + k | X \geq n) = P(X \geq k)$
es:

$$(Y \geq 60 + 30 | Y \geq 60) = (Y \geq 30) = (1-p)^{30}$$

14 Binominali negative

T_n= istante dell'n-esimo successo

$T_1 := inf\{i \geq 1 : \omega_i = 1\}$
$T_{n+1} := inf\{i \geq T_n : \omega_i = 1\}$ $n \geq 1$

X è una variabile aleatoria binominale negativa (o di pascal) di parametri n e p se è il numero di insuccessi precedenti all'n-ennesimo successo di uno schema di bernoulli di parametro p $X \sim NB(n,p)$

$$pnk \in \mathbb{N}\varphi_x(k) \begin{cases} = P(x=k) = P(T_n = k+n) \\ = P(\omega_{n+k} = 1, \sum_{j=1} \omega_j = n-1) \\ = p\binom{k+n-1}{n-1}p^{n-1}(1-p)^k \end{cases} \Rightarrow \binom{k+n-1}{n-1}p^n(1-p)^k \tag{2}$$

15 Riproducibilità

22/04/

Una famiglia di variabili aleatorie si dice riproducibile se sommando 2 variabili aleatorie indipendenti appartenenti a quella famiglia abbiano ancora una variabile aleatoria della medesima famiglia

Prop: La famiglia delle binominali a parametro p fissato è riproducibile. Se $X \sim bin(n,p)$, $Y \sim bin(m,p)$, X e Y indipendenti allora:

$$X + Y \sim bin(n+m, p)$$

16 Ipergeometriche

Data un urna con n biglie bianche e n biglie nere, contiamo le bianche:
- con reimmissione abbiamo $bin(k, \frac{n}{m+n})$
- senza reimmissione usiamo un'ipergeometrica

Def: Si chiama ipergeometrica di parametri k, n, m la variabile aleatoria che conta il numero di bianche tra le estratte senza reimmissione

$$X \sim hyp(k,n,m)$$

$$\varphi_x(b) : \begin{cases} \frac{\binom{m}{b}\binom{n}{k-b}}{\binom{n+m}{k}} \\ 0 \quad altrimenti \end{cases}$$

Prop: Siano $\{a_i\}_{i\in\mathbb{N}},\ \{b_i\}_{i\in\mathbb{N}}$ interi non negativi che tendono in modo monotono a $+\infty$, $\lim_{i\to\infty} a_i = \lim_{i\to\infty} b_i = +\infty$ o tali che
$\lim_{i\to\infty} \frac{a_i}{b_i+a_i} = \alpha,\ \alpha \in [0,1]$, allora:

$$\frac{\binom{a_i}{k}\binom{b_i}{n-k}}{\binom{a_i+b_i}{n}} \to_{i\to\infty} \binom{n}{k}\alpha^k(1-\alpha)^{n-k}$$

17 Poisson

Def: X è variabile aleatoria di Poisson di parametro $\lambda > 0$ se:

$$\varphi_x(k) \begin{cases} \frac{\lambda^k}{k!} \cdot e^{-\lambda} & k \in \mathbb{N} \\ 0 & altrimenti \end{cases}$$

e si denota come $X \sim Pois(\lambda)$

Es:
in una partita di calcio vengono segnati 2.5 gol di media. X determina la probabilità di fare gol in un intervallo:
$\Rightarrow$ dividiamo 90' in 5 intervalli: $X \sim bin(5, 1/2)$
$\Rightarrow$ dividiamo in 20 intervalli: $X \sim bin(20, 1/8)$
$\Rightarrow$ dividiamo in 90 intervalli: $X \sim bin(90, 1/36)$

Questa successione tende a una variabile aleatoria di Poisson

Oss: Poisson viene a volte utilizzato come descrizione di una binomiale con n, p piccoli o grandi, non precisi

Prop: $\{p_n\}_n$ successione di numeri in $[0,1]$ tale che $\lim_{x\to\infty} n \cdot p_n = \lambda \in \mathbb{R}^+$ allora $\forall k \in \mathbb{N}$:

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

Prop: Le variabili aleatorie di Poisson sono riproducibili. $X \sim Pois(\lambda_1),\ Y \sim Pois(\lambda_2)$:

$$X + Y \sim Pois(\lambda_1 + \lambda_2)$$

18 Speranza/Valore atteso/Media

Caso variabile aleatoria discreta

Def:
La speranza di una variabile aleatoria discreta è il baricentro della sua distribuzione

$$E[X] = \textstyle\sum_{x \in \mathcal{R}_x}\ x \cdot \varphi_x(x)$$

Oss: se prendo un esito $Y = y$ nella mia tribù e considerò $P(.|Y)$

$$E[X|Y] = \textstyle\sum_{x \in \mathcal{R}_x}\ x \cdot P(X = x|Y = y) = \sum x \cdot \varphi_{x|y}(X|Y)$$

Teorema
Sia X una variabile aleatoria discreta con densità discreta φ_x e sia $Y = g(x)$, allora:

$$E[Y] = \textstyle\sum_{k \in \mathcal{R}_x}\ g(k) \cdot \varphi_x(x)$$

Teorema
Sia (X, Y) un vettore aleatorio discreto con densità congiunta $\varphi_{x,y}$ e sia $Z = g(x, y),\ g : \mathbb{R}^2 \to \mathbb{R}$, allora:

$$E[Z] = \textstyle\sum_{j \in \mathcal{R}_y} \sum_{k \in \mathcal{R}_x}\ g(k, j) \cdot \varphi_{xy}(k, j)$$

Prop 29/04/

il valore atteso possiede le seguenti proprietà:

- Linearità: Siano X, Y variabili aleatorie discrete e $a, b \in \mathbb{R}$, allora:

$$E[aX + bY] = aE[X] + bE[Y]$$

- Prodotto di variabili aleatorie indipendenti: Siano X, Y variabili aleatorie discrete e indipendenti, allora:

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

- Monotonia: X variabile aleatoria discreta, se $X \geq 0$ allora $E[X] \geq 0$. L'uguaglianza vale solo se $X \equiv 0$

Corollario: Se X, Y variabili aleatorie discrete tali che $P((X \geq Y) = 1)$ allora $E[X] \geq E[Y]$. In più se $E[X] = E[Y]$ Allora $X = Y$

Speranza di variabili aleatorie discrete note

Bernoulliane	$X \sim bin(1, p)$	$\Rightarrow$	$E[X] = p$
Binomiali	$X \sim bin(n, p)$	$\Rightarrow$	$E[X] = n \cdot p$
Poisson	$X \sim pois(\lambda)$	$\Rightarrow$	$E[X] = \lambda$
Ipergeometriche	$X \sim hyp(k, m, n)$	$\Rightarrow$	$E[X] = k \cdot \frac{m}{n+m}$
Geometriche	$X \sim geom(p)$	$\Rightarrow$	$E[X] = \frac{1-p}{p}$
Binomiali Negative	$X \sim NB(n, p)$	$\Rightarrow$	$E[X] = n \cdot \frac{1-p}{p}$

19 Variabile aleatoria assolutamente continua

X è variabile aleatoria assolutamente continua $P(X = a) = 0$

$$P(X \in [a, b]) = \int_a^b f_x(x)\, dx$$

Def

X variabile aleatoria assolutamente continua allora $E[X] = \int_{-\infty}^{+\infty} x \cdot f_x(x)\, dx$

Teorema
X variabile aleatoria assolutamente continua e $Y = g(x)$ allora:

$$E[Y] = \int_{\mathbb{R}} g(x) f_x(x)\, dx$$

Teorema
(X, Y) vettlre aleatorio assolutamente continuo e $g : \mathbb{R}^2 \to \mathbb{R}$ allora se $Z = g(x, y)$:

$$E[Z] = \int\int_{\mathcal{R}^2} \ g(x, y) f_{X,Y}(x, y)\, dx\, dy$$

Teorema
(X, Y) vettlre aleatorio misto con X discreta e Y assolutamente continua e densità ibrida (o mista) $f_{X,Y}$. Se $g : \mathbb{R}^2 \to \mathbb{R}$ e $Z = g(X, Y)$ allora:

$$E[Z] = \sum_{x \in \mathcal{R}_x} \int_{-\infty}^{+\infty} \ g(x, y) f_{X,Y}(x, y)\, dx\, dy$$

Proprietà
Il valore atteso gode delle seguenti proprietà:

- Linearità
- prodotto di variabili aleatorie indipendenti
- Monotonia

20 momenti di una variabile aleatoria

Def
per ogni $n \in \mathbb{N}/\{0\}$ si dice momento n-esimo di X variabile aleatoria il numero reale $E[X^n]$.
Si dice *momento centrale* di X il numero reale $E[(X - E[X])^n]$
Il momento secondo centrale di X prende il nome di *varianza* di X: $Var[X]$

NB: la varianza misura la larghezza della distribuzione, al quadrato

Prop

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

Proprietà della varianza

- $Var[X] \geq 0, \ Var[X] = 0 \Leftrightarrow X \equiv const$
- $a, b \in \mathbb{R}, \ Var[aX + b] = a^2 Var[X]$

Prop

Siano X, Y variabili aleatorie indipendeti, allora $Var[X + Y] = Var[X] + Var[Y]$

21 Deviazione standard

Chiamiamo *deviazione standard* σ_x di una variabile aleatoria X, il numero $\sigma_x = \sqrt{Var[X]}$
Sia $Y = \alpha X$ allora $\sigma_Y = \alpha \cdot \sigma_X$

Varianza di modelli discreti noti

Bernoulliane	$X \sim bin(1,p)$	$\Rightarrow$	$Var[X] = p(1-p)$
Binomiali	$X \sim bin(n,p)$	$\Rightarrow$	$Var[X] = n \cdot p(1-p)$
Geometriche	$X \sim geom(p)$	$\Rightarrow$	$Var[X] = \frac{(1-p)}{p^2}$
Binomiali neagative	$X \sim NB(p)$	$\Rightarrow$	$Var[X] = n \cdot \frac{(1-p)}{p^2}$
Poisson	$X \sim pois(\lambda)$	$\Rightarrow$	$Var[X] = \lambda$

22 Diseguaglianze

Disuguaglianza di Markov
Sia X una variabile aleatoria non negativa, allora $\forall a > 0$:

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Disuguaglianza di Chebychev
Sia X variabile aleatoria. Per ogni $a > 0$:

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

Oss: posso prendere $a\sigma_x$ al posto di a e la funzione diventa:

$$P(|X - E[X]| \geq a\sigma_x) \leq \frac{\sigma_x^2}{a^2\sigma_x^2}$$

23 Covarianza e correlazione

Date X, Y variabili aleatorie, chiamiamo *covarianza* di X e Y il numero:

$$Cpv[X,Y] = E[(X - E[X])(Y - E[Y])]$$

Oss: Se $X = Y$ Allora

$$Cov[X,Y] = E[(X - E[X])^2] = Var[X]$$

$$Conv[X,Y] = E[X \cdot Y] - E[X] \cdot E[Y]$$

Proprietà

- Se X e Y sono indipendenti allora $Cov[X, Y] = 0$
- La covarianza è simmetrica: $Cov[X, Y] = Cov[Y, X]$

Def
Date X e Y variabili aleatorie tali che $Cov[X, Y] = 0$, allora si defininiscono *scorrelate*

Proprietà

- $Var[X + Y] = Var[X] + Var[Y] + 2 \cdot Conv[X, Y]$
- In ogni componente la Cov è lineare:
  $Cov[aX + bY, Z] = a \cdot Cov[X, Z] + b \cdot Cov[Y, Z]$
- La convarianza è bilineare:
  Dati $(a_i)_{i=1}^n, (b_j)_{j=1}^m$ vettori reali, $(X_i)_{i=1}^n, (Y_j)_{j=1}^m$ vettori aleatori:
  $Cov[\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j] = \sum_{i,j} a_i b_j \; Conv[X_i, Y_j]$

Def
La matrice $Cov[X_i, Y_i]$ si chiama *matrice di covarianza* ($\sum$)

$$Cov[\vec{a} \cdot \vec{X}, \vec{b} \cdot \vec{Y}] = \vec{a}^t \sum \vec{b}$$

Prop

$$-\sqrt{Var[X]Var[Y]} \leq Cov[X, Y] \leq \sqrt{Var[X]Var[Y]}$$

Def
Date X e Y variabili aleatorie, chiamiamo *Correlazione* il numero:

$$\mathcal{P}(X, Y) = Corr[X, Y] = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}}$$

24 Mediane

06/05/

Si dice *Mediane* di una variabile aleatoria X un numero mx tale che:

$$P(X \leq mx) = P(X \geq mx)$$

Oss:

$$P(X \leq mx) = F_X(mx)$$
$$P(X \geq mx) = 1 - F_X(mx) - P(X = mx)$$

cioé, per X assolutamente continua mx tale che:

$$F_X(mx) = 1 - F_X(mx) \text{ ossia } F_X(mx) = 1/2 \qquad mx \in F_X^{-1}(\{1/2\})$$

Per X assolutamente continua esiste una mediana, ma può essere non unica

Oss: Se X è una variabile aleatoria discreat, la mediana può non essere unica o non esistere

Def
Chiamiamo mediana impropria un reale $\widetilde{m}x$ tale che:

$$P(X \leq \widetilde{m}x) \geq 1/2 \quad \text{e} \quad P(X \geq \widetilde{m}x) \geq 1/2$$

25 Quantile

Dato X con legge F_X e $p \in (1, 0)$, chiamiamo *p-quantile* il numero reale $Q_X(p)$:

$$Q_X(p) = inf\{x \in \mathbb{R} : \ F_X(x) \geq p\}$$

Oss: la funzione quantile $Q_X : p \rightarrow Q_x(p)$ $((0, 1) \rightarrow \mathbb{R})$ è qualcosa di simile all'inversa della F_X

26 Moda

Chiamiamo *moda* di una variabile aleatoria X un numero $x \in \mathcal{R}_x$ tale che:

- Se X è discreta, φ_x è massima in x, cioé $x \in argmax\ \varphi_x(y)$
- Se X è continua, fx è massima in x, cioé $x \in argmax\ f_x(y)$

Se la moda è unica X è unimodale. Se ha 2 mode è bimodale. Se ha più di 2 mode è multimodale

Modelli assolutamente continui

27 Uniformi

Dati due numeri reali $a < b$ chiamiamo un variabile aleatoria X uniforme su $[a, b]$ se la sua densità f_X è costante in (a, b) e nulla altrove

$$X \sim unif(a, b) \quad \text{o} \quad X \sim unif[a, b]$$

$$\frac{1}{b-a}$$

Indicatori:

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

la mediana coincide con la media e la moda coincide con qualunque valore in (a, b)

28 Esponenziali

X è esponenziale di parametro $\lambda > 0$ se

$$f_x(x) = \begin{cases} c \cdot e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$c = \lambda$, $F_X(x) = 1 - e^{-\lambda x}$

Indicatori

- $E[X] = \frac{1}{\lambda}$
- $Var[X] = \frac{1}{\lambda^2}$
- la moda è 0
- la mediana è $\frac{log(2)}{\lambda}$

Le esponenziali hanno assenza di memoria, cioé per $s, t > 0$:

$$P(X > s + t \mid X > s) = P(X > t)$$

29 Gaussiane o Normali

X è normale standard se ha densità:

$$X \sim \mathcal{N}(0,1) \Rightarrow f_x(x) = \frac{1}{\sqrt{2\pi}}^{(1)} \cdot e^{\frac{-x^2}{2}} \ ^{(2)}$$

(1) è costante di normalizzazione, mentre (2) da la forma a campana

Proprietà

- f_X è simmetrica rispetto a $x = 0,\ f_X(x) = f_X(-x)$
- f_x ha massimo in $x = 0$. Tale massimo è $\frac{1}{\sqrt{2\pi}}$
- ha punti di flesso in ± 1
- in ± 2 ha valore ≈ 0.05 e in ± 3 vale ≈ 0.004

Funzione di ripartizione

$$\Phi(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}} dt$$

Proprietà

- in 0 vale $\frac{1}{2}$
- è simmetrica rispetto a $(0, \frac{1}{2}) \rightarrow \Phi(x) = 1 - \Phi(x)$
- in -3 vale ≈ 0.0013, in 3 vale ≈ 0.9987
- $\Phi(-2) \approx 0.0228$, $\Phi(2) \approx 0.9772$
- è monotona strettamente crescente

Φ^{-1} è funzione quantile

$$\Phi^{-1}(p) = x \Leftrightarrow \Phi(x) = p \Leftrightarrow P(X \leq x) = p$$

La funzione quantile è simmetrica rispetto a $(\frac{1}{2}, p)$, cioé $\Phi^{-1}(p) = -\Phi^{-1}(1-p)$

Indicatori

- $E[X] = 0$
- $Var[X] = 1$

Def
Sia $Z \sim \mathcal{N}(0,1)$ allora X è una variabile aleatoria normale di parametri $\mu \in \mathbb{R}, \sigma > 0$ se $X = \sigma Z + \mu$, e scriveremo $X \sim \mathcal{N}(\mu, \sigma)$

$$F_X(x) = P(z \leq \frac{x-\mu}{\sigma}) = \Phi(\frac{x-\mu}{\sigma})$$

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

- $E[X] = \mu$
- $Var[X] = \sigma^2$

Proprietà
$X \sim \mathcal{N}(\mu, \sigma)$ eredità le proprietà di $Z \sim \mathcal{N}(0, 1)$ tenendo conto di trasformazioni e deviazione

Prop: la famiglia Gaussiana è riproducibile. Dati $X_1 \sim \mathcal{N}(\mu_1, \sigma_1)$ e $X_2 \sim \mathcal{N}(\mu_2, \sigma_2)$, allora:

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$$

30 Chi quadro

Se X è una somma di n quadrati di guassiane standard indipendenti, diciamo che X è una chi quadro con n gradi di libertà, e scriveremo $X \sim X_n^2$ oppure $X \sim \mathcal{X}^2(n)$

$$(Y_i)_{i=1}^n \quad iid \qquad Y_i \sim \mathcal{N}(0, 1) \quad X = \sum_{i=1}^n Y_i^2$$

Oss: Se sommiamo 2 quadrati, la distribuzione che otteniamo è una esponenziale con $\lambda = \frac{1}{2}$

Oss: le chi quadro sono riproducibili

La densità $f_X(x) = c_n \cdot x^{\frac{n}{2}-1} \cdot e^{\frac{-x}{2}}$ con c_n opportuna rinormalizzazione

Indicatori:

- $E[X] = n$
- $Var[X] = 2n$

31 T (di Student)

Data $Z \sim \mathcal{N}(0, 1)$, $W \sim \mathcal{X}^2(n)$ indipendenti, X è una *t di student* a n gradi di libertà se $X = \frac{Z}{\sqrt{w/n}}$ e scriviamo $X \sim t(n)$ o $X \sim t_n$

Proprietà

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

Indicatori

- $E[X] = 0 \quad \forall n$
- $Var[X] = \begin{cases} \text{non definita} & n = 1 \\ +\infty & n = 2 \\ \frac{n}{n-2} & n > 2 \end{cases}$

Oss: $Var[X] \to 1$ con lim $n \to \infty$

- $E[\frac{W}{n}] = 1$
- $Var[\frac{W}{n}] = \frac{2}{n} \qquad n \to \infty = 0$

17/05/

32 Convergenza variabili aleatorie

32.1 convergenza quasi certa

$(\Omega, \mathcal{F}, P)$ spazio di probabilità e X variabile aleatoria su di esso e $(X_n)_{n\in\mathbb{N}}$ una successione di variabili aleatorie sullo spazio di probabilità. Diciamo che $(X_n)_n$ *converge quasi certamente* o *puntualmente* a X e scriviamo $X_n \xrightarrow[n\to\infty]{\text{qc}}$ se esiste $E \in \mathcal{F}$ con $P(E) = 1$ tale che $\forall w \in E$:

$$\lim_{n\to\infty} \quad X_n(w) = X(w)$$

32.2 Convergenza in probabilità

$(\lambda_n)_n$ successione di variabili aleatorie e X variabile aleatoria su $(\Omega, \mathcal{F}, P)$. Diciamo che $(X_n)_n$ converge *in probabilità* a X e scriviamo $X_n \xrightarrow[n\to\infty]{p} X$ se $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} \quad P(|X_n - X| > \varepsilon) = 0$$

32.3 Convergenza in media quadratica

Diciamo che $(X_n)_n$ converge in *media quadratica* on in $\mathcal{L}^2$ a X e scriviamo $X_n \xrightarrow[n\to\infty]{\mathcal{L}^2} X$ se:

$$\lim_{n\to\infty} \quad E[|X_n - X|^2] = 0$$

Prop: la convergenza in media quadratica implica la convergenza in probabilità, cioé

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}^2} X \qquad \text{allora} \qquad X_n \xrightarrow[n\to\infty]{p} X$$

32.4 Convergenza in distribuzione

Sia $(X_n)_{n\in\mathbb{N}}$ successione su $(\Omega, \mathcal{F}, P)$ e X variabile aleatoria su $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. Diciamo che $(X_n)_n$ converge *in distribuzione* o *in legge* o *debolmente* a X ($X_n \xrightarrow[n\to\infty]{\mathfrak{L}} X,\ X_n \xrightarrow[n\to\infty]{d} X$
o $X_n \xrightarrow[n\to\infty]{w} X$) se $\forall x \in \mathbb{R}$:

$$\lim_{n\to\infty} \quad P(X_n \leq x) = P(X \leq x)$$
ossia
$$\lim_{n\to\infty} \quad F_{X_n}(x) = F_X(x)$$

Prop: Le convergenze quasi certe implicano la convergenza in probabilità

$$X_n \xrightarrow{\alpha} X \Leftarrow X_n \xrightarrow{p} X \Leftarrow \begin{cases} X_n \xrightarrow{qc} X \\ X_n \xrightarrow{\mathcal{L}^2} X \end{cases}$$

Oss: La convergenza in $\mathcal{L}^2$ e qc non sono confrontabili

Prop: la convergenza in probabilità implica la convergenza debolmente

33 Teoremi limite

Sia $(X_1 \_\_ X_n)$ un vettore aleatorio con componenti indipendenti, di media comune μ e varianza comune σ^2. Sia $S_n = \sum_{i=1}^n X$ la variabile aleatoria somma. Allora:

$$E[\tfrac{S_n}{n}] = \mu \quad \text{e} \quad Var[\tfrac{-S_n}{n}] = \tfrac{\sigma^2}{n}$$

34 Teorema della legge debole dei grandi numeri

Sia $(X_n)_{n\in\mathbb{N}}$ variabile aleatoria ciascuna di media μ e varianza σ^2. Sia inoltre $S_n = \sum_{i=1}^n X_i$. Allora $\frac{S_n}{n}$ converge in probabilità a μ, cioé $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} \quad P(|\tfrac{S_n}{n} - \mu| > \varepsilon) = 0$$

35 Teorema centrale del limite

$$\tfrac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{\mathfrak{L}} \mathcal{N}(0,1)$$
cioé
$$\lim_{n\to\infty} \quad P(\tfrac{S - n\mu}{\sigma\sqrt{n}} \leq x) = \Phi(x)$$

Oss: $\frac{S_n - n\mu}{\sigma}$ è nell'ordine di $\sqrt{n}$ per $n \to \infty$

Oss: il modo in cui usiamo il TLC se n è sufficientemente grande, allora:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0,1) \qquad \frac{S_n}{2} \sim \mathcal{N}(\mu, \frac{\sigma}{\sqrt{n}})$$

$$S_n \sim \mathcal{N}(n\mu, \sigma\sqrt{n})$$

n è sufficientemente grande quando:

- $X_i \sim \mathcal{N} \rightarrow\ n \geq 1$
- $X_i \sim Unif \rightarrow\ n \geq 5$
- $X_i simgeom$ o $X_i \sim exp \rightarrow\ n \geq 15$
- $X_i \sim X^2 \rightarrow n \geq 25$

35.1 Correzione di continuità

Se le X_i sono discrete nella versione approssimata usiamo la correzione di continuita, ossia:

$$F_{S_n}(x) \simeq \Phi\left(\frac{x+1/2-n\cdot E[X_i]}{\sqrt{n\cdot Var[X_i]}}\right)$$

Oss:$bin(n,p) \sim \mathcal{N}(np, \sqrt{np(1-p)})$ purché p sia lontano da 0 e 1. Cioé se $np(1-p) \gtrsim 3$:

$$Pois(\lambda) \sim \mathcal{N}(\lambda, \sqrt{\lambda}) \qquad \text{per } \lambda \geq 30$$

36 Popolazione di riferiento

La *popolazione di riferimento* è un insieme di elementi sui quali conduciamo un'indagine. Gli elemtni si chiamano *individui, esemplari* o *unità statistiche.*

Il *Campione* è un sottoinsieme della popolazione

37 Campionamento

Esistono dievrsi tipi di campionamento: *campionamento casuale semplice, campionamento stratificato, campionamento a grappoli*

38 Variabili

Le caratteristiche che misuriamo sono definite *variabili.* I loro valori si chiamano *modalità* o *livelli*:

- qualitative o categoriche
  - nominali (senza un ordine naturale)
  - ordinali (con ordine naturale)
- quantitative o numeriche
  - discrete
  - continue

Le variabili quantitative possono essere quantificate in scale per intervallo o rapporto

39 Statistica

Chiamiamo *statistica* una funzione calcolabile dalle misurazioni del campione.

Si dice *stimatore* di un parametro di una variabile aleatoria discreta che sia una statistica e il cui valore sia "spesso vicino" al parametro d'interesse.
Il valore deterministico assunto dallo stimatore usando la realizzazione del campione si chiama *stima* del parametro

Notazione: Se θ è uno stimatore di ϑ, possiamo scrivere θ_n per evidenziare la numerosità del campione e $\hat{\vartheta} = \theta$.
Chiedere che θ sia vicino a ϑ significa che l'errore di stima sia piccolo

Def
Uno stimatote θ di ϑ è:

- *corretto* o *non distorto* se $E[\theta] = \vartheta$
- *distorto* se $E[\theta] \neq \vartheta$. Chiameremo allora *bias* $E[\theta] - \vartheta$

Se $Lim_{n\to\infty}\ E[\theta_n] = \vartheta$ allora θ è *asintoticamente non distorto*

40 Errore quadratico medio

Si dice *Errore quadratrico medio*, o MSE, di θ la quantità $MSE[\theta] = E[(\theta - \vartheta)^2]$:

$$MSE[\theta] = Var[\theta] + bias^2$$

Def
θ è *consistente* se $\theta_n \xrightarrow{p} \theta$. θ è *consistente in media quadratica* se $\theta_n \xrightarrow{\mathcal{L}^2} \theta$

Prop
Se θ è asintoticamente non distorto e $lim_{n\to\infty}\ Var[\theta_n] = 0$ allora θ è consistente in media quadratica e quindi consistente

Oss
Stimatori corretti ma non consistenti $(X_1, ..., X_n)$. Allora una qualunque X_i è stimatore della media $\mu \Rightarrow E[X_i] = \mu$. Se $\sigma^2 = Var[X_i] \neq 0$ allora $P(|X_i - \mu| > \varepsilon) > 0$ e rimane tale per $n \to \infty \Rightarrow$ non conv in P