

SUNTI MDP

Alessandro Pioggia

28 dicembre 2021

Indice

0.1	Combinatoria	4
0.1.1	Cos'è la cardinalità di un insieme? Cos'è una corrispondenza biunivoca tra insiemi? Cosa significa insieme numerabile?	4
0.1.2	Esempio di insieme numerabile e non numerabile	4
0.1.3	Che cos'è il prodotto cartesiano?	4
0.1.4	Quanti sottoinsiemi ha un insieme con n elementi?	4
0.1.5	Che cos'è una lista?	5
0.1.6	Parla dei prodotti condizionati	5
0.1.7	Che cosa sono le disposizioni?	5
0.1.8	Definisci le combinazioni, illustra inoltre la correlazione che c'è fra esse e le disposizioni, di conseguenza mostra come contare il numero di combinazioni	6
0.1.9	Come possiamo contare con precisione il numero di anagrammi di una parola?	6
0.1.10	Definisci la funzione ricorsiva di Stifel	6
0.1.11	Parlami dei numeri di Fibonacci	7
0.1.12	Raccontami il principio di inclusione-esclusione	7
0.1.13	Come faccio a contare il numero di funzioni iniettive?	8
0.1.14	Come faccio a contare il numero di funzioni suriettive?	8
0.1.15	Che cosa sono gli scombussolamenti?	9
0.1.16	Che cosa sono le partizioni?	9
0.1.17	Come si contano le partizioni?	9
0.1.18	Che cosa sono e come si ottengono i numeri di Stirling?	10
0.1.19	Che cosa sono e come si ottengono i cammini reticolari?	10
0.2	Statistica descrittiva	10
0.3	Probabilità	11
0.3.1	Introduci la probabilità	11
0.3.2	Cos'è uno spazio di probabilità?	11
0.3.3	A cosa mi riferisco quando parlo di casi possibili su casi totali?	12

0.3.4	Parlami della probabilità condizionale	12
0.3.5	Qual'è la formula delle probabilità totali? Come funzione? A cosa serve?	12
0.3.6	Cos'è una variabile aleatoria?	12
0.3.7	Parlami della densità	12
0.3.8	Parlami della densità uniforme	13
0.3.9	Parla della densità di Bernoulli e della funzione caratteristica	13
0.3.10	Schema successo-insuccesso	14
0.3.11	Raccontami la densità binomiale	14
0.3.12	Che cos'è la variabile ipergeometrica?	14
0.3.13	Raccontami la variabile geometrica e la geometrica modificata	14
0.3.14	Definisci la proprietà di mancanza di memoria, chi ne gode?	15
0.3.15	Parlami della densità di Poisson	15
0.3.16	Parla delle variabili multidimensionali	15
0.3.17	Invece per quanto riguarda la densità delle variabili multidimensionali?	16
0.3.18	Come si trova la densità marginale a partire dalla congiunta?	16
0.3.19	Definisci la funzione di ripartizione	16
0.3.20	Quale è e come si ottiene la funzione di ripartizione del massimo? del minimo invece?	16
0.3.21	Se abbiamo due densità marginali uguali, possiamo dedurre la congiunta? Quando non possiamo? Mi fai un esempio?	17
0.3.22	Considerando una variabile aleatoria X , se gli viene applicata una funzione che la trasforma, come varia la sua densità?	17
0.3.23	Parlami del valore atteso	17
0.3.24	Quali sono le proprietà e le principali caratteristiche del valore atteso?	18
0.3.25	Definisci la varianza	19
0.3.26	Definisci la covarianza	20
0.3.27	Definisci la proprietà più importante della varianza	20
0.3.28	Spiega la legge dei grandi numeri	20
0.3.29	Quale è il valore atteso e la varianza di una variabile di Bernoulli?	22
0.3.30	Definisci la variabile continua	23
0.3.31	Quale è la densità continua?	23

0.3.32	Che legame c'è fra una funzione di ripartizione continua e la densità	23
0.3.33	Come si trova il valore atteso nell'ambito delle variabili continue?	23
0.3.34	Come si determina la varianza nell'ambito delle variabili continue?	24
0.3.35	Definisci la densità continua uniforme	24
0.3.36	Definisci la densità continua astratta e la funzione di ripartizione astratta	24
0.3.37	Parlami della densità esponenziale	25
0.3.38	Raccontami la densità normale	25

0.1 Combinatoria

0.1.1 Cos'è la cardinalità di un insieme? Cos'è una corrispondenza biunivoca tra insiemi? Cosa significa insieme numerabile?

Supponendo di avere un insieme A con n elementi, n è la cardinalità. Supponendo di avere due insiemi A e B , definiamo una funzione $f(A) : B$, se è sia suriettiva che iniettiva i due insiemi sono in corrispondenza biunivoca (ovvero è possibile associare a ciascun elemento di un insieme uno ed uno solo elemento dell'altro). Un insieme viene detto numerabile se i suoi elementi sono in numero finito o se possono essere messi in corrispondenza biunivoca con i numeri naturali. Se riesco ad elencare un insieme tramite un elenco infinito è numerabile e viceversa.

0.1.2 Esempio di insieme numerabile e non numerabile

Esempio di insiemi numerabili:

- $P = \{0, 2, 4, 6, \dots\} \leftrightarrow \mathbb{N} = \{0, 1, 2, 3, \dots\}$ con $f(x) = 2x$
- $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\} \leftrightarrow \mathbb{N} = \{0, 1, 2, 3, \dots\}$ con $f(z) = 2z - 1$ se $z \geq 0$,
 $f(z) = -2z$ se $z \leq 0$

Esempio di insiemi non numerabili:

- L'insieme \mathbb{R}
- $A = \{ \text{Sequenze binarie di lunghezza infinita} \}$

0.1.3 Che cos'è il prodotto cartesiano?

Dati due insiemi A e B , il prodotto cartesiano rappresentato da $A \times B$ non è altro che l'insieme delle coppie ordinate (a, b) con $a \in A$ e $b \in B$ (Dunque $(a, b) \neq (b, a)$). Formalmente : $A \times B = \{(a, b) | a \in A, b \in B\}$

0.1.4 Quanti sottoinsiemi ha un insieme con n elementi?

Ne ha 2^n , il risultato lo si ottiene conoscendo l'insieme delle parti (si indica con $p(A)$) e le sue proprietà. L'insieme delle parti è appunto il numero di sottoinsiemi di un dato insieme, per giungere alla risposta occorre sfruttare una logica (supponendo di considerare un insieme A con cardinalità n). Vogliamo dimostrare che : se $|A| = n$, allora $|p(A)| = 2^n$, dunque ragioniamo in questi termini :

- fra $p(A)$ e $\{0, 1\}^n$ c'è una corrispondenza biunivoca.
- dal principio di uguaglianza abbiamo che, dal momento che esiste una corr. biunivoca, $|p(A)| = |\{0, 1\}^n|$
- dal principio della moltiplicazione otteniamo che $|\{0, 1\}^n| = |\{0, 1\}|^n = 2^n$

0.1.5 Che cos'è una lista?

Gli elementi di una potenza cartesiana si dicono liste o sequenze in A e hanno cardinalità n . Esempio :

A^3 ha una lunghezza delle sequenze o liste pari a 3

0.1.6 Parla dei prodotti condizionati

$S \subseteq A \times B$, S è un prodotto condizionato \leftrightarrow posso scegliere la prima coordinata di un elemento di S in n modi e la seconda coordinata, una volta fissata la prima, in m .

Se $S \subseteq A \times B$ è un prodotto condizionato di tipo (n, m) allora $|S| = n \cdot m$.

Dimostrazione :

- dividiamo S in n sottoinsiemi disgiunti tali che : $\{S_1 \cup S_2 \cup \dots \cup S_n\}$
- $S_i = \{ \text{tutti gli elementi di } S \text{ con prima coordinata } a_i \in A \}$
- per definizione abbiamo dunque che $|S_i| = m$
- dal momento che i sottoinsiemi considerati sono disgiunti (dal momento che rispettano le condizioni del prodotto condizionato), otteniamo che $S = \{S_1 + S_2 + \dots + S_n\}$
- $|S| = \{|S_1| + |S_2| + \dots + |S_n|\} = \{m + m + \dots + m\} = m \cdot n$

0.1.7 Che cosa sono le disposizioni?

Una disposizione di lunghezza k su un insieme di cardinalità n , è una sequenza in cui all'interno non ci sono valori ripetuti. Formalmente : $\{a_1, a_2, \dots, a_k\}$ con $a_i \neq a_j$ e $j = i + 1$ è una disposizione di lunghezza k .

Una disposizione lunga k su un insieme di n elementi è esprimibile come prodotto condizionato di tipo : $(n, n - 1, n - 2, \dots, n - k)$, questo perchè una volta che seleziono la prima coordinata, se la seconda deve essere distinta ho $n-1$ scelte e così via...

Dalla proprietà dei prodotti condizionati dimostrata in precedenza abbiamo che $|\text{disposizioni}| = (n \cdot n - 1 \cdot n - 2 \cdot \dots \cdot n - k)$. Se ho $k = n \rightarrow |\text{disposizioni}| = n!$, vengono chiamate permutazioni.

0.1.8 Definisci le combinazioni, illustra inoltre la correlazione che c'è fra esse e le disposizioni, di conseguenza mostra come contare il numero di combinazioni

Considerando un insieme A con $\{1, \dots, n\}$ elementi, le combinazioni di A sono tutti i suoi sottoinsiemi di cardinalità k . Per contare il numero di combinazioni lunghe k di un insieme con n elementi è possibile sfruttare la correlazione che c'è fra disposizioni e combinazioni. Per ogni combinazione lunga k esistono esattamente $k!$ disposizioni, infatti se mettiamo a funzione disposizioni e combinazioni otteniamo una funzione suriettiva (disposizioni \rightarrow combinazioni). A questo proposito deduciamo che : $|\text{combinazioni}| = \frac{n(k)}{k!}$.

0.1.9 Come possiamo contare con precisione il numero di anagrammi di una parola?

Possiamo sfruttare le combinazioni di tipo (a, b, c) , con $|a| + |b| + |c| = n$. Gli anagrammi di tipo (a, b, c) sono le sequenze ordinate in cui appare lo 0 a volte, l'1 b volte, il 2 c volte. Inoltre è possibile mettere in biezione le combinazioni di tipo (a, b, c) con le sequenze ternarie ordinate di sottoinsiemi (s_1, s_2, s_3) , in cui $|s_1| = a$, $|s_2| = b$, $|s_3| = c$. (In s_1 inserisco gli indice posizionali degli 0 che appaiono, di conseguenza lo faccio anche con s_2 e s_3). Inoltre abbiamo che $s_1 \cup s_2 \cup s_3 = \{0, \dots, n\}$

0.1.10 Definisci la funzione ricorsiva di Stifel

La funzione ricorsiva di Stifel è la seguente : $\sum_{k=0}^{n/2} \left(\binom{n-1}{k} + \binom{n-1}{k-1} \right)$

Per dimostrarne la correttezza è sufficiente suddividere l'insieme preso in considerazione in due sottoinsiemi, che chiameremo A e B . $A =$ insiemi che contengono n $B =$ insiemi che non contengono n

Gli insiemi che non contengono n sono $\binom{n-1}{k}$, ovvero escludo dalla conta del numero dei sottoinsiemi lo stesso n . Per quanto riguarda A invece, possiamo ottenere il valore $\binom{n-1}{k-1}$ perchè considero tutti i sottoinsiemi di cardinalità $k - 1$ con n escluso, dopodichè n verrà aggiunto a ciascun sottoinsieme.

0.1.11 Parlami dei numeri di Fibonacci

I numeri di Fibonacci sono l'insieme delle sequenze di bit che non presentano due 1 consecutivi.

Formula : $f_n = f_{n-1} + f_{n-2}$

Per dimostrare la correttezza della formula è sufficiente suddividere le sequenze fra quelle che:

- iniziano con 1 : prendo n-2 elementi ed aggiungo all'inizio qualsiasi coppia che non sia (1, 1).
- iniziano con 0 : prendo n-1 elementi ed aggiungo 0 all'inizio.

C'è anche una formula alternativa per calcolare i numeri di fibonacci :
$$\sum_{k=0}^{n/2} \binom{n-k+1}{k}$$

Dimostrazione :

In pratica in questo caso prendiamo una sequenza e togliamo tutti gli zeri, quindi abbiamo tutti piccoli insiemi che contengono sequenze con soli 1. Funziona perchè so che per rispettare le condizioni, ogni 1 della sequenza deve essere affiancato da uno 0, tranne eventualmente se l'1 ricopre l'ultima posizione, dunque ho n - k zeri - 1 (che sta in fondo).

0.1.12 Raccontami il principio di inclusione-esclusione

Il principio di inclusione-esclusione permette di calcolare la cardinalità di un insieme, espresso come unione di n insiemi finiti.

In formule (Si ricorda che I è l'insieme di sottoinsiemi di un insieme con 1, ..., n elementi) :

$$|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_{I \in 0, \dots, n} (-1)^{|I|-1} \cdot \prod_{i \in I} |A_i|$$

Detto ciò sappiamo che molto spesso viene sfruttata la formula che ci permette di arrivare al complementare. Per fare ciò è opportuno conoscere l'insieme universo U, che contiene tutti gli elementi e tutti gli insiemi esistenti (quindi anche sè stesso e l'insieme vuoto). Prima di enunciare la formula è importante sapere che : $|U| = |\prod_{i \in 0} A_i|$.

$$|(A_1 \cup A_2 \cup \dots \cup A_n)^c| = |U| - \sum_{I \in 0, \dots, n} (-1)^{|I|-1} \cdot \prod_{i \in I} |A_i|$$

$$= |\prod_{i \in 0} A_i| - \sum_{0 \neq I \in 0, \dots, n} (-1)^{|I|-1} \cdot \prod_{i \in I} |A_i|$$

$$= |\prod_{i \in 0} A_i| + \sum_{0 \neq I \in 0, \dots, n} (-1)^{|I|} \cdot \prod_{i \in I} |A_i|$$

$$= \sum_{I \in 0, \dots, n} (-1)^{|I|} \cdot \prod_{i \in I} |A_i|$$

0.1.13 Come faccio a contare il numero di funzioni iniettive?

RECAP : Una funzione iniettiva è definita tale se elementi distinti del dominio hanno immagini distinte.

Supponendo di avere una funzione iniettiva $f : A \rightarrow B$ con $m \leq n$ (poniamo $|A| = m$ e $|B| = n$), il numero di funzioni iniettive è n_m infatti, piccolo appunto, non è altro che un prodotto condizionato di tipo $(n - 1, n - 2, \dots, n - m)$.

0.1.14 Come faccio a contare il numero di funzioni suriettive?

RECAP : Una funzione suriettiva è definita tale se ogni elemento del codominio ha come immagine almeno un elemento del dominio. Supponendo di avere una funzione suriettiva $f : A \rightarrow B$ con $m \geq n$, (poniamo ancora $|A| = m$ e $|B| = n$), il numero di funzioni suriettive è possibile calcolarlo sfruttando il principio di inclusione-esclusione. Supponiamo di avere un insieme A composto da n sottoinsiemi, ovvero $A = \{A_1 \cup A_2 \cup \dots \cup A_n\}$

in cui definiamo $A_i = \{ \text{tutte le funzioni in cui l'elemento } i \notin \text{Im} f \}$, la sua cardinalità è : $|A_i| = (n - 1)^m$, stessa regola la usiamo per l'intersezione, ovvero $|A_i \cap A_j| = (n - 2)^m$ e dunque generalizzando avremo che :

$|\prod_{i \in 0 \dots n} A_i| = (n - |I|)^m$ In seguito a questa deduzione osserviamo che se, la nostra funzione appartiene anche solo ad uno degli $|A_i|$ insiemi, non può essere definita suriettiva, dal momento che ogni elemento del codominio deve avere come immagine almeno un elemento del dominio. Quindi per contare il numero di funzioni suriettive occorre calcolare il complementare dell'insieme che abbiamo preso in considerazione prima, ovvero $|(A_1 \cup A_2 \cup \dots \cup A_n)^c|$, sfruttando il principio di inclusione-esclusione.

$$|(A_1 \cup A_2 \cup \dots \cup A_n)^c| = |\prod_{i \in 0 \dots n} A_i| - \sum_{0 \neq i \in I} (-1)^{(|I|-1)} |\prod_{i \in 0 \dots n} A_i|$$

$$= \sum_{i \in I} (-1)^{(|I|)} \cdot (n - |I|)^m$$

$$= \sum_{k=0}^n (-1)^k \cdot \binom{n}{k} \cdot (n - k)^m$$

0.1.15 Che cosa sono gli scombussolamenti?

Gli scombussolamenti rappresentano tutte le possibili permutazioni in cui nessun elemento rimane al proprio posto. Per contarli è sufficiente sfruttare la logica che è stata usata per contare il numero di funzioni suriettive, ovvero :

Supponendo di avere un insieme $A = (A_1 \cup A_2 \cup \dots \cup A_n)$,

in cui $A_i = (n-1)! = \{\text{tutte le permutazioni in cui } i \text{ rimane al proprio posto}\}$; ne vogliamo calcolare il complementare (in modo da trovare solo quelli che non cambiano mai di posto!).

Dunque torniamo al calcolo di :

$$\begin{aligned} |(A_1 \cup A_2 \cup \dots \cup A_n)^c| &= \sum_{I \in 0, \dots, n} (-1)^{|I|} \cdot |\prod_{i \in I} A_i| \\ &= \sum_{k=0}^n (-1)^k \cdot \binom{n}{k} \cdot (n-k)! \\ &= \sum_{k=0}^n (-1)^k \cdot \frac{n!}{k!(n-k)!} \cdot (n-k)! \\ &= \sum_{k=0}^n (-1)^k \cdot n_{n-k} \end{aligned}$$

0.1.16 Che cosa sono le partizioni?

Le partizioni rappresentano la suddivisione di un insieme in blocchi disgiunti, rigorosamente non vuoti.

0.1.17 Come si contano le partizioni?

Per contare le partizioni si utilizzano i numeri di Bell, per fare ciò abbiamo a nostra disposizione una formula ricorsiva : $\sum_{k=1}^n \binom{n-1}{k-1} \cdot B_{n-k}$

Dimostrazione della formula :

- Considerando un insieme di cardinalità k in cui è contenuto anche n , noi vogliamo contare tutti i possibili sottoinsiemi formati dai restanti $k-1$ elementi, sono $\binom{n-1}{k-1}$.
- Per ogni sottoinsieme ottenuto al punto precedente, noi vogliamo calcolarne tutte le possibili partizioni, che otterremo con il passo ricorsivo : B_{n-k}
- In conclusione otteniamo un prodotto condizionato del tipo : $(\binom{n-1}{k-1}, B_{n-k})$

0.1.18 Che cosa sono e come si ottengono i numeri di Stirling?

I numeri di stirling servono per contare il numero di partizioni di una precisa cardinalità k su un insieme di n elementi. Per contare è sufficiente usare la formula ricorsiva così definita :

$$S_{n,k} = k \cdot S_{n-1,k} + S_{n-1,k-1}$$

Dimostrazione della formula :

- n in compagnia : prendo in considerazione tutte le partizioni che contengono n , quindi avrò $S_{n-1,k}$, a cui aggiungerò ad ogni blocco n , quindi moltiplico il tutto per k
- n da solo : Devo partizionare tutti i numeri da 1 a $k-1$ (escluso n che è da solo) in $k - 1$ blocchi $\rightarrow S_{n-1,k-1}$

0.1.19 Che cosa sono e come si ottengono i cammini reticolari?

0.2 Statistica descrittiva

0.3 Probabilità

0.3.1 Introduci la probabilità

La probabilità studia i fenomeni aleatori che, a differenza di quelli deterministici, non permettono di prevedere con esattezza il risultato, è necessario dunque effettuare delle stime.

Ω = "insieme di tutti i possibili risultati di un fenomeno aleatorio"

Evento : sottoinsieme di Ω , dunque è l'insieme dei possibili risultati di un fenomeno aleatorio.

Famiglia coerente di eventi (E_1, E_2, \dots, E_n) : collezione di sottoinsiemi di Ω che rispettano le seguenti condizioni (lo indichiamo con A)

- $0, \Omega \in A$
- Se $E \in A$ anche $E^c \in A$
- Se E_1, E_2, \dots, E_n sono eventi, allora sia la loro intersezione che unione è un evento

Valutazione di probabilità : Funzione che associa ad ogni evento un numero, scelto in base alle probabilità che esso accada.

Volendola definire formalmente :

Data una famiglia coerente di eventi A , una valutazione di probabilità su A è una funzione $P : A \rightarrow \mathbb{R}$ che soddisfa le seguenti proprietà :

- $P(\Omega) = 1$
- $P(A) \geq 0$
- $P(A_1 \cup A_2, \dots, \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$

L'ultima proprietà ci dice che gli eventi devono essere disgiunti.

0.3.2 Cos'è uno spazio di probabilità?

Uno spazio di probabilità è una terna di tipo (Ω, A, P) in cui Ω è un insieme, A è una famiglia coerente di eventi e P è una probabilità su A . Parliamo invece di spazio di probabilità uniforme quando, Ω è un insieme finito, A è un insieme di possibili sottoinsiemi di Ω e tutti i risultati sono equiprobabili. In formule :

$\Omega = \{a_1, a_2, a_3, \dots, a_n\}$; $A = \{\text{Possibili sottoinsiemi di } \Omega\}$; inoltre richiediamo che $P(a_1) = P(a_2) = P(a_3) = \dots = P(a_n) = p$

0.3.3 A cosa mi riferisco quando parlo di casi possibili su casi totali?

Considerando uno spazio di probabilità uniforme (Ω, A, P) ed $A \in \Omega$ è un evento, allora : $P(A) = \frac{|A|}{|\Omega|}$ e si tratta di casi possibili su casi totali.

Dimostrazione :

Prendiamo un qualunque sottoinsieme di Ω con cardinalità k ,

$$P(A) = \frac{1}{n} + \dots + \frac{1}{n} = \frac{k}{n} = \frac{|A|}{|\Omega|}$$

0.3.4 Parliamo della probabilità condizionale

Considerando due eventi distinti A e B , il calcolo della probabilità condizionale consta nel definire la probabilità che accada l'evento B sapendo che è già avvenuto A (o viceversa). Questa probabilità la definiamo come $P(B|A)$ e si calcola nella seguente maniera: $P(B|A) = \frac{P(A \cap B)}{P(A)}$, questo perchè non è altro che la proporzione di $A \cap B$ in A .

0.3.5 Qual'è la formula delle probabilità totali? Come funziona? A cosa serve?

La formula delle probabilità totali ci consente di calcolare la probabilità di un evento, conoscendo le probabilità condizionali, suddividendolo in casi. Formalmente, supponendo di avere 3 eventi A , B e C , con C che è intersecato sia ad A che a B : $P(C) = P(C \cap A) + P(C \cap B) = P(C) \cdot P(A|C) + P(C) \cdot P(B|C)$

0.3.6 Cos'è una variabile aleatoria?

E' una quantità che dipende dal risultato del fenomeno aleatorio. Data una funzione $X : \Omega \rightarrow \mathbb{R}$, X è una variabile aleatoria se:

$$\{\omega \in \Omega | X(\omega) \leq a\}$$

è un evento per ogni $a \in \mathbb{R}$.

0.3.7 Parliamo della densità

Partiamo differenziando due tipi di densità, quella discreta e quella astratta. Densità discreta:

Data una variabile aleatoria discreta X , prende il nome di densità discreta

la funzione $p_x : \mathbb{R} \rightarrow \mathbb{R}$ tale che $p_x(h) = P(x = h)$ con h numero intero. Per essere definita tale deve soddisfare 3 proprietà:

- $p_x(h) \geq 0$ per ogni $h \in \mathbb{R}$
- $p_x(h) \neq 0$ per una quantità al più numerabile di valori h
- $\sum_{h \in \mathbb{R}} p_x(h) = 1$

Densità astratta:

Prende il nome di densità astratta la funzione $p : \mathbb{R} \rightarrow \mathbb{R}$ che soddisfa le 3 proprietà:

- $p(h) \geq 0$
- $p(h) \neq 0$
- $\sum_{h \in \mathbb{R}} p(h) = 1$

0.3.8 Parlami della densità uniforme

Supponendo di avere un evento $A = \{x_1, x_2, \dots, x_n\}$, abbiamo che X è una variabile uniforme, se tutte le x_i hanno la seguente densità, detta uniforme

$$\rightarrow \begin{cases} \frac{1}{n} & k = x_1, x_2, x_3, \dots, x_n \\ 0 & \text{altrimenti} \end{cases}$$

0.3.9 Parla della densità di Bernoulli e della funzione caratteristica

Una variabile di bernoulli può assumere solo due valori: $\{1, 0\}$ e la sua densità

$$\text{è la seguente: } d_x(k) = \begin{cases} P(A) & k = 1 \\ 1 - P(A) & k = 0 \\ 0 & \text{altrimenti} \end{cases}$$

Se A è un evento la sua funzione caratteristica X_A data da :

$$X_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

è una variabile aleatoria, perchè soddisfa la definizione, ovvero che $X_A \leq a$ è un evento per ogni $a \in \mathbb{R}$. Deduciamo quindi che la seguente relazione è rappresentabile attraverso una variabile di Bernoulli, che assume 1 con probabilità p e 0 con probabilità $(p - 1)$.

0.3.10 Schema successo-insuccesso

Uno schema successo-insuccesso è una sequenza di un numero infinito di fenomeni aleatori, ognuno dei quali può dare successo o insuccesso. Si può studiare il numero di successi (variabile binomiale, ogni successo avviene con probabilità p ed ogni fenomeno aleatorio è indipendente dall'altro) o il numero di insuccessi prima di ottenere un successo (variabile geometrica).

0.3.11 Raccontami la densità binomiale

Le variabili binomiali, aventi densità binomiale, rappresentano uno schema successo-insuccesso a prove indipendenti (significa che la probabilità di avere successo è la stessa per ogni tentativo, indipendentemente da ciò che è successo prima). La variabile binomiale, nel caso generale, determina la probabilità che io ottenga k successi su n tentativi. Supponendo di avere una variabile aleatoria

X = "numero di successi su n tentativi", la probabilità che

$$P(X = k) = \begin{cases} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} & k = 1, \dots, n \\ 0 & \text{altrimenti} \end{cases}$$

0.3.12 Che cos'è la variabile ipergeometrica?

La variabile ipergeometrica, rappresenta un particolare schema successo-insuccesso a prove dipendenti (Ovvero la probabilità che ottenga successo in un tentativo è influenzato dai risultati precedenti). Per definire la densità di questa variabile è sufficiente prendere in considerazione una pescata di n palline da una urna, con palline rosse e nere, senza rimpiazzo, con successo="estrazione di una rossa", dunque avrò una variabile X = "numero di successi su n tentativi". La densità risulterà così:

$$P(X = k) = \begin{cases} \frac{\binom{r}{k} \binom{nere}{n-k}}{\binom{r+nere}{n}} & k = 1, \dots, n \\ 0 & \text{altrimenti} \end{cases}$$

0.3.13 Raccontami la variabile geometrica e la geometrica modificata

La variabile geometrica modificata ci permette di contare il numero di tentativi necessari prima di ottenere un successo. La sua densità è: $(1-p)^{k-1} \cdot p$

La variabile geometrica invece, ha come obiettivo quello di contare il numero di insuccessi che si ottengono prima di raggiungere il successo. Si nota che non è altro che la geometrica modificata shiftata di uno, dunque la sua densità sarà: $(1 - p)^k \cdot p$

0.3.14 Definisci la proprietà di mancanza di memoria, chi ne gode?

La proprietà di mancanza di memoria afferma che la probabilità che un evento accada non è influenzata da tutto ciò che è avvenuto prima. Formalmente:

$$P(X \geq k + m | X \geq m) = P(X \geq k)$$

Dimostrazione (con variabile geometrica):

$$\begin{aligned} P(X \geq k + m | X \geq m) &= \frac{P(X \geq k + m \cap X \geq m)}{P(X \geq m)} = \frac{P(X \geq k + m)}{P(X \geq m)} \\ &= \frac{(1 - p)^{k+m} \cdot p}{(1 - p)^m \cdot p} = (1 - p)^k = P(X \geq k) \end{aligned}$$

0.3.15 Parliamo della densità di Poisson

Quando abbiamo una variabile binomiale con n molto grande ($n \geq 50$) e una probabilità molto piccola ($p \leq 0.02$), abbiamo la necessità di sfruttare la densità citata, in quanto con la binomiale verrebbe un calcolo troppo complicato. La densità di Poisson è la seguente:

$$\begin{cases} e^{-\lambda} \cdot \frac{\lambda^k}{k!} & k = 1, 2, \dots \\ 0 & \text{altrimenti} \end{cases}$$

Occorre dimostrare che essa sia una densità, ovvero occorre che $\sum_{k=0}^{\infty} d(k) = 1$

$$\sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

Ciò è stato possibile sfruttando il fatto che: $e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$ grazie ad una proprietà dello sviluppo di Taylor.

0.3.16 Parla delle variabili multidimensionali

Una variabile multidimensionale è una cosa del tipo $X = (x_1, x_2, \dots, x_m)$ dove x_1, \dots, x_m sono m variabili aleatorie. L'unica condizione è che queste variabili siano definite sullo stesso spazio (Ω) .

0.3.17 Invece per quanto riguarda la densità delle variabili multidimensionali?

La densità multidimensionale è una funzione $d_x : \mathbb{R}^m \rightarrow \mathbb{R}$. Ad esempio con 3 variabili risulta: $d_x(k, h, l) = P(x_1 = k, x_2 = h, x_3 = l)$, vogliamo che simultaneamente queste tre variabili assumano questi tre valori (k, h, l).

0.3.18 Come si trova la densità marginale a partire dalla congiunta?

Prendendo come punto di riferimento il caso più comune, ovvero il caso bidimensionale, occorre sommare le varie densità congiunte in cui la coordinata corrispondente alla marginale che vogliamo trovare rimane fissa mentre le altre variano. Formalmente:

$$d_x(k) = \sum_{h \in R} d_{x,y}(k, h_1, h_2, \dots, h_n)$$
$$d_y(h) = \sum_{k \in R} d_{x,y}(k_1, k_2, \dots, k_n, h)$$

0.3.19 Definisci la funzione di ripartizione

La funzione di ripartizione di una variabile aleatoria x è una funzione $F_x : \mathbb{R} \rightarrow \mathbb{R}$ in cui $F_x(t) = P(x \leq t)$.

0.3.20 Quale è e come si ottiene la funzione di ripartizione del massimo? del minimo invece?

Considerando due variabili aleatorie indipendenti x e y e $z = \max(x, y)$ abbiamo che la funzione di ripartizione di z è la seguente:

$$F_z(t) = F_x(t) \cdot F_y(t)$$

Dimostrazione:

$$F_z(t) = P(z \leq t) = P(\max(x, y) \leq t) = P(x \leq t, y \leq t)$$
$$= P(x \leq t) \cdot P(y \leq t) = F_x(t) \cdot F_y(t)$$

Invece la funzione di ripartizione del minimo, con $w = \min(x, y)$:

$$(1 - F_w(t)) = (1 - F_x(t)) \cdot (1 - F_y(t))$$

Dimostrazione:

$$(1 - F_w(t)) = P(w \geq t) = P(\min(x, y) \geq t)$$
$$= P(x \geq t, y \geq t) = P(x \geq t) \cdot P(y \geq t)$$
$$(1 - F_w(t)) = (1 - F_x(t)) \cdot (1 - F_y(t))$$

0.3.21 Se abbiamo due densità marginali uguali, possiamo dedurre la congiunta? Quando non possiamo? Mi fai un esempio?

Posso ricavare una densità congiunta a partire dalle marginali solo in caso di indipendenza perchè vale la relazione:

$$P(x_1 = k, \dots, x_n = k) = P(x_1 = k) \cdot \dots \cdot P(x_n = k)$$

Un esempio può essere il lancio di tre dadi, dove consideriamo x_i il risultato ottenuto. Ponendo $Y = x_1 + x_2 + x_3$, abbiamo che moltiplicando le densità marginali otteniamo la densità di Y , quindi la congiunta, questo perchè i lanci di dadi sono indipendenti.

0.3.22 Considerando una variabile aleatoria X , se gli viene applicata una funzione che la trasforma, come varia la sua densità?

Caso monodimensionale

Consideriamo una funzione $\phi : \mathbb{R} \rightarrow \mathbb{R}$, t.c. $\phi(x) = x^2 = z$, abbiamo che:

$$d_z(h) = \sum_{h:\phi(h)=z} d_x(h)$$

Caso bidimensionale

Considerando una funzione $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$, e una variabile aleatoria X di densità $d_x(h)$, supponendo che $\phi(h, k) = h + k = z$, abbiamo che:

$$d_z(h, k) = \sum_{h,k:\phi(h,k)=z} d_{x,y}(h, k)$$

0.3.23 Parlami del valore atteso

Supponendo di avere una variabile aleatoria X , di densità $d_x(h)$, il valore atteso di X è:

$$E[X] = \sum_h h \cdot d_x(h).$$

Ciò non è altro che una media ponderata.

0.3.24 Quali sono le proprietà e le principali caratteristiche del valore atteso?

Teorema

Sia $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ una qualsiasi trasformazione di variabile e sia $X = (x, y)$
Allora (ponendo $z = \phi(X)$) :

$$E[Z] = \sum_{h,k \in \mathbb{R}} \phi(h, k) \cdot d_{x,y}(h, k)$$

Dimostrazione

Per dimostrarlo prendiamo come riferimento una funzione $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$, t.c
 $\phi(x, y) = x + y = z$

$$E[z] = \sum_l l \cdot d_z(l)$$

$$= \sum_l l \cdot \sum_{h,k=\phi(h,k)=l} d_{x,y}(h, k)$$

$$= \sum_{h,k} \phi(h, k) \cdot d_{x,y}(h, k)$$

Corollario

Siano X, Y due variabili aleatorie, non necessariamente indipendenti, allora:

$$E[X + Y] = E[X] + E[Y]$$

Dimostrazione

Consideriamo una funzione $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$, la quale ci permette di ottenere: $\phi(x, y) = x + y = Z$, quindi, dal teorema visto precedentemente otteniamo:

$$\begin{aligned} E[Z] &= \sum_{h,k} \phi(h, k) \cdot d_{x,y}(h, k) \\ &= \sum_{h,k} h \cdot d_{x,y}(h, k) + \sum_{h,k} k \cdot d_{x,y}(h, k) \\ &= \sum_h h \cdot \sum_k d_{x,y}(h, k) + \sum_k k \cdot \sum_h d_{x,y}(h, k) \\ &= \sum_h h \cdot d_x(h) + \sum_k k \cdot d_y(k) \\ &= E[X] + E[Y] \end{aligned}$$

Prodotto di variabili

Siano X, Y due variabili aleatorie indipendenti, allora:

$$E[X \cdot Y] = E[X] \cdot E[Y] \rightarrow \text{non vale il viceversa!!}$$

Dimostrazione

Torniamo a sfruttare la funzione $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$, per il quale $\phi(x, y) = x \cdot y = Z$

$$\begin{aligned} E[Z] &= \sum_{h,k \in \mathbb{R}} \phi(h, k) \cdot d_{x,y}(h, k) \\ &= \sum_{h,k} h \cdot k \cdot d_x(h) \cdot d_y(k) \\ &= E[X] \cdot E[Y] \end{aligned}$$

0.3.25 Definisci la varianza

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

Si osserva che:

$$Var(x) = E[(x - E[x])^2] = E[x^2 - 2 \cdot x \cdot E[x] + E[x]^2] = E[x^2] - E[x]^2$$

0.3.26 Definisci la covarianza

La covarianza permette di osservare la dipendenza presente fra due variabili aleatorie o statistiche, a questo proposito se le variabili sono indipendenti la covarianza sarà $= 0$.

Formalmente:

$$Cov(x, y) = E[x \cdot y] - E[x] \cdot E[y]$$

0.3.27 Definisci la proprietà più importante della varianza

$$Var(x + y) = var(x) + var(y) - 2 \cdot Cov(x, y)$$

Dimostrazione:

$$\begin{aligned} Var(x + y) &= E[(x + y)^2] - E[x + y]^2 \\ &= E[x^2] - 2 \cdot E[x] \cdot E[y] + E[y^2] - (E[x] \cdot E[y])^2 \\ &= E[x^2] - 2 \cdot E[x] \cdot E[y] + E[y^2] - (E[x]^2 - 2 \cdot E[x] \cdot E[y] + E[y]^2) \\ &= Var(x) + Var(y) - 2 \cdot Cov(x, y) \end{aligned}$$

0.3.28 Spiega la legge dei grandi numeri

Per arrivare ad enunciarla abbiamo bisogno di una famosissima formula di probabilità, che si chiama disuguaglianza di Chebyshev. Ci dà una formula quantitativa che esprime il fatto che la varianza è una misura di dispersione. Nel dettaglio:

Disuguaglianza di Chebyshev

Sia X una variabile aleatoria, allora:

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

(fissiamo $\epsilon = 0.1$, qui stiamo calcolando la probabilità che la differenza fra X e la sua media sia maggiore di 0.1. Cosa ci aspettiamo? Se la varianza di questa X è piccola, ci aspettiamo che la probabilità in questione sia piccola. Se X non si disperde assumerà, con buona probabilità, valori vicini al suo valore medio e di conseguenza il valore della probabilità sarà piccolo. La

proposizione ci dice che la probabilità in questione è sempre \leq della varianza, anzi è direttamente proporzionale con la varianza. Per $\epsilon = 1$, la probabilità è minore uguale della varianza. E' una formula quantitativa del fatto che la varianza è una dispersione.)

Dimostrazione

Consideriamo l'evento $A = "|X - E[X]| \geq \epsilon$ ". Consideriamo la variabile caratteristica, la χ_A (La χ_A è una variabile di Bernoulli che vale 1 se A accade e vale 0 se A non accade. La probabilità che valga 1 è $= P(A)$).

Poniamo $Y = \epsilon^2 \chi_A$ e $Z = (X - E[X])^2$

Osserviamo che sussiste sempre una relazione fra Y e Z , cioè che $Y \leq Z$. Infatti:

Caso 1: se A non accade $\rightarrow Y = 0$ e $Z \geq 0$

Caso 2: se A accade $\rightarrow Y = \epsilon^2$ (se A accade, la relazione $|X - E[X]| \geq \epsilon$ è verificata, dunque di conseguenza $|X - E[X]|^2 \geq \epsilon^2$, ma il quadrato di quello di sinistra è proprio Z , dunque si può affermare che $Z \geq Y$.)

Se una variabile è sempre minore uguale di un'altra, anche il suo valore atteso sarà minore uguale del valore atteso dell'altra. Da questa considerazione segue che $\rightarrow E[Y] \leq E[Z]$. Ma ora quale è il valore atteso di Y ? Il valore atteso di Y è $\epsilon^2 E[\chi_A]$. Quello di Z ? Per definizione è la $Var(x)$.

Dunque concludiamo con $\epsilon^2 E[\chi_A] \leq Var(X)$. Abbiamo $E[\chi_A] \leq \frac{Var(X)}{\epsilon^2}$, con $E[\chi_A] = P(A)$ Cosa abbiamo fatto? Abbiamo introdotto due variabili aleatorie, Y e Z , abbiamo osservato che, qualunque cosa accada Y è sempre minore di Z e di conseguenza anche $E[Y] \leq E[Z]$. Infine esplicitando i valori attesi si ottiene la disuguaglianza che si vuole dimostrare.

Abbiamo ora questa uguaglianza che vogliamo andare ad utilizzare per dimostrare la legge dei grandi numeri.

Cosa dice intuitivamente la legge dei grandi numeri?

Supponiamo di lanciare una moneta, supponiamo di effettuare 100 lanci e in questi 100 lanci di ottenere k volte testa. Ora cosa possiamo dire di k ? beh k non è detto che sia 50. Ci aspettiamo che k sia vicino ad 50, cioè il rapporto tra teste e lanci è vicino ad $\frac{1}{2}$. Se invece facessi 1000 o 10.000 lanci ci aspetteremmo che aumenti la probabilità che questo rapporto sia vicino ad $\frac{1}{2}$. La legge dei grandi numeri dà proprio una interpretazione quantitativa che permette di esprimere questo fatto in modo rigoroso. Stiamo parlando di

una convergenza di una successione di variabili aleatorie.

Formalizzando:

Abbiamo n lanci, risultati lanci = x_1, x_2, \dots, x_n . Ognuna che vale 1 o 0 a seconda se ottengo testa o croce.

Stiamo considerando il rapporto $\frac{x_1 + x_2 + \dots + x_n}{n} := \bar{x}_n$ (numero di teste ottenute su n lanci). Questa definizione la possiamo considerare per una qualunque successione x_1, x_2, \dots, x_n di variabili aleatorie.

Definizione

Prendiamo una successione infinita x_1, x_2, \dots di variabili aleatorie. Diciamo che questa successione converge ad un numero a se $\forall \epsilon$ se io vado a calcolare:

$\lim_{n \rightarrow \infty} P(|X_n - a| \geq \epsilon) = 0$ (l'ennesimo termine della successione sia più distante da ϵ che da a) Noi vogliamo dire che le variabili \bar{x}_n , che danno il rapporto di teste, si avvicinano ad $\frac{1}{2}$.

Legge dei grandi numeri

Supponiamo di avere una successione infinita di variabili aleatorie indipendenti e aventi tutte la stessa densità di media μ : x_1, x_2, \dots . Allora: Abbiamo che la successione \bar{x}_n converge a μ . Nel caso delle monete stiamo dicendo che il numero medio di teste ottenute converge ad $\frac{1}{2}$.

Dimostrazione

Basta applicare Chebyshev alla variabile \bar{x}_n . $E[\bar{x}_n] = E\left[\frac{x_1 + x_2 + \dots + x_n}{n}\right] = \mu$

$Var(\bar{x}_n) = Var\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) = \frac{1}{n^2} n Var(x_1)$ (Quando una varianza ha una costante è possibile portarla fuori elevata alla seconda).

0.3.29 Quale è il valore atteso e la varianza di una variabile di Bernoulli?

$E[x] = p, var(x) = p \cdot (1 - p)$

0.3.30 Definisci la variabile continua

Una variabile aleatoria X è continua se la sua funzione di ripartizione è a sua volta continua.

Gode di una importante proprietà, ovvero che $P(X = t) = 0$ per ogni numero reale t .

Dimostrazione

$$P(X = t) \leq P(X \leq t) - P(X \leq t - \frac{1}{n})$$

Per n che tende ad infinito diventa:

$$P(X = t) \leq 0 \rightarrow P(X = t) = 0$$

0.3.31 Quale è la densità continua?

La densità continua di una variabile aleatoria continua X è la funzione $f_x : \mathbb{R} \rightarrow \mathbb{R}$ definita nel seguente modo: $\int_a^b f_x(s)ds$.

Per essere definita tale è necessario che soddisfi due proprietà:

- $f_x(s) \geq 0$
- $\int_{-\infty}^{+\infty} f_x(s)ds = 1$

0.3.32 Che legame c'è fra una funzione di ripartizione continua e la densità

Dalla densità è possibile ottenere la funzione di ripartizione attraverso l'integrazione, viceversa dalla funzione di ripartizione è possibile ottenere la funzione della densità attraverso il processo di derivazione. Questo è dimostrabile nel seguente modo:

$$F_x(t) = P(x \leq t) = P(-\infty \leq x \leq t) = \int_{-\infty}^t f(s)$$

$$\int_a^b f_x(s) = P(a \leq x \leq b) = F_x(b) - F_x(a) \text{ (teorema fondamentale del calcolo integrale)}$$

0.3.33 Come si trova il valore atteso nell'ambito delle variabili continue?

$$E[x] = \int_{-\infty}^{+\infty} f_x(s) \cdot s \cdot ds$$

Le proprietà di cui gode sono le seguenti:

- $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$
- $E[X + Y] = E[X] + E[Y]$ (in caso di indipendenza)

0.3.34 Come si determina la varianza nell'ambito delle variabili continue?

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

Le proprietà di cui gode sono le seguenti:

- $Var(X + Y) = Var(X) + Var(Y)$
- $Var(\alpha X) = \alpha^2 Var(X)$

0.3.35 Definisci la densità continua uniforme

Una variabile continua uniforme, definita in un intervallo $[a, b]$, assume valori, in maniera uniforme, solo all'interno dell'intervallo considerato, al di fuori di esso la densità è $= 0$. Dunque vogliamo che la densità sia costante all'interno dell'intervallo selezionato.

$$f_x(s) = \frac{1}{b-a} \text{ e } F_x(s) = \int_a^b f(s)ds = \left[\frac{s}{b-a}\right]_a^b = \frac{s-a}{b-a}$$

Notiamo che la funzione di ripartizione di una variabile aleatoria continua non è altro che una retta che congiunge l'estremo sinistro con quello destro.

0.3.36 Definisci la densità continua astratta e la funzione di ripartizione astratta

La densità continua astratta è una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ che gode delle seguenti proprietà:

- $f_x(s) \geq 0$
- $\int_{-\infty}^{+\infty} f_x(s)ds = 1$

La funzione di ripartizione astratta è una funzione $F : \mathbb{R} \rightarrow \mathbb{R}$ che gode delle seguenti proprietà:

- $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow \infty} F(x) = 1$
- è debolmente crescente

0.3.37 Parlami della densità esponenziale

La variabile esponenziale è una v. aleatoria continua che gode della proprietà di mancanza di memoria, dunque la probabilità che un evento accada non dipende da ciò che è avvenuto prima. La sua densità è la seguente: $a \cdot e^{-as}$ per $s \geq 0$ (assume valore 0 altrimenti).

La funzione di ripartizione la si ottiene attraverso l'integrazione:

$$F_x(t) = \int_0^t a \cdot e^{-at} dt = [-e^{-at}]_0^t = 1 - e^{-at}$$

Inoltre è interessante sapere che $E[x] = \frac{1}{a}$, questo ci permette di determinare a senza problemi. Infine concludiamo affermando che $Var(x) = \frac{1}{a^2}$

0.3.38 Raccontami la densità normale

Le variabili normali sono definite in funzione di due parametri, μ e σ^2 , ovvero il valore atteso e la varianza. Si tratta di una distribuzione di probabilità continua che è spesso usata come prima approssimazione per descrivere variabili aleatorie a valori reali che tendono a concentrarsi attorno a un singolo valore medio. Il grafico della sua densità è una campana simmetrica, dunque assume valori con frequenza più alta al centro e tende a diminuire spostandosi negli estremi. La variabile normale standard, che indichiamo con ζ_0 non è altro che una variabile normale con $\mu = 0$ e $\sigma^2 = 1$, la sua funzione di ripartizione astratta è la seguente: $f(s) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-s^2}{2}}$.