

Domande Orali M.D.P (Caselli) v.2.3

Ce la faremo tutti, forse.

Questo documento è una collezione di domande raccolte da diversi orali e da diversi appelli. Le risposte alle domande sono frutto dell'unione di definizioni di Wikipedia, appunti vari presi a lezione, note di Caselli stesso. Potrebbero esserci errori o imprecisioni. Non considerate questo documento un riassunto su cui potete studiare. Usatelo piuttosto per ripassare insieme a qualcuno, o per testare se siete pronti per l'orale.

INOLTRE

Se volete aiutare il documento a migliorare potete trovare la repository su bitbucket:
<https://bitbucket.org/AngeloFilaseta/domande-orale-caselli/src/master/>
Contattatemi pure su bitbucket se volete ereditare la repo, voi degli anni a venire.
Aiuterete un sacco (ed intendo **UN SACCO**) di altra gente.
Grazie mille!

Grazie a:

Filaseta Angelo
Tronetti Elisa
Associazione Sprite
Wikipedia (it.wikipedia.org)

1. Cos'è la cardinalità di un insieme? Cos'è una corrispondenza biunivoca tra insiemi? Cosa vuol dire insieme numerabile?(x2)

Sia A un insieme, sia n il numero degli elementi contenuti in A . La cardinalità di A è n . Due insiemi hanno la stessa cardinalità se esiste una corrispondenza biunivoca tra essi, ovvero se è possibile associare a ciascun elemento di un insieme a uno ed un solo elemento dell'altro. Un insieme è numerabile se ha la stessa cardinalità di \mathbb{N} .

2. Esempio di insieme numerabile e non numerabile.

Un insieme numerabile è per esempio \mathbb{Z} . Un insieme non numerabile è \mathbb{R} , o anche un infinito insieme di infinite sequenze di bit.

3. Quanti sottoinsiemi ha un insieme con n elementi?

Si tratta della cardinalità dell'insieme delle parti. $P(A)$ è l'insieme i cui elementi sono sottoinsiemi di A , e la sua cardinalità è 2^n , perchè c'è una corrispondenza biunivoca con $\{0, 1\}^n$.

4. Se siamo solo interessati ai sottoinsiemi di una certa cardinalità invece come facciamo?

Potremmo usare il principio di inclusione-esclusione, o calcolare il numero di combinazioni della lunghezza interessata.

5. Cos'è una lista? Cos'è una disposizione? C'è una permutazione?

Una lista è una sequenza di numeri in cui ci possono essere ripetizioni. Una disposizione è una lista in cui non compaiono due elementi uguali.

$(2, 5, 1, 2)$ è una lista, ma non una disposizione. $(2, 5, 1, 7)$ è una disposizione. Le disposizioni di lunghezza k in un insieme di n elementi si scrivono $(n)_k$ e sono $(n) \cdot (n-1) \cdot \dots \cdot (n-k+1)$. Le disposizioni lunghe n in un insieme di n elementi si dicono permutazioni e $k!$.

6. Cos'è il prodotto cartesiano?

Siano A e B insiemi.

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Per esempio, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = (0, 0), (0, 1), (0, 2), \dots$

7. Cos'è uno scambussolamento? (x2)

Uno scambussolamento è una permutazione in cui nessun elemento rimane al suo posto. Per esempio (3 2 1 4) è uno scambussolamento di (1 2 3 4). Il numero di scambussolamenti di lunghezza n è

$$\sum_{i=0}^n (-1)^i (n)_{n-i}$$

.

8. Cosa sono i numeri di Fibonacci?(x2)

ATTENZIONE: NON USARE LA DEFINIZIONE DI WIKIPEDIA!

Non è sbagliata, è semplicemente definita in modo diverso.

$\forall n \geq 0$ denotiamo con F_n il numero di liste binarie in cui non compaiono due 1 consecutivi.

$$F_0 = 1, \quad F_1 = 2, \quad F_2 = 3, \quad F_n = F_{n-1} + F_{n-2}$$

Dimostrazione, presa dalle note:

Chiamiamo G_n il numero di sequenze binarie di lunghezza n in cui non compaiono due 1 consecutivi e che terminano con 0 e H_n il numero di sequenze binarie di lunghezza n in cui non compaiono due 1 consecutivi e che terminano con 1. Si ha evidentemente $F_n = G_n + H_n$. Ora, le sequenze utilizzate per G_n sono sequenze di lunghezza $n-1$ senza 1 consecutivi a cui aggiungiamo uno 0: abbiamo quindi $G_n = F_{n-1}$. Le sequenze utilizzate per H_n devono terminare necessariamente con 01 e sono quindi formate da sequenze di lunghezza $n-2$ senza 1 consecutivi a cui aggiungiamo 01. Abbiamo quindi $H_n = F_{n-2}$. \square

9. Spiegami il principio di inclusione-esclusione. (x3)

Vogliamo calcolare la cardinalità dell'unione di n insiemi. Sommiamo quindi la cardinalità dei singoli insiemi. Togliamo poi tutti quelli che erano comuni a due insieme (ovvero gli elementi nell'intersezione), perché contati due volte. Vanno poi aggiunti di nuovo gli elementi in comune a tre insiemi, che abbiamo tolto due volte nella scorsa operazione e così via. La formula è la seguente:

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{S=\{1..n\}} (-1)^{|S|+1} \left| \bigcap_{i \in S} A_i \right|$$

.

10. Cos'è una combinazione? Cos'è un coefficiente binomiale? (x2)

Ad ogni disposizione di lunghezza k , possiamo associare un sottoinsieme di cardinalità k , "dimenticandoci" dell'ordine. Per esempio: $(4, 1, 2), (1, 2, 4), (2, 1, 4) \rightarrow \{1, 2, 4\}$

Ogni sottoinsieme è associato ad esattamente $k!$ disposizioni, cioè tutte le sue permutazioni. Le disposizioni sono quindi $k!$ volte i sottoinsiemi, e le denotiamo con:

$$\binom{n}{k} = \frac{(n)_k}{k!}$$

11. Quanto fa $\binom{4}{0} + \binom{4}{1} + \binom{4}{2} + \binom{4}{3} + \binom{4}{4}$ senza fare conti?

Si tratta di tutti i sottoinsiemi di un insieme con 4 elementi, quindi la somma è 2^4 che è $P(\{1, 2, 3, 4\})$ (insieme delle parti).

Oppure facendo i conti...

$\binom{4}{0}$ e $\binom{4}{4}$ sono uguali a 1. $\binom{4}{1}$ è 4, $\binom{4}{2}$ è 6 perchè posso piazzare 2 numeri vicini 3 volte in 4 posti, moltiplicato per 2 che sono le permutazioni dei 2 numeri. $\binom{4}{3} = \binom{4}{1} = 4$

12. Quante sono le combinazioni di tipo (a,b,c)?

Le combinazioni di tipo (a,b,c) in A è una terna ordinata $(A1, A2, A3)$ di sottoinsiemi di A tali che:

$$|A1| = a, |A2| = b, |A3| = c \quad \text{ed} \quad A1 \cup A2 \cup A3 = A$$

Se $a, b, c \in \mathbb{N}$ ed $a + b + c = n$ allora le combinazioni di tipo (a,b,c) formano un prodotto condizionato di tipo:

$$\left(\binom{n}{a}, \binom{n-a}{b}, 1 \right)$$

13. Qual'è la formula di Stifel? A cosa serve? Dimostrazione.(x4)

Se $0 < k < n$ allora

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

Dimostrazione presa dalle note:

Questa dimostrazione si può fare in modo algebrico ma preferiamo farla in modo combinatorio. Infatti abbiamo che $\binom{n-1}{k-1}$ è il numero di sottoinsiemi con k elementi e che contengono n , mentre $\binom{n-1}{k}$ è il numero di sottoinsiemi con k elementi e che NON contengono n . Il risultato segue. \square

La dimostrazione è scritta un po' male, con "Infatti abbiamo che $\binom{n-1}{k-1}$ è il numero di sottoinsiemi con k elementi" intende dire i sottoinsiemi con $k-1$ elementi a cui va aggiunto n come ultimo valore. Per capire meglio provate a scrivere tutte le combinazioni $\binom{3}{2}$ e aggiungete 4 alla fine di ognuna, poi scrivete l'unica di $\binom{3}{3}$.

Magia.

Generalizzando:

$$\binom{n}{a, b, c} = \binom{n-1}{a-1, b, c} + \binom{n-1}{a, b-1, c} + \binom{n-1}{a, b, c-1}$$

14. Cos'è un cammino reticolare?

Consideriamo i cammini che uniscono l'origine con il punto (m, n) , con $m, n \in \mathbb{N}$ nel piano cartesiano costituiti da una sequenza di passi unitari verso destra e verso l'alto. Quanti sono tali cammini? È evidente che ad ogni passo verso destra l'ascissa aumenta di 1 e ad ogni passo verso l'alto l'ordinata aumenta di 1.

Ne segue che complessivamente bisogna effettuare m passi verso destra e n passi verso l'alto. Abbiamo quindi che ogni cammino corrisponde ad una sequenza binaria in cui gli 0 corrispondono ai passi verso destra e gli 1 ai passi verso l'alto in cui 0 compare esattamente m volte, e 1 compare esattamente n volte.

Ogni cammino resta quindi individuato dalla scelta delle m posizioni (tra le $m+n$ disponibili) in cui inserire gli 0 e quindi tali cammini sono esattamente $\binom{m+n}{m}$.

15. Parlami delle funzioni iniettive e suriettive.

Riassuntazzo: una funzione iniettiva è una funzione che associa, a elementi distinti del dominio, elementi distinti del codominio. In altre parole: una funzione da un insieme A a un insieme B è iniettiva se ogni elemento di B non può essere ottenuto in più modi diversi partendo dagli elementi di A .

Una funzione si dice invece suriettiva quando ogni elemento del codominio è immagine di almeno un elemento del dominio. In tal caso si ha che l'immagine coincide con il codominio. Siano m il numero di elementi dell'insieme A ed n il numero degli elementi in B . Esistono in totale n^m funzioni. Quante le iniettive?

Le funzioni iniettive da A a B sono tante quante le disposizioni in B di lunghezza m , cioè $(n)_m$, ovviamente se $m \leq n$.

Per trovare il numero di funzioni suriettive, è richiesto l'uso del principio di inclusione-esclusione.

È importante notare che $m \geq n$ in questo caso.

Denotiamo con A_1 l'insieme contenente tutte le funzioni che non hanno l'elemento 1 nell'immagine. Denotiamo A_2 allo stesso modo fino ad A_n . Se K è l'insieme contenente tutte le funzioni suriettive, allora $K = (A_1 \cup A_2 \cup \dots \cup A_n)^C$.

Quindi:

$$\begin{aligned} |A_i| &= (n-1)^m & |A_i \cap A_j| &= (n-2)^m & \dots \\ |K| &= \sum_{S=\{1\dots n\}}^n (-1)^{|S|} \left| \bigcap_{i \in S} A_i \right| = \\ |K| &= \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)^m \end{aligned}$$

16. Cos'è una partizione? Cosa sono i numeri di Bell? E quelli di Stirling?
(x4)

Sia A un insieme.

Una partizione di A in k blocchi è un insieme di k sottoinsiemi non vuoti

$$\{S_1, S_2, \dots, S_k\} \quad t.c. \quad S_i \cap S_j = \emptyset \quad \forall i \neq j \quad \text{ed} \quad S_1 \cup \dots \cup S_k = A$$

I numeri di Bell denotati come B_n sono il numero di partizioni di $\{1, 2, \dots, n\}$. La formula ricorsiva per calcolare i numeri di Bell è la seguente:

$$B_n = \sum_{k=1}^n \binom{n-1}{k-1} B_{n-k}$$

I numeri di Stirling denotati con $S_{n,k}$ sono le partizioni in k blocchi di $\{1, 2, \dots, n\}$. La formula per calcolare i numeri di Stirling è la seguente:

$$S_{n,k} = \frac{1}{k!} \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n$$

Che equivale alla formula per calcolare il numero di funzioni suriettive da n a k diviso k!

17. Cos'è un indice di dispersione?

Un indice di dispersione serve per descrivere sinteticamente una distribuzione statistica quantitativa, e in modo particolare la misura con la quale i suoi valori sono distanti da un valore centrale (identificato con un indice di posizione, solitamente media o mediana).

18. Cos'è la varianza? Come si trova? Dimostrazione

La Varianza è un indice di dispersione ed un indicatore della variabilità di un insieme di dati. Un valore basso significa che i dati sono raggruppati molto vicini fra loro, mentre una varianza elevata indica dei dati più distribuiti. Scegliamo un valore t e consideriamo gli scostamenti dei valori x_1, x_2, \dots, x_n da t .

$$\frac{(x_1 - t) + (x_2 - t) + \dots + (x_n - t)}{n}$$

Per praticità usiamo i quadrati:

$$\frac{(x_1 - t)^2 + (x_2 - t)^2 + \dots + (x_n - t)^2}{n}$$

Si sceglie t in modo che questa quantità sia minima. Fissiamo quindi t come \bar{x} . La varianza, nella sua forma finale è quindi;

$$\sigma_x^2 = \frac{1}{n} ((x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2)$$

Si può anche scrivere:

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2$$

Dimostrazione sulle note...

19. Covarianza. Cos'è? Come si trova? (x2)

La covarianza di due variabili statistiche (x e y) o variabili aleatorie è un numero che fornisce una misura di quanto le due varino assieme, ovvero della loro dipendenza. La formula è:

$$\sigma_{x,y} = \frac{1}{n}((x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y}))$$

Può essere positiva negativa o nulla. Se la covarianza è positiva ci aspettiamo che i termini $(x_1 - \bar{x})(y_1 - \bar{y})$ siano più facilmente positivi (a valori di x grandi corrispondono valori grandi di y).

Se la covarianza è negativa, ci si aspetta che al crescere dei valori di x, quelli di y decrescono. Si può anche scrivere:

$$\sigma_{x,y} = \overline{xy} - \bar{x} \bar{y}$$

20. Cos'è un indice di correlazione? Come capisco se ho una buona approssimazione?

Date due variabili statistiche X e Y, l'indice di correlazione di Pearson è definito come la loro covarianza divisa per il prodotto delle deviazioni standard (varianza sotto radice quadrata) delle due variabili.

$$\rho_{x,y} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y}$$

$\rho_{x,y}$ è un valore per descrivere la bontà dell'approssimazione della retta ai minimi quadrati. È un numero sempre compreso tra -1 e 1. Se siamo vicini allo 0 l'approssimazione è pessima, se siamo vicini a -1 o 1 invece abbiamo una buona approssimazione.

21. Cos'è un indice di posizione? A cosa serve la media campionaria? Come si calcola? Dimostrazione (x2)

Un indice di posizione è un indice usato per dare un'idea approssimata dell'ordine di grandezza (la posizione sulla scala dei numeri, appunto) dei valori esistenti e scelti. Sono indici di posizione la Media, la Moda e la Mediana. Se un carattere x assume i valori x_1, \dots, x_n su un campione di n elementi, la media campionaria è:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

22. A cosa serve la media geometrica? Come si calcola?(x3)

La media geometrica di n termini è la radice n -esima del prodotto degli n valori:

$$M_g = \sqrt[n]{\prod_{i=1}^n x_i}$$

La media geometrica trova impiego soprattutto dove i valori considerati vengono per loro natura moltiplicati tra di loro e non sommati. Esempio tipico sono i tassi di crescita, come i tassi d'interesse o i tassi d'inflazione.

Una caratteristica è che valori piccoli sono molto più influenti rispetto a quanto succede nella media aritmetica. In particolare, è sufficiente la presenza di un unico valore nullo per annullare la media geometrica.

23. Retta ai minimi quadrati, qual'è il suo scopo? Come si trova? Che cos'è? Definizione. (x2)

Dati x e y due caratteri. Decidiamo di riportare sulle ascisse i valori di x e sulle ordinate quelli di y , perché vogliamo studiare come x influenzi y . Se avessimo voluto studiare che influenza ha y su x avremmo fatto la scelta opposta. Lo scopo è quello di determinare una retta che approssimi nel migliore dei modi possibili i dati. La retta ai minimi quadrati è la retta di equazione $y = ax + b$ che minimizza la media (o equivalentemente la somma) dei quadrati degli errori che si commettono stimando i valori y_i del secondo carattere con $ax_i + b$. Considerando una retta generica di equazione $y = ax + b$ dipendente dai parametri a e b si vuole cioè minimizzare la quantità:

$$S(a, b) = \frac{1}{n} \sum_{i=1}^n (y_i - ax_i - b)^2$$

Si sceglie $a = \frac{\sigma_{x,y}}{\sigma_x^2}$ e $b = \bar{y} - a\bar{x}$. *Dimostrazione sul perché sulle note...*

24. Cos'è uno spazio di probabilità?

Sia dato un insieme Ω (di possibili risultati di un fenomeno aleatorio). Gli eventi saranno per noi i sottoinsiemi di Ω di cui vorremo e potremo calcolare le probabilità. Pertanto, NON HA SENSO calcolare la probabilità di un sottoinsieme di Ω che non sia un evento. È chiaro quindi che ci piacerebbe che unioni, intersezioni e complementari di eventi siano ancora eventi. Formalizziamo questo desiderio dopo aver introdotto la seguente terminologia. Se E_1, E_2, \dots (finiti o infiniti) sono eventi diciamo che formano una successione di eventi.

Sia A una collezione di sottoinsiemi di Ω . Diciamo che A è una famiglia coerente di eventi su Ω se:

$\emptyset, \Omega \in A$

$E^C \in A$ se $E \in A$

Se E_1, E_2, \dots sono in A allora anche la loro unione e la loro intersezione stanno in A .

25. Qual'è la formula delle probabilità totali? Come funziona? A cosa serve? (x2)

Se $\{A_1, A_2, \dots\}$ è una partizione in eventi di Ω si ha per ogni evento B :

$$\sum_{i=1}^n P(B|A_i)P(A_i)$$

Possiamo vedere B come un evento in Ω , che può presentarsi in qualsiasi blocco A_i . Per calcolare la probabilità di B serve quindi calcolare la probabilità che questo possa accadere per ogni blocco.

26. Cos'è una variabile aleatoria? (x5)

È una quantità che dipende dal risultato del fenomeno aleatorio. Formalmente è una funzione

$$X : \Omega \rightarrow \mathbb{N}$$

Per esempio, il fenomeno è il lancio di due dadi:

Ω sono le coppie da 1 a 6.

X è la somma dei due valori usciti.

Per essere considerata variabile aleatoria, X deve soddisfare la seguente proprietà:

$$\{\omega \in \Omega : X(\omega) \leq t\}$$

è un evento $\forall t \in \mathbb{R}$

Vedi le note per maggiori dettagli, è importante sapere bene la risposta a questa domanda.

27. Cos'è una variabile aleatoria continua?(x2)

Intuitivamente, le variabili casuali continue sono quelle che possono assumere un insieme continuo di valori, al contrario delle distribuzioni discrete, per le quali l'insieme dei possibili valori ha cardinalità al più numerabile.

Una variabile aleatoria x si dice continua se $F_x(t) = P(x \leq t)$ è una funzione continua. Allora la funzione di ripartizione è sempre crescente e compresa tra 0 e 1.

Se x è continua e fissiamo $t \in \mathbb{R}$, allora $P(x = t) = 0$.

28. Cosa intendiamo per schema successo insuccesso?

Uno schema successo-insuccesso è una sequenza di un numero finito 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).

Esempio: L'estrazione di n palline colorate da un'urna in cui vogliamo pescare k palline blu è uno schema successo- insuccesso. Successo: pescò una pallina blu. Insuccesso: non la pescò.

29. Cos'è una variabile aleatoria geometrica? (x2)

In uno schema successo-insuccesso a prove indipendenti.

X = numero di successi prima di ottenere un successo.

Si può considerare ad esempio il lancio di un dado in cui il successo è dato dal risultato 6. Allora il parametro p è $\frac{1}{6}$ e scriviamo la variabile geometrica $X \sim G(p)$. La densità è:

$$pX(k) = \begin{cases} p(1-p)^k & \text{se } k = 0, 1, 2, \dots \\ 0 & \text{altrimenti} \end{cases}$$

Esiste anche la variabile geometrica modificata, denotata con $X \sim \tilde{G}(p)$ la cui densità è:

$$pX(k) = \begin{cases} p(1-p)^{k-1} & \text{se } k = 1, 2, 3, \dots \\ 0 & \text{altrimenti} \end{cases}$$

Non è altro che una variabile geometrica standard a cui è stato effettuato uno shift di uno.

In altre parole, la variabile geometrica ci dice che dopo k insuccessi avremo un successo, la geometrica modificata ci dice che il k -esimo tentativo è il successo.

Se la variabile è geometrica standard $E[X] = \frac{1-p}{p}$

Se la variabile è geometrica modificata $E[X] = \frac{1}{p}$

30. Qual'è il valore atteso di una variabile aleatoria di Bernoulli? (x2)

$$X \sim B(1, p), \quad E[X] = (1)p + 0(1-p) = p$$

31. Cos'è una variabile aleatoria binomiale? La sua densità? Il valore atteso? (x3)

In uno schema successo insuccesso a prove indipendenti, siamo interessati a studiare il numero dei successi. Allora X è una variabile binomiale di parametro p (probabilità del

successo) e si scrive $X \sim B(n, p)$ dove n è il numero tentativi effettuati. La sua densità è:

$$pX(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{se } k = 0, 1, 2 \dots n \\ 0 & \text{altrimenti} \end{cases}$$

$E[X] = np$ (dimostrazione sulle note, usando la linearità del valore atteso)

32. Cos'è una variabile ipergeometrica? Qual'è la sua densità? Quale il valore atteso?(x4)

È uno schema successo-insuccesso senza ripetizione. Per esempio il fenomeno aleatorio consiste nell'estrazione di n palline, senza rimpiazzare di volta in volta la pallina estratta, da una scatola contenente b bianche e r rosse. Diciamo che la prova dà successo se viene estratta una pallina bianca e consideriamo la variabile X data dal numero di successi ottenuti, cioè dal numero di palline bianche estratte. Lo spazio Ω è dato da tutti i possibili sottoinsiemi di n palline scelte da un insieme di $b + r$ palline (tutte), con probabilità uniforme.

L'evento $X = k$ è dato quindi da tutti i sottoinsiemi delle $b + r$ palline costituiti da k palline bianche e $n - k$ palline rosse. Le k bianche le posso scegliere in $\binom{b}{k}$ modi. Le $n - k$ rosse in $\binom{r}{n-k}$ modi.

La densità è quindi:

$$pX(k) = \begin{cases} \frac{\binom{b}{k} \binom{r}{n-k}}{\binom{b+r}{n}} & \text{se } k = 0, 1, 2 \dots \min(n, b) \\ 0 & \text{altrimenti} \end{cases}$$

e scriviamo $X \sim H(n, b, r)$

$E[X] = \frac{nb}{b+r}$ (dimostrazione sulle note, usando la linearità del valore atteso).

33. Cos'è una variabile di Poisson?(x2)

Pensiamo a una variabile binomiale con n molto grande e p molto piccolo. Usare una variabile binomiale risulterebbe scomodo, perchè la formula della densità binomiale ha un carico computazionale altissimo per questi tipi di numeri. Introduciamo quindi $\lambda = np$,

$$p = \frac{\lambda}{n}$$

$$X \sim B(n, \frac{\lambda}{n})$$

$$P_X(k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

Per $n \rightarrow \infty$, questo valore tende a $\frac{\lambda^k}{k!} e^{-\lambda}$ X allora è una variabile di Poisson di parametro λ e si scrive $X \sim P(\lambda)$, la cui densità è:

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

$$E[X] = \lambda \text{ perchè } X \sim B(n, \frac{\lambda}{n}) \rightarrow E[X] = n \frac{\lambda}{n} = \lambda$$

34. Cos'è una variabile aleatoria continua uniforme? La sua densità? Il valore atteso? (x3)

Si tratta di una variabile aleatoria continua che attribuisce la stessa probabilità a tutti i punti appartenenti ad un dato intervallo $[a, b]$. La funzione di ripartizione è:

$$Fx(t) = \begin{cases} 1 & t > b \\ 0 & t < a \\ \frac{t-a}{b-a} & t \in [a, b] \end{cases}$$

La densità è:

$$fx(s) = \begin{cases} 0 & s < a, s > b \\ \frac{1}{b-a} & s \in [a, b] \end{cases}$$

Il valore atteso è il punto medio dell'intervallo $[a, b]$ ovvero $\frac{a+b}{2}$.

35. Cos'è una variabile aleatoria esponenziale? La sua densità? Il valore atteso? (x3)

Una variabile aleatoria esponenziale è una variabile aleatoria continua, utile per modellare alcuni fenomeni naturali, per esempio il "tempo di decadimento di una particella radioattiva" o il tempo di eruzione di un certo vulcano.

Si denota con $X \sim Exp(a)$ in cui a è il parametro, $a \in \mathbb{R}, a > 0$.

Le variabili aleatorie esponenziali godono della proprietà di mancanza di memoria, come le variabili geometriche:

$P(X \geq t+u | X \geq u) = \frac{P(X \geq t+u)}{P(X \geq u)} = P(X \geq t)$ La funzione di ripartizione è:

$$Fx(t) = \begin{cases} 1 - e^{-at} & t > 0 \\ 0 & t < 0 \end{cases}$$

La densità è:

$$fx(s) = \begin{cases} ae^{-as} & s > 0 \\ 0 & s < 0 \end{cases}$$

36. Cos'è una variabile aleatoria bidimensionale?

È una variabile aleatoria multidimensionale in cui $n = 2$.

37. Cos'è una variabile aleatoria n-dimensionale? (Multidimensionali)

Una variabile aleatoria di dimensione n è data da n variabili aleatorie standard che denotiamo con x_1, \dots, x_n . Le densità delle variabili x_i si dicono marginali, mentre la densità di (x_1, \dots, x_n) si dice congiunta.

38. Se abbiamo due densità marginali uguali, possiamo dedurre la congiunta? Quando non possiamo? Mi fai un esempio?(x2)

Si può ricavare la densità congiunta di una variabile multidimensionale solo se tutte le variabili (x_1, \dots, x_n) che la compongono sono indipendenti:

$$p_{x_1, \dots, x_n}(k_1, \dots, k_n) = p_{x_1}(k_1) \cdot \dots \cdot p_{x_n}(k_n)$$

Per esempio, sia $Y = x_1 + x_2 + x_3$, dove le x_i sono i risultati dei lanci di tre dadi. Sapendo le tre densità delle x_i è possibile ricavare la congiunta, ovvero la densità di Y perché i tre lanci sono eventi indipendenti.

39. Come si calcola la densità marginale?

Le densità marginali di una variabile aleatoria multidimensionale non sono altro che le densità delle singole variabili (x_1, \dots, x_n) , quindi dipende dal tipo di variabili aleatorie.

40. Trasformazione di variabili, cos'è il massimo di due variabili aleatorie? E il minimo?

Siano X e Y due variabili aleatorie discrete.

Il massimo si tratta di una variabile multidimensionale (bidimensionale nell'esempio) di tipo $Z = \max(X, Y)$, ed è il numero aleatorio che indica il massimo tra i due precedenti. In altre parole, si tratta di calcolare la probabilità che capitano sia uno che l'altro evento desiderato.

Per definizione di massimo:

$$P(\max(X, Y) \leq t) = P(X \leq t, Y \leq t)$$

Se le due variabili sono indipendenti, allora:

$$P(X \leq t, Y \leq t) = P(X \leq t)P(Y \leq t)$$

Quindi, la probabilità congiunta è uguale al prodotto delle probabilità marginali:

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

Lo stesso principio vale per il minimo. In questo caso vogliamo calcolare la probabilità che almeno uno dei due eventi capiti (ci fermiamo non appena uno dei due si verifica) $W = \min(X, Y)$, e siamo interessati a calcolare $1 - F_W(t)$ dove $F_W(t)$ è comunque il prodotto delle due funzioni di ripartizione.

$$1 - F_W(t) = P(W > t) = P(\min(X, Y) > t) = P(X > t, Y > t) = (1 - F_X(t))(1 - F_Y(t))$$

(Dimostrazione ed esempio nelle note...)

41. Come trovo la varianza di una variabile aleatoria discreta? Che informazioni ne ricavo?(x2)

In statistica descrittiva, la varianza di un carattere è la media dei quadrati degli scarti rispetto alla media:

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

o equivalentemente:

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

(Dimostrazione sulle note... Comunque basta sviluppare il quadrato.)

Come la varianza nella statistica descrittiva, un valore basso significa che i dati sono raggruppati molto vicini fra loro, mentre una varianza elevata indica dei dati più distribuiti.

42. Come trovo, generalmente, il valore atteso di una variabile aleatoria discreta? (x5)

Nel caso di variabile casuale discreta:

$$E[X] = \sum_{i=1}^{\infty} x_i p_i$$

Il valore atteso non è altro che una media pesata in cui il peso è la probabilità.

43. Cosa intendiamo per mancanza di memoria? Come scriviamo questa proprietà? Dimostrazione. (x4)

La mancanza di memoria è una proprietà di alcune variabili aleatorie (geometriche e esponenziali):

$$P(X \geq t + u | X \geq u) = \frac{P(X \geq t + u)}{P(X \geq u)} = P(X \geq t)$$

Sostanzialmente questa formula cela il fatto che se un certo evento non è ancora accaduto, sapere per quanto tempo non è accaduto non modifica la probabilità che questo accada.

(Dimostrazione nelle note, dipende dalla variabile...)

44. Come trovo la varianza della somma di due variabili aleatorie discrete? Dimostrazione

In generale, non vale $Var(x + y) = Var(x) + Var(y)$. Vale solo se x e y sono incorrelate (la loro covarianza è 0) e sono indipendenti.

45. Come trovo la densità di una variabile continua?

Osserviamo che $F(b) - F(a) = \int_a^b f(s)ds$. Il teorema fondamentale del calcolo integrale ci dice che tale formula vale se F è derivabile e $F'(t) = f(t)$. Abbiamo quindi il seguente fatto: se una variabile continua X ha funzione di ripartizione $F(t)$ derivabile, allora $f(t) = F'(t)$ è la densità di X .

46. Cosa vuol dire probabilità condizionale (condizionata)? Qual'è la formula di Bayes? (x2)

Siano A e B eventi. Denotiamo con $P(A|B)$ la probabilità che si verichi A , sapendo che B è verificato.

$P(A|B) = \frac{P(A \cap B)}{P(B)}$. Spesso questa formula viene usata anche al contrario:

$P(A \cap B) = P(A|B) \cdot P(B)$. La formula di Bayes viene utilizzata per calcolare $P(A|B)$ quando è molto facile calcolare $P(B|A)$:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

(Dimostrazione per sostituzioni...)

47. Cos'è una funzione di ripartizione? Un esempio?

La funzione di ripartizione di una variabile aleatoria X a valori reali è la funzione che associa a ciascun valore x la probabilità del seguente evento: "la variabile aleatoria X assume valori minori o uguali ad x ".

In altre parole, è la funzione $F: \mathbb{R} \rightarrow [0, 1]$ con dominio la retta reale e immagine nell'intervallo $[0, 1]$ definita da $F(x) = P(X \leq x)$. Una funzione F è una valida funzione di ripartizione se è non decrescente, continua a destra e:

$$F(x) \geq 0, \quad \forall x$$

$$\lim_{x \rightarrow +\infty} F(x) = 1$$

$$\lim_{x \rightarrow -\infty} F(x) = 0$$

Vedi una qualsiasi variabile aleatoria continua per l'esempio.

48. Cos'è la linearità di un valore atteso? Ha altre proprietà?

Un'importante caratteristica del valore atteso è la linearità: ovvero per ogni variabile casuale X e coppia di numeri reali a e b si ha: $E[aX + b] = aE[X] + b$. Questa proprietà è facilmente dimostrabile: ad esempio, nel caso di una variabile casuale discreta, si ha

$$E[aX + b] = \sum_{i=1}^{\infty} (ax_i + b)P(X = x_i) = a \sum_{i=1}^{\infty} x_i P(X = x_i) + b \sum_{i=1}^{\infty} P(X = x_i) = aE[X] + b$$

perché la somma delle probabilità è 1, in quanto consideriamo la somma di tutti i possibili eventi.

Questa proprietà ha la conseguenza importante che date due variabili casuali qualsiasi X e Y (non necessariamente indipendenti) si ha:

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

Questa proprietà non vale per il prodotto: in generale, $E[XY]$ è diverso da $E[X]E[Y]$. Quando queste due quantità sono uguali, si dice che X e Y sono non correlate. In particolare, due variabili aleatorie indipendenti sono non correlate.

49. Cos'è la distribuzione gaussiana (o normale)?(x3)

La distribuzione normale dipende da due parametri, la media μ e la varianza σ^2 , ed è indicata tradizionalmente con:

$N(\mu, \sigma^2)$. 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 funzione di densità di probabilità associata è simmetrico e ha una forma a campana.

Viene usata principalmente attraverso il teorema centrale del limite.

La distribuzione normale è caratterizzata dalla seguente funzione di densità di probabilità:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{con } x \in \mathbb{R}.$$

Dove μ è il valore atteso e σ^2 la varianza.

50. Cos'è il teorema centrale del limite?

Siano x_1, \dots, x_n (con n GRANDE) variabili aleatorie indipendenti aventi tutte la stessa densità (discreta o continua). Siano $\mu = E[X_1]$, $\sigma^2 = Var(x_1)$.

Allora $x_1 + \dots + x_n \sim N(n\mu, n\sigma^2)$ Questo perché più aumenta il numero di variabili, più ci avviciniamo a una densità normale.

51. Cos'è la legge dei grandi numeri? Dimostrazione.

Se lanciamo una moneta n volte e otteniamo k volte testa, è difficile ottenere $\frac{k}{n} = \frac{1}{2}$. Tuttavia, l'intuizione ci suggerisce che se n è molto grande, il rapporto k/n non debba discostarsi troppo da $\frac{1}{2}$. Abbiamo a che fare in questo caso con una sequenza di variabili X_1, X_2, \dots tutte di densità $B \sim (1, \frac{1}{2})$.

La legge dei grandi numeri permette di dare un senso preciso a questo tipo di intuizione. Abbiamo bisogno di un concetto di convergenza per le variabili aleatorie.

La disuguaglianza di Chebyshev dice:

X variabile aleatoria discreta. $\epsilon > 0$

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

Dimostrazione

$A = \{|x - E[x]| > \epsilon\}$, prendiamo $Y = \epsilon^2 \chi_A$ e $Z = (x - E[x])^2$. Sappiamo che $Z \geq Y$ sempre, sia se A accade sia non accade, e quindi $E[Z] \geq E[Y]$. Di conseguenza:

$$\epsilon^2 E[\chi_A] \leq E[(x - E[x])^2]$$

$$P(A) \leq \frac{Var(x)}{\epsilon^2} \square$$

Sia ora x_1, x_2, \dots, x_n una successione di variabili aleatorie discrete indipendenti, aventi tutte la stessa densità.

Allora

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{x_1 + \dots + x_n}{n} - E[x_1]\right| > \epsilon\right) = 0$$

Dimostrazione applicando Chebyshev a $\bar{x} = \frac{x_1 + \dots + x_n}{n}$