Variabili Aleatorie Discrete - Sommario

X

Tutto sulle v.a. discrete (vettori aleatori esclusi)

X

A. INTRODUZIONE ALLE VARIABILI ALEATORIE

A1. Considerazioni preliminari

Considerazione Preliminare per le Variabili Aleatorie

X

Considerazioni Preliminari per le Variabili Aleatorie

X

0. Preambolo

PREAMBOLO. Finora abbiamo sempre *costruito* i modelli probabilistici $(\Omega, \mathcal{A}, p)$ per rispondere alle domande. Tuttavia, vogliamo semplificarci la vita trovando una maniera per *evitare* di dover costruire questi spazi probabilistici da capo. Abbiamo due giustificazioni per farlo: una giustificazione *matematica* e una di natura *pratica*.
PUNTO DI VISTA MATEMATICO. Vogliamo spostare *tutto* ciò che sappiamo sui spazi probabilistici sulla *retta reale* $\mathbb{R}$ (o volendo anche sullo spazio $\mathbb{R}^N$), dato che in questo modo potremo usare gli strumenti dell'*analisi matematica*.
PUNTO DI VISTA PRATICO. Possiamo *misurare* le probabilità su *intervalli*.
SOLUZIONE. Una soluzione per affrontare questa questione sono le c.d. *"variabili aleatorie"*, delle funzioni invertibili del tipo $X: \Omega \longrightarrow E$. Vedremo questo concetto fino a fondo nelle prossime pagine.

A2. Definizione di variabile aleatoria

Definizione di Variabile Aleatoria

X

Definizione generale di variabile aleatoria. Lemma di caratterizzazione delle variabili aleatorie. Notazione per gli eventi generati dalle variabile aleatorie. Osservazione sui spazi probabilistici discreti.

X

0. Voci correlate

- Spazio di Probabilità Discreto
- Strutture Matematiche della Probabilità

1. Definizione di Variabile Aleatorie

#Definizione

Definizione (variabile aleatoria).

Sia $(\Omega, \mathcal{A}, p)$ uno *spazio di probabilità*. Definiamo *variabile aleatoria* (abbreviato come *v.a.*) un'applicazione $X : \Omega \longrightarrow E$ tale che per ogni $B \in \mathcal{B} \subseteq E$ (dove $\mathcal{B}$ è un *"insieme decente"*) vale l'applicazione inversa

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$$

In parole, una *variabile aleatoria* ci permette di prendere un *"sottoinsieme decente"* di E e riportarlo su Ω dentro la famiglia $\mathcal{A}$, potendo così misurare la sua probabilità con p.

In particolare, se $E = \mathbb{R}$ allora si dice che X è una *"variabile aleatoria reale"* e si sceglie $\mathcal{B}$ come la *sigma-algebra di Borel* in $\mathbb{R}$ (1).

#Osservazione

Osservazione (ogni funzione di una variabile aleatoria per uno spazio discreto è v.a.).

Osserviamo che quando $(\Omega, \mathcal{A}, p)$ è uno *spazio di probabilità discreto* con $\mathcal{A}$ la *famiglia di tutti i sottoinsiemi* di Ω, abbiamo che $X : \Omega \longrightarrow \mathbb{R}$ è *sempre* una *variabile aleatoria*, dato che $\mathcal{A}$ comprende *tutti* i sottoinsiemi del tipo $\{X \in B\}, \forall B \in \mathcal{B}$.

2. Lemma di Caratterizzazione per v.a. reali

#Lemma

Lemma (di caratterizzazione delle v.a. reali).

Una funzione $X : \Omega \longrightarrow \mathbb{R}$ è una *variabile aleatorie reale* se e solo se una delle seguenti condizioni è soddisfatta:

$$\begin{cases} X^{-1}((a,b)) \in \mathcal{A} = \{\omega \in \Omega : X(\omega) \in (a,b)\}, \forall a,b \in \mathbb{R} \\ X^{-1}([a,b]) \in \mathcal{A} = \{\omega \in \Omega : X(\omega) \in [a,b]\}, \forall a,b \in \mathbb{R} \\ X^{-1}((-\infty,a)) \in \mathcal{A} = \{\omega \in \Omega : X(\omega) \in (-\infty,a)\}, \forall a \in \mathbb{R} \\ X^{-1}((-\infty,a]) \in \mathcal{A} = \{\omega \in \Omega : X(\omega) \in (-\infty,a]\}, \forall a \in \mathbb{R} \end{cases}$$

#Dimostrazione

DIMOSTRAZIONE del Lemma 3 (di caratterizzazione delle v.a. reali)
N.B. Questa è un'idea della dimostrazione
Per dimostrare che tutte le quattro condizioni sono *equivalenti*, basta considerare che tutti gli intervalli elencati sopra sono degli *"insiemi approssimabili"* con degli intervalli del tipo (m,n); dopodiché si vede che l'*unione* o *l'intersezione* o il *passaggio al complementari* commutano con X^{-1}; ad esempio si ha

$$X^{-1}((-\infty,a)) = X^{-1}\left(\bigcup_{k=1}^{\infty}(a-k,a)\right) = \bigcup_{k=1}^{\infty} X^{-1}(a-k,a)$$

che prova il primo punto implica il terzo punto. ■

3. Eventi generati dalle V.A.

#Definizione

Definizione (eventi generati da una v.a.).

Per tradizione, usiamo la seguente notazione per indicare *uno o più elementi* $\omega \in \Omega$ che soddisfano la condizione della definizione di una variabile aleatoria (1).

$$\begin{gathered} \{X \in B\} \iff \{\omega \in \Omega : X(\omega) \in B\} \\ \{a < X < b\} \iff \{\omega \in \Omega : X(\omega) \in (a,b)\} \\ \{x < a\} \iff \{\omega \in \Omega : X(\omega) \in (-\infty,a)\} \\ \{X = a\} \iff \{\omega \in \Omega : x(\omega) = a\} \end{gathered}$$

gli eventi $X \in B$ o $a < X < b$, eccetera... si dicono *"eventi generati da X"*.

4. Osservazione sui Spazi Probabilistici Discreti

#Osservazione

Osservazione (le applicazioni dei spazi probabilistici discreti sono sempre v.a. discrete).

Osserviamo che se uno *spazio probabilistico* $(\Omega, \mathcal{A}, p)$ è *discreto* (1), allora una qualsiasi applicazione del tipo $X : \Omega \to \mathbb{R}$ è una *variabile aleatoria*, dato che per la sigma-algebra $\mathcal{A}$ possiamo scegliere la *famiglia di tutti i sottoinsiemi per* Ω. Infatti, abbiamo

$$\{X \in B\} \in \mathcal{A}, \forall B \in \mathcal{B}$$

A3. Definizione di Distribuzione

Definizione di Distribuzione

X

Definizione di distribuzione di una v.a.. ... (?)

X

0. Voci correlate

- Strutture Matematiche della Probabilità
- Spazio di Probabilità Discreto
- Definizione di Variabile Aleatoria

1. Definizione di Distribuzione

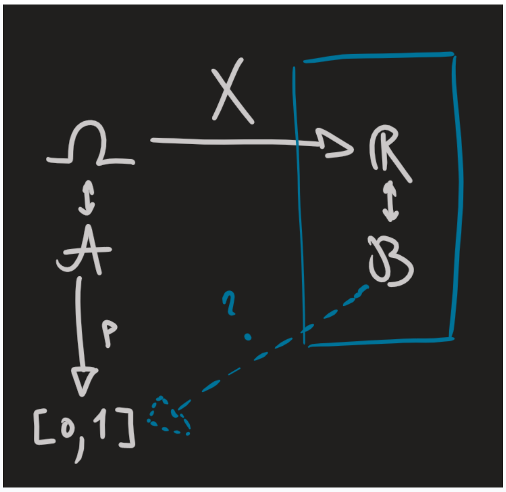
#Osservazione

Osservazione (la situazione con le variabili aleatorie).

Con le *variabili aleatorie* abbiamo creato la situazione nella *figura 1.1.*; ovvero abbiamo una funzione che trasforma $\Omega \to \mathbb{R}$, un'altra funzione che trasforma $\mathcal{A} \to [0,1]$. Inoltre, abbiamo creato uno nuovo spazio che comprende $\mathbb{R}, \mathcal{B}$.

Cosa manca? Esatto, un modo per associare $\mathcal{B} \to [0,1]$. A seguito vedremo un modo per definire tale funzione.

FIGURA 1.1. (*La situazione con le v.a.*)



#Definizione

Definizione (distribuzione su una v.a.).

Data una variabile aleatoria X, definiamo una *distribuzione* (o *legge*) di X come una funzione definita come la composizione

$$p_x = p \circ X^{-1}$$

oppure, usando le notazioni della probabilità, abbiamo

$$p_x(B) = p(\{X \in B\}), \forall B \in \mathcal{B}$$

(*così si vede che la definizione è ben posta*)

A4. Definizione di Indipendenza tra Variabili Aleatorie

Indipendenza tra Variabili Aleatorie

X

Definizione di indipendenza tra variabili aleatorie, composizione di variabili aleatorie indipendenti.

X

0. Voci correlate

- Eventi Indipendenti
- Introduzione ai Vettori Aleatori

1. Definizione di Variabili Aleatorie Indipendenti

IDEA. Da un lato ho la nozione di *indipendenza tra eventi* (1), e dall'altro ho *eventi generati* da variabili aleatorie, denotate da una notazione del tipo $\{X \in E\}$. Allora posso unire le nozioni per *definire* l'idea di *variabili aleatorie indipendenti*.

#Definizione

Definizione (variabili aleatorie indipendenti).

Siano $X_1, \ldots, X_N$ delle *variabili aleatorie* su uno spazio di probabilità $(\Omega, \mathcal{A}, p)$. Questi si dicono *indipendenti* se per ogni scelta degli insiemi *"target"* $E_1, \ldots, E_N \in \mathcal{B}(\mathbb{R})$ si ha la condizione

$$p\{X_1 \in E_1, \ldots, X_N \in E_N\} = \prod_{k=1}^{n} p\{X_k \in E_k\}$$

ovvero *"posso scrivere che la probabilità di avere tutti gli eventi nei loro target equivale al prodotto delle loro probabilità"*.

Usando la nozione di *vettori aleatori*, possiamo pensare all'indipendenza come quella proprietà per cui abbiamo che *gli output marginali* confluiscono in modo indipendente su quello *globale*.

#Osservazione

Osservazione (non è necessario verificare l'altra condizione).

Notiamo che per *eventi* dovevamo anche verificare che *tutti gli eventi* siano indipendenti *due a due*, *tre a tre*, *quattro a quattro*, ed eccetera...
Come mai non lo facciamo qui? Questo perché avendo verificata la condizione

$$p\{X_1 \in E_1, \ldots, X_N \in E_N\} = \prod_{k=1}^{N-1} p\{X_k \in E_k\}$$

abbiamo già gratis

$$p\{X_1 \in E_1, \ldots, X_{N-1} \in E_{N-1}\} = \prod_{k=1}^{N-1} p\{X_k \in E_k\} \cdot \underbrace{p\{X_n \in \mathbb{R}\}}_{1}$$

Questo ci è garantito scegliendo $E_N = \mathbb{R}$ (ovvero l'ultima variabile aleatoria si comporta come vuole): in questo modo si avrebbe $\{X_N \in \mathbb{R}\} = \Omega \implies p(\{X_N \in \mathbb{R}\}) = 1$, che è l'elemento neutro del prodotto. Per induzione si ha che vale su tutte le scelte di eventi.

2. Composizione su Variabili Aleatori Indipendenti

Abbiamo che le *variabili aleatorie* si comportano bene su due aspetti: uno dal punto di vista *vettoriale* (lo vedremo con nozione di *densità marginale e congiunta* (1, 2)) e l'altro dal punto di vista delle *composizioni* (ovvero rielaborazione dei dati). Prima di enunciare la proposizione, facciamo una breve considerazione sulle *composizioni di variabili aleatorie*.

COMPOSIZIONE DI VARIABILI ALEATORIE. Non sempre abbiamo sempre che la *composizione su una variabile aleatoria* rimanga variabile aleatoria. Abbiamo due scelte: o sia lo *spazio di probabilità* discreto (così tanto possiamo prendere la sigma algebra come tutti i sotto insiemi), o sia la *funzione di composizione* abbastanza *"regolare"*.

#Definizione

Definizione (funzione misurabile).

Sia $f:\mathbb{R}\longrightarrow\mathbb{R}$ invertibile. Sia X una variabile aleatoria sullo spazio di probabilità $(\Omega,\mathcal{A},p)$.
Se prendendo la *funzione inversa* di un qualsiasi boreliano su $\mathbb{R}$ si ottiene ancora un boreliano.

$$f^{-1}(B)\in\mathcal{B}(\mathbb{R}),\forall B\in\mathcal{B}(\mathbb{R})$$

Allora f si dice *misurabile*.

Notiamo che una classe notevole di funzioni misurabili sono le continue $\mathcal{C}^0$.

#Proposizione

Proposizione (composizione su v.a.i.).

Siano X_1,X_2 due *variabili aleatorie indipendenti* e $f_1,f_2:\mathbb{R}\longrightarrow\mathbb{R}$ delle funzioni *misurabili*.

Allora le *variabili aleatorie composte* $Y_1:=f_1\circ X_1$ e $Y_2:=f_2\circ X_2$ sono *indipendenti*.

#Dimostrazione

DIMOSTRAZIONE della Proposizione 4 (composizione su v.a.i.)
Supponiamo $E_1,E_2\in\mathcal{B}(\mathbb{R})$ gli eventi target su Y. Allora possiamo scrivere

$$\{Y_1\in E_1,Y_2\in E_2\}=\{f_1\circ X_1\in E_1,f_2\circ X_2\in E_2\}$$

Possiamo applicare l'inversa su $f\circ X$, ottenendo l'evento

$$\{X_1 \in f_1^{-1}(E_1), X_2 \in f_2^{-1}(E_2)\}$$

Adesso per indipendenza delle X, abbiamo

$$p(\{X_1 \in f_1^{-1}(E_1), X_2 \in f_2^{-1}(E_2)\}) = p(X_1 \in f_1^{-1}(E_1)) \cdot p(X_2 \in f_2^{-1}(E_2))$$

e *"annullando l'inversione"* delle f_1, f_2 riotteniamo

$$p(\{Y_1 \in E_1\}) \cdot p(\{Y_2 \in E_2\})$$

da cui si ha la tesi

$$\boxed{p(\{Y_1 \in E_1, Y_2 \in E_2\}) = p(\{Y_1 \in E_1\}) \cdot p(\{Y_2 \in E_2\})}$$

che dimostra la proposizione. ■

Potremo generalizzare questo risultato per *vettori aleatori* (1).

#Corollario

> Corollario (caso vettoriale).
>
> Siano $X_1, \dots, X_N$ delle variabili aleatorie. Ponendo $f(x_1, \dots, x_k) = \sum_{1 \leq j \leq k} x_j$ e $g(x_{k+1}, \dots, x_N) = \sum_{k+1 \leq j \leq N} x_j$ con $k < N$ allora otteniamo che le variabili aleatorie
>
> $$Y_1 = f \circ (X_1, \dots, X_k); Y_2 = g \circ (X_{k+1}, \dots, X_N)$$
>
> sono *variabili aleatorie indipendenti*.

X

B. VARIABILI ALEATORIE DISCRETE: DEFINIZIONI E PROPRIETA'

B1. Definizione di variabile aleatoria discreta

Variabile Aleatoria Discreta

X

Variabili aleatorie discrete: definizione e lemma di caratterizzazione per le v.a. discrete.

X

0. Voci correlate

- Definizione di Variabile Aleatoria
- Spazio di Probabilità Discreto

1. Definizione di Variabile Aleatoria Discreta

#Definizione

Definizione (variabile aleatoria discreta).

Sia $(\Omega, \mathcal{A}, p)$ uno *spazio di probabilità*. Si dice *"variabile aleatoria discreta"* una *variabile aleatoria* (1) che assume al più *una quantità numerabile* di valori, ovvero

$$X(\Omega) = \{x_1, \dots, x_n, \dots\}$$

In particolare prendiamo il caso in cui i valori assunti sono *numeri reali*, ovvero abbiamo un'applicazione del tipo

$$X : \Omega \longrightarrow \mathbb{R}$$

2. Lemma di Caratterizzazione

#Lemma

Lemma (lemma di caratterizzazione delle v.a. discrete).

Siano dati uno spazio di probabilità $(\Omega, \mathcal{A}, p)$ e un'applicazione $X : \Omega \longrightarrow \mathbb{R}$ che assume un numero di valori al più numerabile.

X è una *variabile aleatoria discreta* se e solo se vale $\{X \in x\} \in \mathcal{A}, \forall x \in \mathbb{R}$.
(ovvero *"tutti gli eventi generati da questa variabile aleatoria fanno parte della sigma-algebra"*).

#Dimostrazione

DIMOSTRAZIONE del Lemma 2 (lemma di caratterizzazione delle v.a. discrete)
" $\Longrightarrow$ ": Sia X una *variabile aleatoria discreta* e x un qualsiasi numero della retta reale; allora vale che $\{x\} \in \mathcal{B}$ dato che un singolo punto della retta reale può essere considerata come un *"intervallo degenere"* (infatti $\{x\} = (-\infty, x] \cap [x, +\infty)$). Allora per definizione vale che $\{X = x\} \in \mathcal{A}$.
" $\Longleftarrow$ ": Dato $B \in \mathcal{B}$ qualsiasi, posso scomporre questo *"intervallino"* in una *quantità finita (o al più numerabile) di pezzi*. Possiamo quindi scrivere

$$\{x \in B\} = \bigcup_{x \in B \cap X(\Omega)} \{X = x\}$$

(*nota*: scriviamo $B \cap X(\Omega)$ perché guardiamo *solo* le immagini inverse della X; non ci interessano i valori per cui abbiamo insiemi vuoti)
In particolare questo *appartiene alla sigma-algebra* $\mathcal{A}$, dato che abbiamo un'*unione al più numerabile* di elementi della stessa sigma-algebra: questo prova X essere una *variabile aleatoria* (ed ovviamente è *discreta*). ■

B2. Proprietà delle Densità Associate a V.A. Discrete

Densità associate a Variabili Aleatorie Discrete

X

Definizione di densità associata a v.a. discrete. Proprietà della densità: valori necessari. Osservazione: tramite densità posisamo calcolare le probabilità. Proposizione: condizione necessaria per la densità.

X

0. Voci correlate

- Spazio di Probabilità Discreto
- Variabile Aleatoria Discreta

1. Definizione di Densità

#Definizione

Definizione (densità associata ad una variabile aleatoria discreta).

Sia $(\Omega, \mathcal{A}, p)$ uno *spazio probabilistico discreto*, sia $X : \Omega \longrightarrow \mathbb{R}$ una *variabile aleatoria discreta-reale* (1) e sia una funzione q del tipo

$$q : \mathbb{R} \longrightarrow [0, 1]$$

Chiamiamo q come *densità associata a* X se quest'applicazione viene definita come

$$q(x) := p(\{X = x\}) = p(\{\omega \in \Omega : X(\omega) = x\})$$

2. Proprietà delle Densità

#Proposizione

Proposizione (proprietà della densità).

Sia q una *densità* su X, che assume i valori $X(\Omega) = \{x_1, \dots, x_n, \dots\}$. Allora valgono che:
i. la densità assume sempre valore *nullo* per valori diversi da $x \in X(\Omega)$

$$q(x) = 0, \forall x \notin X(\Omega)$$

ii. la sommatoria di tutte le densità è *unitaria*

$$\sum_{x \in \mathbb{R} \cap X(\Omega)} q(x) = 1$$

#Dimostrazione

DIMOSTRAZIONE della Proposizione 2 (proprietà della densità)
i. Banale, basta usare le definizioni.
ii. Dato che X assume una quantità *al più numerabile* di valori, possiamo passare dalla sommatoria all'*unione* infinita di elementi dello spazio campionario.

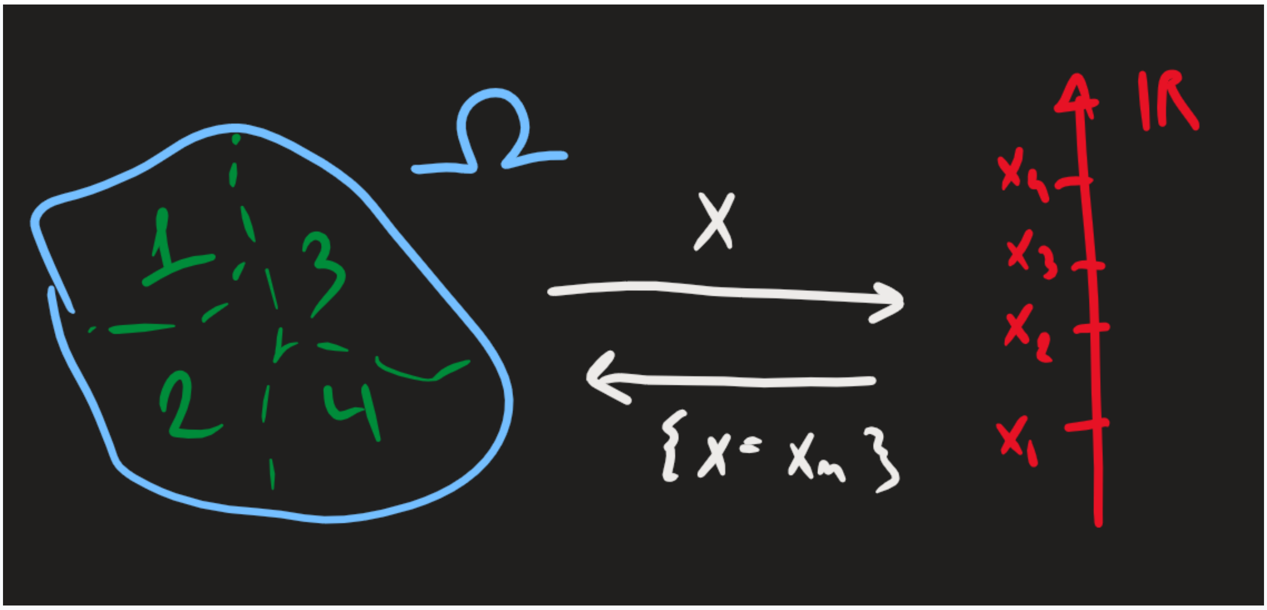
$$\sum_{x \in \mathbb{R}} q(x) = \sum_n q(x_n) = \sum_n p(\{X = x_n\}) = p\left(\bigcup_n \{X = x_n\}\right)$$

Adesso, tenendo conto che gli elementi $\{X = x_n\}$ *formano una partizione* di Ω (per convincerci di questo possiamo rappresentare Ω come uno spazio diviso in pezzi finiti, che poi devono essere biunivocamente associati a elementi x_n), abbiamo che

$$p\left(\bigcup_n \{X = x_n\}\right) = p(\Omega) = 1$$

che è la tesi. ■

FIGURA 2.1. (*Ultimo passaggio*)



#Osservazione

Osservazione (nesso densità-probabilità).

Notiamo che tramite la densità q possiamo *calcolarci* la p_x di un qualsiasi valore, senza dover necessariamente passare a *spazi di probabilità* o conoscere la *variabile aleatoria* stessa. Infatti abbiamo che, dato un qualsiasi intervallo $B \in \mathcal{B}$,

$$p_x(B) := p\{X \in B\} = \sum_{x \in B} p(\{X = x\}) = \sum_{x \in B} q(x)$$

3. Condizione Necessaria delle Densità

#Proposizione

Proposizione (condizione necessaria per le densità).

Sia $q : \mathbb{R} \longrightarrow [0,1]$ una funzione che gode delle seguente proprietà:
i. $q(x) = 0$ tranne per una quantità di x_n *al più numerabile*.
ii. $\sum_{\mathbb{R}} q(x) = 1$ (*nota: si tratta di una somma finita*)

Allora esiste uno *spazio di probabilità* $(\Omega, \mathcal{A}, p)$ e una *variabile aleatoria discreta* $X : \Omega \longrightarrow \mathbb{R}$ avente q come *densità*.

In pratica, avendo una funzione q che soddisfa le proprietà della ipotesi, possiamo *sempre* trattarla come una *densità* per una variabile aleatoria discreta per uno spazio di probabilità.

#Dimostrazione

DIMOSTRAZIONE della Proposizione 4 (condizione necessaria per le densità)
Si tratta di trovare le strutture Ω, $\mathcal{A}$ e p per definire tale spazio di probabilità.

Sia $\Omega := \{x \in \mathbb{R} : q(x) \neq 0\}$ (ovvero stiamo considerando i valori x per cui ho q non nullo). Ovviamente Ω è *al più numerabile*, per il primo punto. Quindi su Ω consideriamo la sigma-algebra $\mathcal{A}$ come la *famiglia di tutti i suoi sottoinsiemi*. Adesso consideriamo la probabilità $p : \mathcal{A} \longrightarrow [0,1]$ definita come

$$p(\{x\}) = q(x), \forall x \in \Omega \implies p(A) = \sum_{x \in A} p(x), \forall A \in \mathcal{A}$$

Per il secondo punto abbiamo che p è effettivamente una *probabilità*. Infatti, $p(\emptyset) = 0$ e $p(\Omega) = 1$.
Infine troviamo la *variabile aleatoria* $X : \Omega \longrightarrow \mathbb{R}$ ponendo

$$X(\omega) = \omega, \forall \omega \in \Omega$$

Dato che abbiamo uno *spazio di probabilità discreto*, segue che X dev'essere necessariamente una *variabile aleatoria* (Osservazione 5 (le applicazioni dei spazi probabilistici discreti sono sempre v.a. discrete)).
Dopodiché per costruzione abbiamo

$$p(\{X = x\}) = p(\{x\}) = q(x)$$

che prova q essere una *densità* su X. ■

X

C. ELENCO DI DENSITA' DISCRETE

C1. Densità Binomiale

Densità Binomiale

X

Definizione di densità binomiale di parametri n e q; caso tipico della densità binomiale (numero di successi di uno schema di prove ripetute). Esempio di densità binomiale.

X

0. Voci correlate

- Densità associate a Variabili Aleatorie Discrete
- Definizione di Schema delle Prove Indipendenti
- Proprietà dello Schema delle Prove Indipendenti

1. Definizione di Densità Binomiale

PREAMBOLO. Adesso iniziamo a vedere *casi specifici* di *densità associate* a v.a. discrete, partendo dalla cosiddetta *"densità binomiale"*.

#Definizione

Definizione (densità binomiale).

Siano $n \in \mathbb{N}$ e $q \in (0,1)$ dei *parametri numerici fissati*. Chiamo la *"densità binomiale di parametri n, q"* la funzione posta come

$$q(x) = \begin{cases} \binom{n}{x} q^x (1-q)^{n-x}, x \in \{0, \dots, n\} \\ 0, x \notin \{0, \dots, n\} \end{cases}$$

e la indichiamo come $B(n, q)$.

#Osservazione

Osservazione (la densità binomiale è una densità).

Si verifica che $B(n, q)$ come definita (1) sopra è una *densità*. Infatti,

$$\sum_{x \in \mathbb{R}} q(x) = \sum_{i=0}^{n} q(i) = \sum_{i=0}^{n} \binom{n}{i} q^i (1-q)^{n-i} = ((1-q)-q)^n = 1^n = 1$$

Parimenti vale che

$$\sum_{x \in \emptyset} q(x) = 0$$

2. Caso Tipico della Densità Binomiale

PREAMBOLO. Adesso vediamo un *caso tipico* di *spazio probabilistico* in cui si ha la *densità binomiale*; ovvero lo *schema delle prove ripetute* (1). In particolare studiamo questo modello in relazione al *numero dei successi ottenuti* (2, punto i.).

#Osservazione

Osservazione (schema delle prove ripetute e il numero dei successi).

Prendiamo lo $(\Omega, \mathcal{A}, p)$ come uno *schema delle prove ripetute*; ovvero abbiamo $\Omega = \{0,1\}^n$ e la probabilità definita come

$$p(\{\omega\}) = q^{\sum \omega_i} (1-q)^{n - \sum \omega_i}, \forall \omega \in \Omega$$

Allora possiamo definirci una *variabile aleatoria discreta* come

$$X(\omega) = \sum_{i=1}^{n} \omega_i$$

In parole, $X(\omega)$ prende la n-upla ω e ritorna il *numero delle volte in cui compare* 1.

Adesso calcolo la densità $q(x) = p(\{X = x\})$. Osserviamo preliminarmente che per $x \notin \{0, 1, \ldots, n\}$ abbiamo $\{X = x\} = \emptyset$ e dunque $q(x) = 0$.

Ora copro invece i casi in cui ho $\{X = x\}$ non vuoto; mi ricordo che per ogni insieme $\{X = x\}$ ho le *combinazioni* di classe x su n oggetti. Allora ho

$$q(x) = p(\{x = X\}) = \sum_{x \in \{X = x\}} q^x (1-q)^{n-x} = \binom{n}{x} q^x (1-q)^{n-x}$$

Notiamo che abbiamo che q è proprio la *densità binomiale* $B(n, q)$.

Facciamo un'ultima osservazione: che vuol dire prendere la distribuzione p_x associata a X? Semplicemente vuol dire che, dato un range di valori $B \subset \mathbb{R}$ stiamo sommando le *probabilità* che il numero delle vittorie stiano dentro in questo range (ovvero le densità). In matematichese, abbiamo

$$p_x(B) = \sum_{x \in B} q(x)$$

3. Esempio di Esercizio con Densità Binomiale

#Esercizio

Esercizio (i bulloni difettosi e sani).

Una fabbrica produce i bulloni in confezioni da tre pezzi ciascuna; ogni bullone prodotto ha una probabilità di essere difettosi con una probabilità del 0.2. Calcolare la probabilità che in una confezione vi sia al più un bullone difettoso.

SVOLGIMENTO. Basta prendere $B(3, 0.8)$, dato che siamo in uno *schema di prove ripetute*, se consideriamo la *"vittoria"* come *"avere un bullone sano"*: in tal caso basta prendere $B = \{2, 3\} \subset \mathbb{R}$ e calcolare $p_x(B)$, ovvero

$$p_x(B) = q(2) + q(3) = \ldots = 0.896$$

Che è il nostro risultato.

C2. Densità ipergeometrica

Densità Ipergeometrica

X

Definizione di densità ipergeometrica. Modello tipico di spazio probabilistico avente densità ipergeometrica.

X

0. Voci correlate

- Densità associate a Variabili Aleatorie Discrete

1. Definizione di Densità Ipergeometrica

Continuiamo con la nostra carrellata delle *densità su v.a. discrete*.

#Definizione

Definizione (densità ipergeometrica).

Siano $a, b, n \in \mathbb{N}^+$ dei *parametri strettamente positivi*, tali che $n \leq a+b$. Ponendo i parametri secondari come

$$k_1 = \max\{n-b, 0\}; k_2 = \min\{a, n\}$$

(si nota che ovviamente segue che $k_1 \leq k_2$) allora si definisce la *densità ipergeometrica di parametri* a, b, n come la funzione definita così

$$q(x) := \begin{cases} \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}, x : (k_1 \leq x \leq k_2) \cap \mathbb{N} \\ 0, \text{altrimenti} \end{cases}$$

Lo si indica con $\text{Iper}(n, a, b)$.

Notiamo che con x che *"appartiene all'iterazione da k_1 a k_2"* segue che

$$k_1 \leq x \leq k_2 \implies \begin{cases} x \geq 0 \land n-x \leq b \\ x \leq a \land n-x \geq 0 \end{cases}$$

(per convincerci di questo consideriamo i casi di k_1, k_2 caso per caso)
quindi abbiamo che i *coefficienti* dei *binomi di Newton* sono ben posti.

#Proposizione

Proposizione (la densità ipergeometrica è una densità).

Si verifica che vale, per una qualsiasi scelta della tripla n, a, b valida, si ha

$$\sum_{k_1 \leq x \leq k_2} q(x) = 1$$

2. Modello Tipico

Per capire il *vero significato* di questa densità, proponiamo il seguente *"modello tipico"* della densità ipergeometrica.

#Esempio

MODELLO. (*Urna senza reimmissione*)

Spazio. Supponiamo di avere un'urna con N palline di due colori: M *bianche* e restante ($N-M$) *nere*. Supponiamo di eseguire n estrazioni senza reimmissione.
Qui notiamo subito che vogliamo porre $n \leq (N-M)+M = N$ per avere una *situazione ben posta*. Questa è uguale alla condizione $n \leq a+b$.
Conoscendo già lo spazio Ω da considerare sono le *combinazioni* di n su N oggetti (1).
Quindi, assumendo l'*equiprobabilità*, abbiamo

$$\forall \omega \in \Omega, p(\omega) = \frac{1}{|\Omega|} = \frac{1}{\binom{N}{n}}$$

Variabili Aleatorie. Adesso iniziamo a considerare le *variabili aleatorie.* Introduciamo la v.a. X come il *"numero di palline bianche estratte"*. Adesso osserviamo che il *codominio* di X ha le seguenti proprietà, prendendo $x = X(\omega)$ qualsiasi. Vediamo di *"interpretare"* i valori derivati k dati nella definizione.
Prima di tutto, dev'essere *minore* di $k_2 := \min\{n, M\}$; sono *"limitato"* dal numero di tentativi disponibili. Notiamo subito che da questo discende $n - x \leq \min\{n, N-M\}$ (stiamo infatti considerando la situazione complementare).
Dopodiché consideriamo il *"complementare"* della maggiorazione appena ottenuta: abbiamo infatti

$$x \leq \min\{0, N-M-n\} \implies x \geq \max\{0, n-(N-M)\}$$

ovvero abbiamo che il *"peggior risultato"* è quello di avere 0 palline, oppure il *minimo numero di palline bianche ottenibili* (ad un certo punto esauriscono le palline nere e saremmo *"costretti"* a pescare quelle bianche).
Derivazione della Densità Ipergeometrica. A questo punto vogliamo calcolare la densità associata ad X. L'evento $\{X = k\}$ corrisponde ad *"aver pescato k palline bianche su una scelta di n palline su N"*, di cui sappiamo la cardinalità (2).

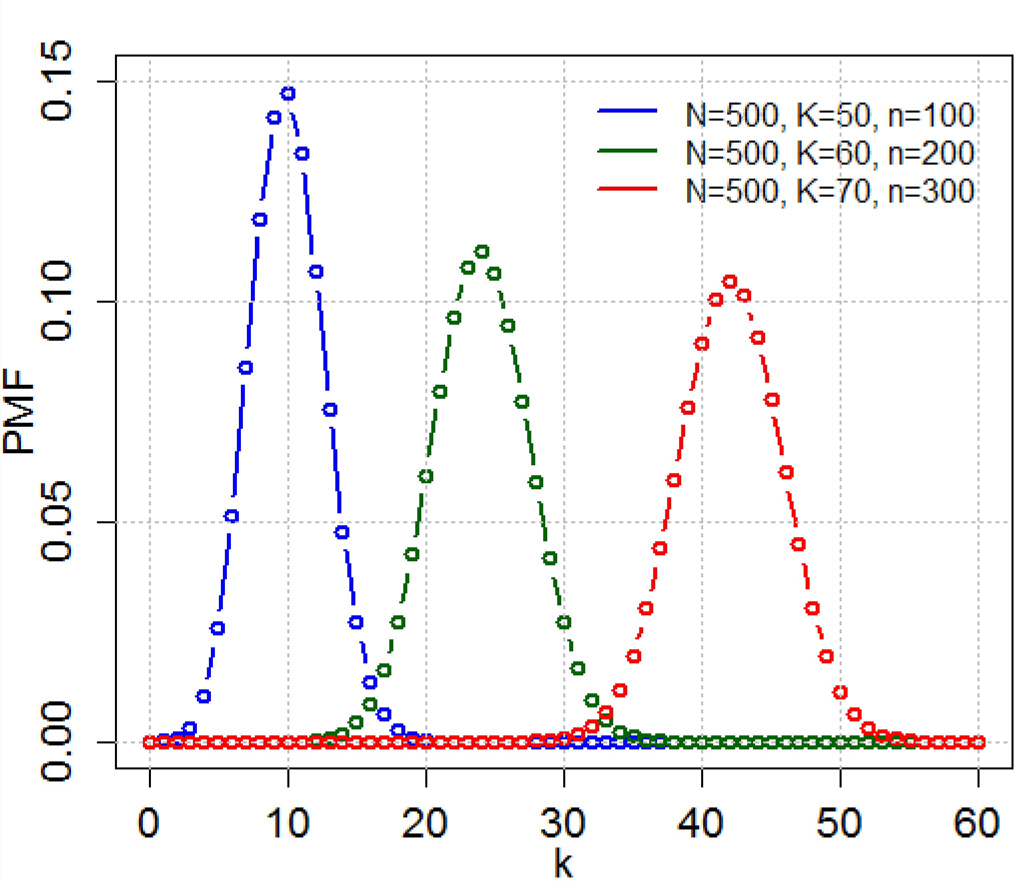
$$|\{X = k\}| = \binom{M}{k}\binom{N-M}{n-k}$$

(possiamo interpretarla come "k *palline devono essere bianche, le restanti* $n-k$ *nere*") da cui abbiamo

$$p(\{X=k\}) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

che è proprio la tesi voluta, considerando le limitazioni su $X(\omega)$. Da qui in poi possiamo generalizzare tutto come $M=a$, $N-M=b$ e $N=n$. ■

FIGURA. (*Densità ipergeometrica*)



C3. Densità Geometrica

Densità Geometrica

X

Definizione di densità geometrica avente un parametro non-estremo. Modello tipico di densità geometrica. L'assenza di memoria delle v.a. associate a densità geometrica.

X

0. Voci correlate

- Densità associate a Variabili Aleatorie Discrete
- Definizione di Schema delle Prove Indipendenti
- Proprietà dello Schema delle Prove Indipendenti

1. Osservazione sullo Schema di Bernoulli

Adesso introduciamo un'altra *densità* importante: la densità geometrica. Prima di capirci qualcosa, si effettua la seguente osservazione preliminare.

#Osservazione

OSSERVAZIONE. (*Sullo schema di Bernoulli*)
MODELLO. Prendiamo $(\Omega, p, \mathcal{A})$ come lo *schema delle prove indipendenti* (1). Dallo studio di questa sappiamo i seguenti risultati (2).
i. *formula per calcolare il primo successo al j-esimo esperimento*

$$P(C_j) = (1-q)^{j-1}q$$

ii. *formula per calcolare k successi*

$$p(C_k) = \binom{n}{k} q^k (1-q)^{n-k}$$

Adesso chiamo C_j come B_j e C_k come A_k.
VARIABILI ALEATORIE. Adesso prendiamo X come la *variabile aleatoria* che ci dice quanti successi abbiamo in un set di ω prove. Si verifica facilmente che

$$p(X=k) = p(C_k)$$

Ora prendendo Y come la *variabile aleatoria* che ci dice il *numero del tentativo in cui abbiamo il primo successo su n tentativi*, abbiamo

$$Y_n(\omega) := \begin{cases} \min\{j \in \mathbb{N} : \omega_j = 1\}, \forall \omega \neq (0, \dots, 0) \\ +\infty, \omega = (0, \dots, 0) \end{cases}$$

e poi

$$p(Y_n = j) = B_j, \forall j \in \{1, \dots, n\}$$

DENSITA' ASSOCIATE. Adesso troviamo le densità associate su X e Y.
La prima è banale, dal momento che basta passare alla definizione: abbiamo

$$q_X(k) := p\{X=k\} = p(A_k) = \binom{n}{k} q^k (1-q)^{n-k}$$

Dopodiché per la densità Y_n abbiamo

$$q_{Y_n}(x) = p(\{Y_n = x\}) = \begin{cases} (1-q)^{x-1}q, x = 1, \ldots, n \\ (1-q)^n, x = +\infty \\ 0, \text{altrimenti} \end{cases}$$

Notiamo che $q_{Y_n}(x) = q_{Y_m}(x)$ per $x \leq \min\{n, m\}$. Ovvero considerando un qualsiasi x che sia minore del minimo, ho che la probabilità di avere il successo in x sarà sempre la stessa: se effettuo $5, 6, 7$ o più lanci, in ogni caso avrò la stessa probabilità di avere successo al primo, secondo, terzo e fino al quinto lancio.

ULTIMA VARIABILE ALEATORIA. Essendo pronti a generalizzare, adesso cambiamo qualcosa sulla variabile aleatoria Y_n. Supponiamo di *non sapere* il numero di tentativi n a priori: cosa cambia? In questo caso possiamo considerare la variabile aleatoria Y come $\lim_n Y_n$ (ovvero l'estensione di Y_n su tutti gli interi $\mathbb{N}$). In questo caso abbiamo

$$q_Y(x) = \begin{cases} (1-q)^{x-1}q, x \in \mathbb{N} \\ 0, \text{altrimenti o } x = +\infty \end{cases}$$

Questo ha perfettamente senso, dato che

$$\lim_n q_{Y_n}(+\infty) = \lim_n (1-q)^n = 0$$

Adesso siamo pronti per definire tutto. ■

2. Definizione di Densità Geometrica

Possiamo definire formalmente la nozione di *densità geometrica*.

#Definizione

Definizione (densità geometrica).

Sia $q \in (0, 1)$ un parametro. Si definisce la *"densità geometrica di parametro q"* come la funzione

$$q(x) = \begin{cases} (1-q)^{x-1}q, \forall x \in \{1, 2, 3, \ldots\} \\ 0, \text{altrimenti} \end{cases}$$

e la si chiama come $\text{Geo}(q)$.

#Proposizione

Proposizione (la densità geometrica è una densità).

Mediante le *nozioni sulle serie*, si verifica che abbiamo convergenza in

$$\sum_{x=1}^{+\infty} q(x) = 1$$

Infatti,

$$\sum_{x=1}^{+\infty} q(x) = q \sum_{y=x-1=0}^{+\infty} (1-q)^y = 1$$

per quanto visto sulle *serie geometriche* (1)

3. Proprietà della Densità Geometrica

Vediamo una proprietà importante della *densità geometrica*.

#Proposizione

Proposizione (l'assenza di memoria).

Sia Y una v.a. con densità geometrica $\text{Geo}\,(q)$. Allora

$$\boxed{p\{Y > n+m \mid Y > n\} = p\{Y > m\}}$$

ovvero, avendo fatto n tentativi e volendo ottenere il successo al $n+m$-esimo tentativo, *"me ne dimentico"* di tutte le n prove appena svolte.

Inoltre vale che le probabilità sono calcolate come

$$\boxed{\begin{gathered} p(\{Y > n\}) = (1-q)^n \\ p(\{Y \leq n\}) = 1-(1-q)^n \end{gathered}}$$

#Dimostrazione

DIMOSTRAZIONE della Proposizione 3 (l'assenza di memoria)

Osserviamo che abbiamo l'inclusione

$$\{Y > n+m\} \subset \{Y > n\}$$

Dunque segue che

$$\{Y > n+m\} \cap \{Y > n\} = \{Y > n+m\}$$

(per convincerci di questo fare un diagramma di Eulero-Venn)
quindi, prendendo la *definizione di probabilità condizionale* si ha

$$p\{Y \geq n+m \mid Y > n\} := \frac{p\overbrace{\{Y > n+m \cap Y > n\}}^{\{Y>n+m\}}}{p(Y > n)}$$

calcolando esplicitamente i termini abbiamo

$$p\{Y > n\} = \sum_{x=n+1}^{+\infty} q(x) = \sum_{x=n+1}^{+\infty} q(1-q)^{x-1}$$

con $x = j + n \implies j = x - n \implies j = (n+1) - (n) = 1, x - 1 = (j-1) + n$ abbiamo

$$\ldots = \sum_{j=1}^{+\infty} \underbrace{q(1-q)^{j-1}}_{1}(1-q)^n = (1-q)^n$$

Similmente si ha

$$p\{Y > n+m\} = (1-q)^{n+m}$$

quindi abbiamo

$$p\{Y \geq n+m \mid Y > n\} = \frac{p(Y > m+n)}{p(Y > n)} = \frac{(1-q)^{n+m}}{(1-q)^n} = (1-q)^n$$

che è la definizione di

$$p(\{Y > n\}) := (1-q)^n$$

il che prova la tesi. Per calcolare $p(\{Y \leq n\})$ basta prendere il complementare. ■

#Osservazione

> Osservazione (forma debole della densità geometrica).
>
> Notando che stiamo sempre agendo su $\mathbb{N}$. Possiamo quindi porre $m' = m - 1$ nella proposizione sull'assenza di memoria e avremo
>
> $$p(\{Y \geq n+m|Y > n\}) = p(\{Y \geq m\})$$
>
> che di dà la *"forma debole"* della proposizione appena enunciata.

#Corollario

> Corollario (assenza di memoria, caso uguaglianza).
>
> Sia Y una v.a. con densità geometrica $\mathrm{Geo}\,(q)$. Allora

$$p(\{Y=n+m|Y>n\}) = p(\{Y=m\})$$

#Dimostrazione

DIMOSTRAZIONE del Corollario 5 (assenza di memoria, caso uguaglianza)
Basta pensare l'insieme a sinistra come

$$\{Y=n+m|Y>n\} = \{Y \geq n+m|Y>n\} \setminus \{Y>n+m|Y>n\}$$

e usare le regole appena osservate (Corollario 5 (assenza di memoria, caso uguaglianza);Osservazione 4 (forma debole della densità geometrica)) si ricava che

$$\begin{aligned} p(\underbrace{\{Y \geq n+m|Y>n\}}_{\{Y\geq m\}} \setminus \underbrace{\{Y>n+m|Y>n\}}_{\{Y>m\}}) &= p(\{Y\geq m\}) - p(\{Y>m\}) \\ &= p(\{Y\geq m\}\setminus\{Y>m\}) \\ &= \boxed{p(\{Y=m\})} \end{aligned}$$

che è la tesi. ■

Conclusione. Il corollario di questo teorema ci dice che *attendere il primo successo per altre n prove* è completamente equivalente ad *attendere il primo successo per nessuna prove effettuata*; ovvero le precedenti prove si comportano come se non esistessero. Questo sfata la *fallacia* (nota come la *"Gambler's Fallacy"*, wikipedia), per cui uno pensa che facendo infinite prove su qualcosa, si ha che la probabilità di avere almeno un successo aumenta all'aumentare di prove effettuate (se sono sempre fallimentari). Certo, per la *scimmia di Borel* potrebbe sembrar vera (1): nella visione globale ho la probabilità di avere almeno un successo aumenta: tuttavia, questo è completamente diverso dal contare il *"primo successo"*! Un conto è *contare la probabilità di avere almeno un successo*, un'altra storia è *contare il primo successo*.
Per spiegare in termini di logica, la *scimmia di Borel* implica di avere un *numero* di successi. Tuttavia, non viene detto nulla su *dove vengono disposti* questi successi.

4. Esercizi sulla Densità Geometrica

#Esercizio

Esercizio (l'urna).

Da un'urna con 10 palline bianche e 15 nere si eseguono estrazioni con reinserimento fino all'estrazione di una pallina nera. Calcolare la probabilità che servano almeno 10 estrazioni per ottenere la pallina nera.

SVOLGIMENTO.
Si ha $q=0.6$. Prendiamo la densità $\operatorname{Geo} 0.6$ e la associamo alla variabile aleatoria Y. La consegna ci sta chiedendo di calcolare

$$p(\{Y \geq 10\})$$

Qui basta considerare il suo complementare e utilizzare il teorema sull'assenza di memoria della densità geometrica. Ovvero

$$p(\{Y \geq 10\}) = 1 - p(\{Y < 10\}) = 1 - p(\{Y \leq 9\}) = 1 - (1 - (1-q)^9)$$

Che ci dà il risultato finale

$$0.0262144\%$$

■

C4. Densità di Poisson

Densità di Poisson

X

Densità di Poisson. Definizione, motivazione e significato. Esercizio.

X

0. Voci correlate

- Densità associate a Variabili Aleatorie Discrete
- Densità Binomiale

1. Definizione di Densità di Poisson

Adesso vediamo una densità importante per *approssimare* le *densità binomiali*.

#Definizione

Definizione (densità di Poisson).

Sia $\lambda > 0$ un *parametro reale* (ovviamente fissato). Chiamo la *densità di Poisson di parametro* λ come la funzione

$$q(x) = \begin{cases} \dfrac{\lambda^x}{x!}e^{-\lambda}, x \in \mathbb{N} \\ 0, \text{altrimenti} \end{cases}$$

Viene indicato con p_λ.

#Teorema

Teorema (la densità di Poisson è una densità).

Si ha che p_λ è sempre una *densità*. Infatti dallo sviluppo di Taylor abbiamo (1)

$$\sum_{x=0}^{+\infty} \frac{\lambda^x}{x!} = e^\lambda$$

Dunque

$$\sum_{x=0}^{+\infty} q(x) = e^{\lambda - \lambda} = 1$$

2. Derivazione della Densità di Poisson

Vogliamo capire da *dove viene* questa densità. L'idea principale è quella di ottenere il *limite delle densità binomiali* $B\left(n, \frac{\lambda}{n}\right)$ su n (1). Adesso osserviamo bene.

#Osservazione

Osservazione (il limite delle densità binomiali).

Si ha la *famiglia delle densità binomiali* posta come

$$(\mathcal{B}_n)_n = B\left(n, \frac{\lambda}{n}\right) \to q_n$$

Sia fissato adesso un $k \in \mathbb{N}$ qualunque, vogliamo calcolare il valore $q_n(k)$.
Per definizione si ha

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

(1). Questo vale per $k \in \{0, \dots, n\}$. Calcolando esplicitamente i coefficienti binomiali, abbiamo

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

Dato che ho k fissato, posso scegliere uno spettro di valori n tale che $n \geq k$. Così risulta

$$
\begin{aligned}
q_n(k) &= \frac{n!}{k!(n-k)!}\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k} \\
&= \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}\frac{n(n-1)\ldots(n-k+1)}{\underbrace{n(n)\ldots(n)}_{k\text{ fattori}}} \\
&= \frac{\lambda^k}{k!}e^{-\lambda}
\end{aligned}
$$

Considerando i limiti

$$
\begin{gathered}
\lim_n \left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda} \\
\lim_n \left(1-\frac{\lambda}{n}\right)^k = 1 \\
\lim_n \frac{n}{n}\cdot\frac{n-1}{n}\cdot\ldots\cdot\frac{n-k+1}{n} = 1\cdot\ldots\cdot 1 = 1
\end{gathered}
$$

In conclusione si enuncierà il seguente teorema per riassumere tutto.

#Teorema

Teorema (proprietà fondamentale della densità di Poisson).

Sia $(\mathcal{B}_n)_n$ la famiglia di densità posta come

$$
(\mathcal{B}_n)_n \sim B\left(n, \frac{\lambda}{n}\right) = q_n
$$

Allora vale che, per un qualsiasi $k \in \mathbb{N}$ fissato,

$$
\boxed{\lim_n \mathcal{B}_n(k) = \frac{\lambda^k}{k!}e^{\lambda} = p_\lambda(k)}
$$

Ovvero la *"famiglia delle densità binomiali converge verso la densità di Poisson"*.

3. Significato pratico

Osserviamo il *significato* della densità di Poisson in *termini pratici*.

#Osservazione

Osservazione (densità di Poisson in termini pratici).

> Prendiamo uno *schema di prove ripetute* con $q_n = \frac{\lambda}{n}$; ovvero la *probabilità* q diminuisce con singola prova. Va notate che il risultato appena provato prima (1) vale anche per un q_n tale che
>
> $$\lim_n nq_n = \lambda$$
>
> Ovvero abbiamo $q_n \sim \frac{1}{n}$.
>
> Posto $\lambda_n = nq_n$, abbiamo per la *continuità* che i limiti si trasmettono (quindi $\lim_n \lambda_n = \lim_n nq_n = \lambda$. Pertanto,
>
> $$\lim_n p_{\lambda_n}(k) = p_\lambda(k)$$
>
> Ovvero c'è uno scarto tra il *"limite approssimato"* e il *"limite target"* di
>
> $$\lim_n |B(n, q_n)(k) - p_{\lambda_n}(k)| = 0$$
>
> Quindi per approssimare $B(n, q_n)$, che potrebbe risultare ostico da calcolare con i coefficienti binomiali, possiamo semplicemente usare p_λ che contiene solo dei fattoriali ed esponenti. Empiricamente si mostra che minimalmente il numero ideale n è $n = 50$

4. Esercizio sulla densità di Poisson

#Esercizio

> Esercizio (gli errori in una pagina).
>
> Ci sono mediamente 1,2 errori di battitura per pagina di un libro. Calcolare la probabilità di trovare in una data di 2000 pagine zero errori e almeno tre errori.

SVOLGIMENTO.
Qui si ha uno *schema di prove ripetute* con $q = 1.2/2000$ che rappresenta la probabilità di avere un carattere erroneo. Volendo possiamo prendere la densità $B(2000, q)$. Tuttavia, questo diventa impossibile da calcolare per una mente umana. Prendiamo dunque p_λ con $\lambda = nq = 1.2$. Adesso diventa tutto facile da calcolare: basta associare X la densità p_λ e calcolare le probabilità

$$p\{X = 0\} = e^{-\lambda} \sim 0.3$$
$$p\{X \geq 3\} = 1 - p\{x < 3\} = 1 - \sum_{j=0}^{2} p\{X = j\} = \ldots \sim 0.12$$

che finisce l'esercizio. ■