Probabilità e Processi Stocastici (455AA)

Lezione 4

Dario Trevisan

3/10/2024

Section 1

Problemi (dalla lezione precedente)

Problema 5

Un'urna da cui un amico effettua estrazioni contiene inizialmente 40 palline di cui 10 rosse e 30 blu. Un amico ci informa che ha estratto la sequenza, nell'ordine

rosso, rosso, blu, rosso.

Sappiamo inoltre sicuramente che l'amico ha effettuato tutte estrazioni o con rimpiazzo o senza rimpiazzo, ma non sappiamo esattamente come (supponiamo a priori probabilità uniforme).

1. Dire se è più probabile che le estrazioni siano state effettuate con rimpiazzo o senza.

Problema 5

Un'urna da cui un amico effettua estrazioni contiene inizialmente 40 palline di cui 10 rosse e 30 blu. Un amico ci informa che ha estratto la sequenza, nell'ordine

rosso, rosso, blu, rosso.

Sappiamo inoltre sicuramente che l'amico ha effettuato tutte estrazioni o con rimpiazzo o senza rimpiazzo, ma non sappiamo esattamente come (supponiamo a priori probabilità uniforme).

1. Dire se è più probabile che le estrazioni siano state effettuate con rimpiazzo o senza.
2. Calcolare la probabilità che la pallina successiva estratta dall'amico sia rossa.

$E_R =$ "estrazioni con rimpiazzo"

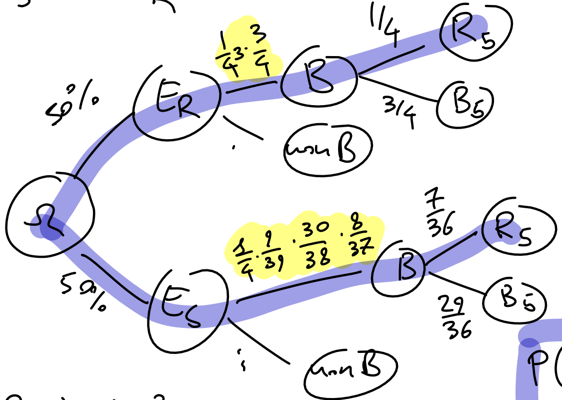
$B =$ "R_1, R_2, B_3, R_4"

$E_S =$ non E_R

$N = 40$

$R = 10$

$B = 30$



$$P(E_R \mid B) = \frac{P(E_R)\cdot \overline{P(B \mid E_R)}}{P(B)}$$

$$P(E_S \mid B) = \frac{P(E_S)\,\overline{P(B \mid E_S)}}{P(B)}$$

$$P(B) = \frac{1}{2}\cdot\frac{3}{4^4} + \frac{1}{2}\cdot\frac{1}{4}\cdot\frac{9}{39}\cdot\frac{30}{38}\cdot\frac{8}{37}$$

$$\boxed{P(R_5 \mid B) = \frac{P(R_5, B \mid \Omega)}{P(B \mid \Omega)}}$$

Section 2

Variabili aleatorie generali

Contenuti

- le variabili aleatorie dal punto di vista elementare (accenniamo all'assiomatizzazione di Kolmogorov)

Contenuti

- le variabili aleatorie dal punto di vista elementare (accenniamo all'assiomatizzazione di Kolmogorov)
- il concetto di legge (o distribuzione) nei due casi più rilevanti (densità discreta e continua)

Contenuti

- le variabili aleatorie dal punto di vista elementare (accenniamo all'assiomatizzazione di Kolmogorov)
- il concetto di legge (o distribuzione) nei due casi più rilevanti (densità discreta e continua)
- composizione di variabili tramite una funzione e trasformazione della densità.

Contenuti

- le variabili aleatorie dal punto di vista elementare (accenniamo all'assiomatizzazione di Kolmogorov)
- il concetto di legge (o distribuzione) nei due casi più rilevanti (densità discreta e continua)
- composizione di variabili tramite una funzione e trasformazione della densità.
- variabile aleatoria congiunta e marginali

Contenuti

- le variabili aleatorie dal punto di vista elementare (accenniamo all'assiomatizzazione di Kolmogorov)
- il concetto di legge (o distribuzione) nei due casi più rilevanti (densità discreta e continua)
- composizione di variabili tramite una funzione e trasformazione della densità.
- variabile aleatoria congiunta e marginali
- formula di Bayes nel linguaggio delle variabili aleatorie

Contenuti

- le variabili aleatorie dal punto di vista elementare (accenniamo all'assiomatizzazione di Kolmogorov)
- il concetto di legge (o distribuzione) nei due casi più rilevanti (densità discreta e continua)
- composizione di variabili tramite una funzione e trasformazione della densità.
- variabile aleatoria congiunta e marginali
- formula di Bayes nel linguaggio delle variabili aleatorie
- indipendenza tra variabili aleatorie

Contenuti

- le variabili aleatorie dal punto di vista elementare (accenniamo all'assiomatizzazione di Kolmogorov)
- il concetto di legge (o distribuzione) nei due casi più rilevanti (densità discreta e continua)
- composizione di variabili tramite una funzione e trasformazione della densità.
- variabile aleatoria congiunta e marginali
- formula di Bayes nel linguaggio delle variabili aleatorie
- indipendenza tra variabili aleatorie
- metodo "grafico" delle reti bayesiane per rappresentare le dipendenze (o l'indipendenza)

Sistemi di alternative e variabili

Una *variabile aleatoria* può essere pensata come una *notazione efficace* per un *sistema di alternative* associato ad una grandezza su cui vi sia incertezza.

- In molti problemi, è richiesto il grado di fiducia che una grandezza X (quantitativa ossia numerica o anche qualitative, come ad esempio colori o sequenze) assuma un determinato valore.

Sistemi di alternative e variabili

Una *variabile aleatoria* può essere pensata come una *notazione efficace* per un *sistema di alternative* associato ad una grandezza su cui vi sia incertezza.

- In molti problemi, è richiesto il grado di fiducia che una grandezza X (quantitativa ossia numerica o anche qualitative, come ad esempio colori o sequenze) assuma un determinato valore.
- Anche se vi è incertezza, si sa che assume uno e un solo valore $x \in E$.

Sistemi di alternative e variabili

Una *variabile aleatoria* può essere pensata come una *notazione efficace* per un *sistema di alternative* associato ad una grandezza su cui vi sia incertezza.

- In molti problemi, è richiesto il grado di fiducia che una grandezza X (quantitativa ossia numerica o anche qualitative, come ad esempio colori o sequenze) assuma un determinato valore.
- Anche se vi è incertezza, si sa che assume uno e un solo valore $x \in E$.
- Possiamo introdurre una alternativa $A_x =$ "X assume il valore x". $(A_x)_{x\in E}$ è un sistema di alternative.

$P(A_x \text{ e } A_{x'}) = 0$ se $x \neq x'$, $P\left(\bigcup_{x\in E} A_x\right) = 1$

Diciamo che X *è una variabile aleatoria a valori in* E, e usiamo la notazione

$$\{X = x\} = A_x.$$

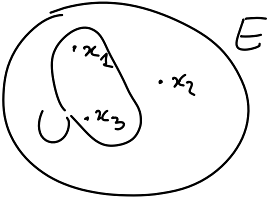
- La variabile X è identificata con il sistema di alternative $(\{X = x\})_{x \in E}$.

Diciamo che X *è una variabile aleatoria a valori in* E, e usiamo la notazione

$$\{X = x\} = A_x.$$

- La variabile X è identificata con il sistema di alternative $(\{X = x\})_{x \in E}$.
- La grandezza X si comporta come una *variabile* matematica. Dato un sottoinsieme di valori $U \subseteq E$, possiamo scrivere

$$\{X \in U\} = \bigcup_{x \in U} \{X = x\}$$



$$\{X \in \{x_1, x_2\}\} = \{X = x_1\} \cup \{X = x_2\}$$

Diciamo che X *è una variabile aleatoria a valori in* E, e usiamo la notazione

$$\{X = x\} = A_x.$$

- La variabile X è identificata con il sistema di alternative $(\{X = x\})_{x \in E}$.
- La grandezza X si comporta come una *variabile* matematica. Dato un sottoinsieme di valori $U \subseteq E$, possiamo scrivere

$$\{X \in U\},$$

- Una scrittura compatta è $X \in E$ oppure, seguendo l'assiomatizzazione di Kolmogorov, $X : \Omega \to E$ (questa notazione sarà chiarita tra poco).

Il vantaggio di disporre di una variabile X è che possiamo effettuare determinate operazioni naturali, che corrispondono in pratica ad operazioni, magari meno evidenti, sul sistema di alternative.

- Ad esempio, dato $U \subseteq E$, possiamo scrivere

$$\{X \in U\} = \text{"}X \text{ assume un qualsiasi valore tra quelli di } U\text{"}.$$

Il vantaggio di disporre di una variabile X è che possiamo effettuare determinate operazioni naturali, che corrispondono in pratica ad operazioni, magari meno evidenti, sul sistema di alternative.

- Ad esempio, dato $U \subseteq E$, possiamo scrivere

$$\{X \in U\} = \text{"}X \text{ assume un qualsiasi valore tra quelli di } U\text{"}.$$

- Più formalmente è la disgiunzione inclusiva (o unione tra insiemi)

$$\{X \in U\} = \bigcup_{x \in U} \{X = x\}$$

.

Esempi $\quad E = \mathbb{R} \quad U = (a, b) \quad a < b$

$$\{X \in U\} = \{X \in (a, b)\} = \{a < X < b\}$$

$$U = (-\infty, t]$$

$$\{X \in (-\infty, t]\} = \{X \leq t\}$$

$$\boxed{\{X^2 = 1\}} = \{X = -1\} \text{ o } \{X = 1\}$$

$$= \{X \in \{-1, 1\}\}$$

Variabili aleatorie secondo Kolmogorov

Dato $(\Omega, \mathcal{A}, P)$ una variabile aleatoria X, a valori in E è una funzione

$$X : \Omega \to E$$

tale l'immagine inversa

$$\{X=x\} = X^{-1}(x) = \{\omega \in \Omega : X(\omega) = x\}$$

sia un evento in $\mathcal{A}$ per ogni $x \in E$.

- La teoria è un po' più complicata in realtà, per trattare E infiniti.

Legge (o distribuzione) di una variabile

Data X a valori in E, si vuole spesso determinare

$$P(X = x|I)$$

I inf. nota

o più in generale

$$P(X \in U|I).$$

Questo si traduce nel concetto di *legge* (o distribuzione), ossia la funzione

$$E \supseteq U \mapsto P(X \in U|I) \in [0,1]$$

- **Osservazione**: la legge dipende dall'informazione nota I.

Legge (o distribuzione) di una variabile

Data X a valori in E, si vuole spesso determinare

$$\boxed{P(X = x|I)}$$

o più in generale

$$P(X \in U|I).$$

Questo si traduce nel concetto di *legge* (o distribuzione), ossia la funzione

$$E \supseteq U \mapsto P(X \in U|I).$$

- **Osservazione**: la legge dipende dall'informazione nota I.
- **Problema**: come determinare la legge di X?

Se E ha k elementi i sottoinsiemi U sono $\boxed{2^k}$

Legge (o distribuzione) di una variabile

Data X a valori in E, si vuole spesso determinare

$$P(X = x|I)$$

o più in generale

$$P(X \in U|I).$$

Questo si traduce nel concetto di *legge* (o distribuzione), ossia la funzione

$$E \supseteq U \mapsto P(X \in U|I).$$

- **Osservazione**: la legge dipende dall'informazione nota I.
- **Problema**: come determinare la legge di X?
- **Soluzione**: densità discreta (o continua), definita su $x \in E$ (non su $U \subseteq E$).

Densità discreta

Ad ogni sistema di alternative (finito) $(A_i)_{i=1}^n$ è associata una densità discreta $P(A_i|I)$.

$\mathbb{N}, \mathbb{Z}$

Per variabili aleatorie X a valori in E finito (o infinito ma discreto), si definisce *densità discreta* (rispetto ad I) la funzione

$$E \ni x \mapsto P(X = x|I).$$

- Assume valori in $[0, 1]$ e

$$\boxed{\sum_{x \in E} P(X = x|I) = 1,}$$

Densità discreta

Ad ogni sistema di alternative (finito) $(A_i)_{i=1}^n$ è associata una densità discreta $P(A_i|I)$.

Per variabili aleatorie X a valori in E finito (o infinito ma discreto), si definisce *densità discreta* (rispetto ad I) la funzione

$$E \ni x \mapsto P(X = x|I).$$

- Assume valori in $[0, 1]$ e
$$\sum_{x \in E} P(X = x|I) = 1,$$
- Utile la notazione $P(X = x|I) \propto f(x)$ per indicare costanti moltiplicative non espresse: si ha esplicitamente
$$P(X = x|I) = f(x) / \sum_{y \in E} f(y).$$

Esempi

X a valori in $\{0,1,2,\dots,n\} = E$

X densità uniforme ossia $P(X=x) \propto 1$

$$P(X=x) = \frac{1}{n+1}$$

NON è possibile definire una densità uniforme su $E = \mathbb{N}$

$P(X=x) = c \quad \forall x \in \mathbb{N}$

$$\frac{1}{\sum_{y \in \mathbb{N}} c} = \frac{1}{\infty} = 0$$

Es Dire se esiste una densità

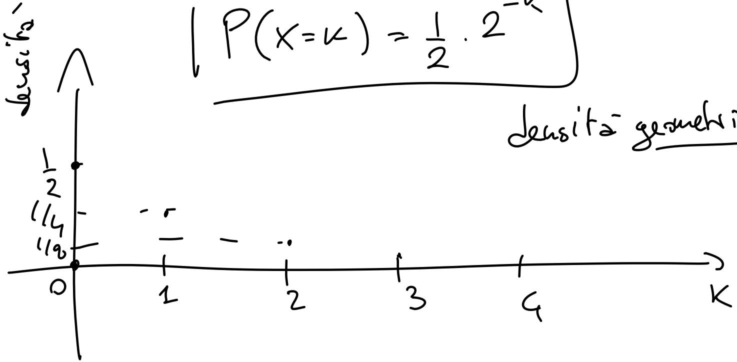
$$\boxed{P(X=x) = \frac{1}{x+1} \quad x \in \mathbb{N}}$$

Esempio $P(X=k) \propto 2^{-k}$ $k = 0, 1, 2, \dots$

$$\sum_{k=0}^{+\infty} 2^{-k} = \frac{1}{1-1/2} = 2$$

$$\boxed{P(X=k) = \frac{1}{2} \cdot 2^{-k}}$$

densità geometrica



parametro

Densità Poisson. Dato $\lambda > 0$, si dice che X a valori in $\mathbb{N}$ ha densità Poisson (di parametro λ) se vale, per ogni $k = 0, 1, \ldots$,

$$P(X = k|I) = e^{-\lambda}\frac{\lambda^k}{k!} \propto \frac{\lambda^k}{k!}$$

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

Serie di Taylor dell'esponenziale

Dalla densità discreta alla legge

Esempio $P(X < 3 \mid X \text{ Poisson}(\lambda)) =$

$$= P(X \in \{0,1,2\} \mid X \text{ Poisson}(\lambda)) = e^{-\lambda}\left(\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!}\right)$$

Se X assume valori in un insieme E finito oppure infinito discreto, vale per ogni $U \subseteq E$,

$$P(X \in U|I) = \sum_{x \in U} P(X = x|I).$$

$$\{X \in U\} = \bigcup_{x \in U} \{X = x\}$$

Densità continua

Problema possibili valori E sono un infinito "continuo", es. $E = \mathbb{R}$, $E = [a, b]$.

- Esempio: X "uniforme" su tutti i valori dell'intervallo $[0, 1]$ *non* possiamo definire

$$P(X = x|I) = c > 0$$

Densità continua

Problema possibili valori E sono un infinito "continuo", es. $E = \mathbb{R}$, $E = [a, b]$.

- Esempio: X "uniforme" su tutti i valori dell'intervallo $[0, 1]$ *non* possiamo definire
$$P(X = x|I) = c > 0$$
- Idea: definire una probabilità "inifinitesima" in x.

Funzione di *densità continua* di probabilità della variabile X,

$$\mathbb{R} \supseteq E \ni x \mapsto f(x)$$

tale che per ogni intervallo $(a, b) \subseteq \mathbb{R}$,

$$P(a < X < b|I) = \int_a^b f(x)dx.$$

- Deve valere $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)dx = 1$.

Funzione di *densità continua* di probabilità della variabile X,

$$E \ni x \mapsto f(x)$$

tale che per ogni intervallo $(a, b) \subseteq \mathbb{R}$,

$$P(a < X < b|I) = \int_a^b f(x) dx.$$

- Deve valere $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x) dx = 1$.
- Non è necessario $f(x) \leq 1$.

Funzione di *densità continua* di probabilità della variabile X,

$$E \ni x \mapsto f(x)$$

tale che per ogni intervallo $(a, b) \subseteq \mathbb{R}$,

$$P(a < X < b|I) = \int_a^b f(x)dx.$$

- Deve valere $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)dx = 1$.
- Non è necessario $f(x) \leq 1$.
- Non è richiesto che f sia continua.

Funzione di *densità continua* di probabilità della variabile X,

$$E \ni x \mapsto f(x)$$

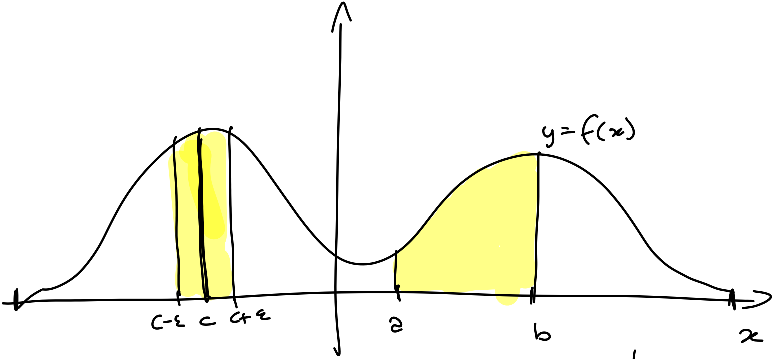
tale che per ogni intervallo $(a, b) \subseteq \mathbb{R}$,

$$P(a < X < b|I) = \int_a^b f(x)dx.$$

- Deve valere $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)dx = 1$.
- Non è necessario $f(x) \leq 1$.
- Non è richiesto che f sia continua.
- Notazione più espressiva: $f(x) = \underline{\underline{p(X = x|I)}}$.

P_X f_X

Probabilità = area del sottografico



$$P(\{a < X < b\}) = \int_a^b f(t)\,dt$$

$$P(X = c \mid I) = \int_c^c f(t)\,dt = 0 \qquad P(X \in (c-\varepsilon, c+\varepsilon)) > 0$$

Esempi

Notazione $\quad P(X=x|I) \propto g(x) \quad x \in E$

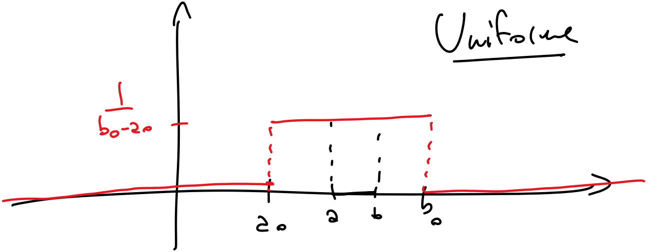
$$\Updownarrow$$

$$P(X=x|I) = \frac{g(x)}{\int_E g(t)\,dt} \qquad \int_E g(t)\,dt < +\infty$$

Densità continua uniforme $\quad E = [a_0, b_0]$

$$P(X=x|I) = \begin{cases} \frac{1}{b_0-a_0} & \text{se } x \in [a_0,b_0] \\ \boxed{0 \quad \text{altrimenti}} \end{cases}$$

Uniforme



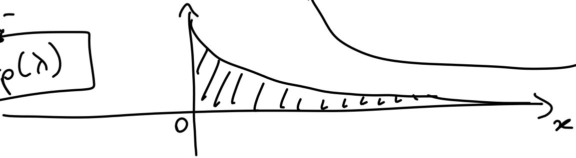
$$P(X \in (a,b)) = \frac{b-a}{b_0-a_0} \qquad (a,b) \subseteq (a_0,b_0)$$

$\lambda > 0$

$$P(X = x \mid I) = \lambda e^{-\lambda x} \qquad x \in E = [0, +\infty)$$

Densità

$\boxed{\mathrm{Exp}(\lambda)}$



$$\int_0^{+\infty} e^{-\lambda t}\, dt \overset{y=\lambda t}{=} \int_0^{+\infty} e^{-y} \frac{dy}{\lambda} = \frac{1}{\lambda}$$

Caso vettoriale

Estendiamo al caso vettoriale (richiede integrali in più variabili, ma non li chiederò negli esercizi).

- X a valori in $\mathbb{R}^d$ ha densità continua f (rispetto all'informazione I) se vale, per ogni "rettangolo"

$$U = \underline{(a_1, b_1) \times \ldots \times (a_d, b_d)} \subseteq \mathbb{R}^d,$$

$$P(X \in U|I) = \int_U f = \int_{a_1}^{b_1} dx_1 \ldots \int_{a_d}^{b_d} dx_d \, f(x_1, \ldots, x_d).$$

Caso vettoriale

Estendiamo al caso vettoriale (richiede integrali in più variabili, ma non li chiederò negli esercizi).

- X a valori in $\mathbb{R}^d$ ha densità continua f (rispetto all'informazione I) se vale, per ogni "rettangolo"

$$U = (a_1, b_1) \times \ldots \times (a_d, b_d) \subseteq \mathbb{R}^d,$$

$$P(X \in U|I) = \int_U f = \int_{a_1}^{b_1} dx_1 \ldots \int_{a_d}^{b_d} dx_d\, f(x_1, \ldots, x_d).$$

- notazione $f(x) = p(X = x|I)$.