Probabilità e Processi Stocastici (455AA)

Lezione 4

Dario Trevisan – *https://web.dm.unipi.it/trevisan*

01/10/2025

Problemi (dalla lezione precedente)

Problema 5 $N=40$ $R=10$ $B=30$

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).

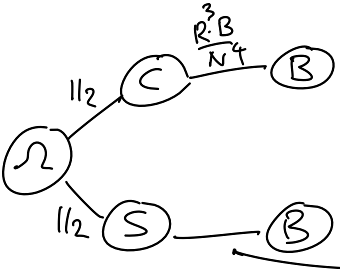
1. Dire se è più probabile che le estrazioni siano state effettuate con rimpiazzo o senza.
2. Calcolare la probabilità che la pallina successiva estratta dall'amico sia rossa.

Alternativa

C = "estrazioni con rimpiazzo"

S = "estrazioni senza rimpiazzo"

B = "R1, R2, B3, R4"



$$P(B|C) = \frac{R}{N}\cdot\frac{R}{N}\cdot\frac{B}{N}\cdot\frac{R}{N} = \frac{R^3\cdot B}{N^4}$$

$$P(B|S) = \frac{R}{N}\,\frac{(R-1)}{(N-1)}\cdot\frac{B}{(N-2)}\,\frac{(R-2)}{N-3}$$

$$P(S|B) = P(S|\Omega)\cdot L(S;B)/P(B)$$

$$P(C|B) = P(C|\Omega)\cdot L(C;B)/P(B)$$

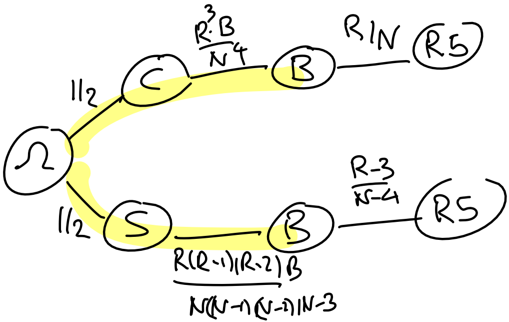
$$L(S;B) = P(B|S) = \frac{R(R-1)(R-2)\cdot B}{N(N-1)(N-2)(N-3)}$$

$$L(C;B) = P(C|S) = \frac{R^3\cdot B}{N^4}$$

$\hookrightarrow$ $R=10$, $B=30$, $N=40$ $\rightsquigarrow$ ----

---

$$P(R5|B) \overset{?}{=} \frac{P(R5,B|\Omega)}{P(B|\Omega)} =$$



$$P(B,R5) = \frac{1}{2} \cdot \frac{R^3 \cdot B}{N^4} \cdot \frac{R}{N} + \frac{1}{2} \frac{R(R-1)(R-2)(R-3) \cdot B}{N(N-1)(N-2)(N-3)(N-4)} \approx$$

$$P(B) = \frac{1}{2} \cdot \frac{R^3 B}{N^4} + \frac{1}{2} \frac{R(R-1)(R-2) \cdot B}{N(N-1)(N-2)(N-3)} \approx$$

$$\longrightarrow \quad \frac{P(B,R5)}{P(B)} \approx$$

Bayes / Disintegration

$$P(E|I) = \sum_{i=1}^{n} P(E|I, A_i)\, P(A_i|I)$$

$$P(R5|B) = \overbrace{P(R5|\cancel{B}, C)}^{R/N}\, P(C|B) + \underbrace{P(R5|B,S)}_{= \frac{R-3}{N-4}}\, P(S|B)$$

Bayes

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.

X = non noto

x = valore possibile

$\{R1, B1\}$ $X_1 \in \{\text{rosso}, \text{blu}\}$

$R1 =$ "$X_1 = \text{rosso}$" $B1 =$ "$X_1 = \text{blu}$"

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 $(\{X = x\})_{x \in E}$.

Se vale $X = 5$ non vale $X = 6$

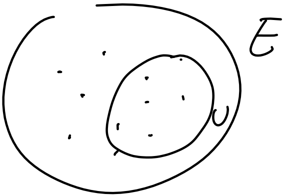
Diciamo che X è *una variabile aleatoria a valori in* E, e usiamo la notazione

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

- 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

$\{X \in U\}$ = "Il valore di X è uno dei valori $x \in U$"

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



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)

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

  .

Se $E = \mathbb{R}$ $\quad X \in E$

$$\{0 < X < 10\} = \{X \in (0, 10)\}$$

intervallo aperto

$$\{X \leq -100\} = \{X \in (-\infty, 100]\}$$

$$\{X^2 = 2\} = \{X \in \{\sqrt{2}, -\sqrt{2}\}\}$$

$$\{X^2 \geq 3,\ X^3 \leq 40\} = \{X \in \cup\}$$

trovare?

Insieme $\{1, 2, 3, 4, 5 \dots\}$

Intervalli $(3, 5) =]3, 5[\qquad [3, 5]$

Variabili aleatorie secondo Kolmogorov

$\boxed{X \in E}$

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

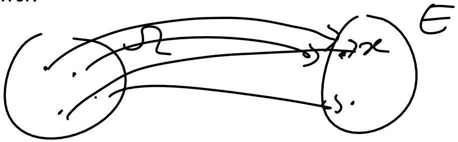
$$X : \Omega \to E$$

tale l'immagine inversa

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

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

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

o più in generale

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

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

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

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

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?

OSS Se $\# E = d$ l'insieme $\mathcal{P}(E)$ dominio della legge contiene 2^d elementi

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 $P(A_i|I)$.

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

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

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

p.m.f.

$[0,1]$

- 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

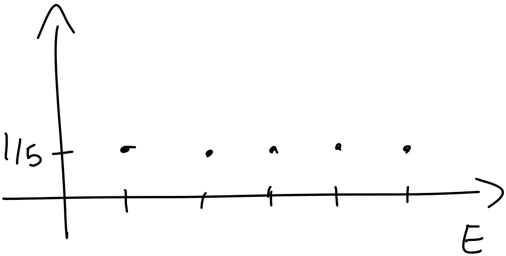
$$P(X = x|I) = f(x) / \sum_{y \in E} f(y).$$

Esempi

Densità Uniforme $\quad X \in E \quad E$ finito $\#E < +\infty$

X ha densità uniforme $\quad P(X = x \mid I) \propto 1$

$$P(X = x \mid I) = \frac{1}{\#E}$$

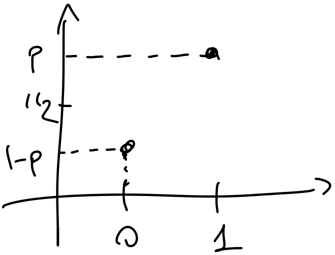


Densità Bernoulli $\quad X \in \{0, 1\}$

$$\left.\begin{array}{l} A = A_1 = \{X = 1\} \\ \text{non}A = A_0 = \{X = 0\} \end{array}\right\}$$ variabile indicatrice di A

$$p = P(A) = P(X = 1) \rightsquigarrow 1 - p = P(X = 0)$$

$p \in [0, 1]$



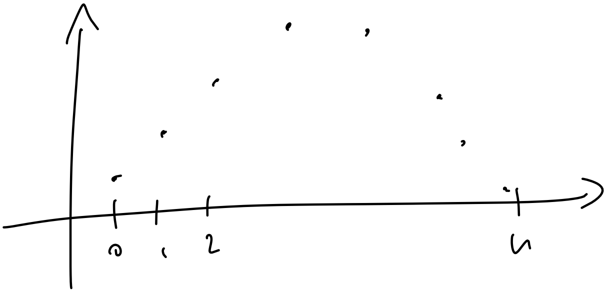
Densità binomiale $\quad X \in \{0,1,2,\dots,n\}$

numero di esperimenti

n

$p \in [0,1] \leftarrow$ prob. di successo

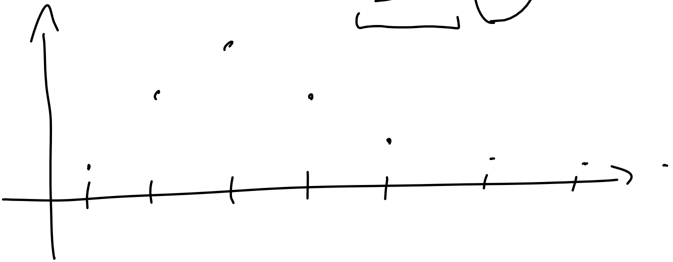
$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{per } k = 0,1,2,\dots,n$$



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

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

$$P(X = k|I) = \underbrace{e^{-\lambda}\frac{\lambda^k}{k!}} \propto \frac{\lambda^k}{k!} > 0$$



Dalla densità discreta alla legge

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

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

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

regola della somma

$P($X sia pari $|$ X è $\mathrm{Bin}(6, 1/3))$ $\longrightarrow$ esercizio

$P($ X sia primo $|$ $X \sim \mathrm{Poisson}(10))$ $\longrightarrow$ esercizio

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

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

$$P(X = x|I) \propto 1$$

$$\sum_{x \in [0,1]} c = +\infty \qquad c > 0$$

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.

Definizione

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

$$E \ni x \mapsto f(x) \qquad [\text{Probabilità}/\text{unità della } X]$$

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

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

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

Non vale sempre $f(x) \leq 1$

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 $f(x) \leq 1$.

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 $(a, b) \subseteq \mathbb{R}$,

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

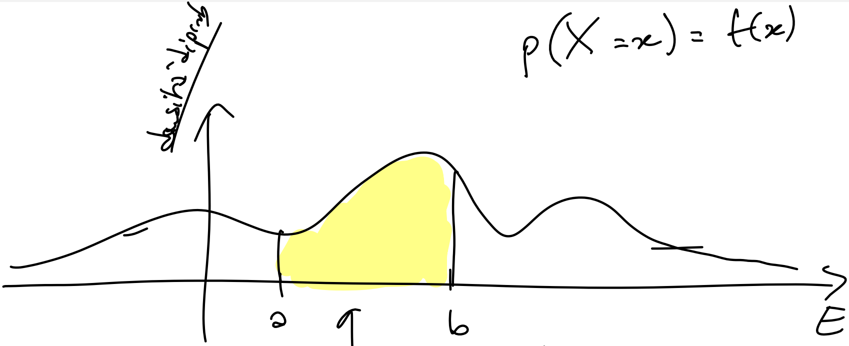
- Deve valere $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)dx = 1$.
- Non è necessario $f(x) \leq 1$.
- Non è richiesto che f sia continua.
- Notazione più espressiva: $f(x) = p(X = x|I)$. p.d.f.

$= \rho_X(x) = \rho(x)$

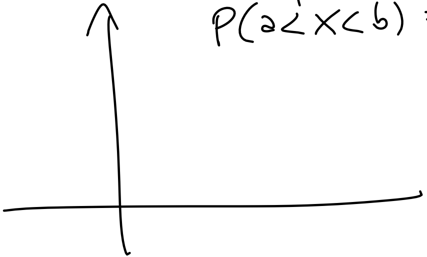
Probabilità = area del sottografico

$E = \mathbb{R}$

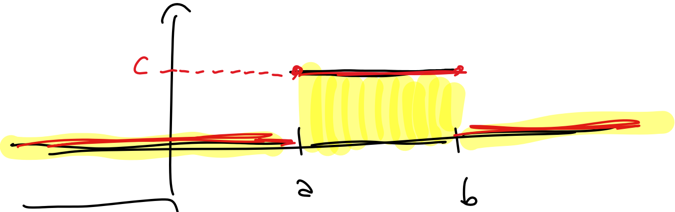
$p(X = x) = f(x)$



$$P(a < X < b) = \int_a^b p(X = x)\,dx$$



Densità uniforme $\quad [a, b] \subseteq \mathbb{R}$



$$\int_{-\infty}^{+\infty} f(x)\,dx = c \cdot (b-a) = 1$$

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

$$b-a = \frac{1}{2} \quad \Downarrow \quad c = 2 > 1$$

$$p(X=x) \propto g(x) \quad \rightsquigarrow \quad p(X=x) = \frac{g(x)}{\int_{-\infty}^{+\infty} g(z)\,dz}$$

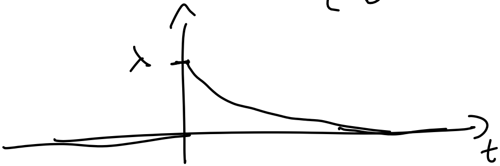
Esempio densità esponenziale

$E = [0, +\infty)$

T ha densità esponenziale di intensità $\lambda > 0$

Se ha densità continua

$$p(T=t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{se } t<0 \end{cases}$$



Calcolate $\int_0^{+\infty} e^{-\lambda t}\,dt = ?$

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"

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

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

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