

Probabilità e Processi Stocastici (455AA)

Lezione 5

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

6/10/2025

- ▶ Una variabile aleatoria X a valori in E identifica un sistema di alternative

$\{X = x\} = \text{“la quantità } X \text{ assume il valore } x\text{”}$

- ▶ Una variabile aleatoria X a valori in E identifica un sistema di alternative

$\{X = x\}$ = “la quantità X assume il valore x ”

- ▶ Notazione generale $\{X \in U\}$, $U \subseteq E$.

- ▶ Una variabile aleatoria X a valori in E identifica un sistema di alternative

$\{X = x\} = \text{“la quantità } X \text{ assume il valore } x\text{”}$

- ▶ Notazione generale $\{X \in U\}$, $U \subseteq E$.
- ▶ Legge di una variabile $U \mapsto P(X \in U|I)$

- ▶ Una variabile aleatoria X a valori in E identifica un sistema di alternative

$\{X = x\}$ = “la quantità X assume il valore x ”

- ▶ Notazione generale $\{X \in U\}$, $U \subseteq E$.

- ▶ Legge di una variabile $U \mapsto P(X \in U|I)$

- ▶ Densità discreta

$$(P(X = x|I))_{x \in E} \Rightarrow P(X \in U) = \sum_{x \in U} P(X = x)$$

- ▶ Una variabile aleatoria X a valori in E identifica un sistema di alternative

$\{X = x\} = \text{“la quantità } X \text{ assume il valore } x\text{”}$

- ▶ Notazione generale $\{X \in U\}$, $U \subseteq E$.

- ▶ Legge di una variabile $U \mapsto P(X \in U|I)$

- ▶ Densità discreta

$$(P(X = x|I))_{x \in E} \Rightarrow P(X \in U) = \sum_{x \in U} P(X = x)$$

- ▶ Densità continua

$$(p(X = x|I))_{x \in E} \Rightarrow P(X \in U) = \int_U P(X = x)dx$$

- ▶ Una variabile aleatoria X a valori in E identifica un sistema di alternative

$\{X = x\} = \text{"la quantità } X \text{ assume il valore } x\text{"}$

- ▶ Notazione generale $\{X \in U\}$, $U \subseteq E$.

- ▶ Legge di una variabile $U \mapsto P(X \in U|I)$

- ▶ Densità discreta

$$(P(X = x|I))_{x \in E} \Rightarrow P(X \in U) = \sum_{x \in U} P(X = x)$$

- ▶ Densità continua

$$(p(X = x|I))_{x \in E} \Rightarrow P(X \in U) = \int_U P(X = x)dx$$

- ▶ **ATTENZIONE:** esistono variabili X né discrete né continue!

Caso vettoriale

Estendiamo la densità continua al caso vettoriale (richiede integrali in più variabili, ma non richiesti negli esercizi).

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

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

l'identità

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

Caso vettoriale

Estendiamo la densità continua al caso vettoriale (richiede integrali in più variabili, ma non richiesti negli esercizi).

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

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

l'identità

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

- ▶ notazione $f(x) = p(X = x | I)$.

Trasformazioni di variabili aleatorie

Composizione tramite funzione

Sia X a valori in E e $g : E \rightarrow F$ una funzione.

- Definiamo la **variabile composta** $g(X)$ tramite il sistema di alternative, per $z \in F$,

$$\{g(X) = z\} = \{X \in g^{-1}(z)\}.$$

Trasformazione della densità: caso discreto

Se X è discreta (e quindi $g(X)$ è discreta)

$$P(g(X) = z|I) = \sum_{x \in g^{-1}(z)} P(X = x|I).$$

Se X ha densità continua e $g(X)$ è discreta,

$$P(g(X) = z|I) = \int_{g^{-1}(z)} p(X = x|I) dx.$$

Trasformazione della densità: caso continuo

Segue da un cambio di variabile il seguente risultato:

- ▶ Sia X una variabile aleatoria a valori in \mathbb{R} , con densità continua $p(X = x|I)$.

Trasformazione della densità: caso continuo

Segue da un cambio di variabile il seguente risultato:

- ▶ Sia X una variabile aleatoria a valori in \mathbb{R} , con densità continua $p(X = x|I)$.
- ▶ Sia $g : \mathbb{R} \rightarrow \mathbb{R}$ una funzione invertibile, derivabile, con derivata continua e mai nulla $g'(x) \neq 0$.

Trasformazione della densità: caso continuo

Segue da un cambio di variabile il seguente risultato:

- ▶ Sia X una variabile aleatoria a valori in \mathbb{R} , con densità continua $p(X = x|I)$.
- ▶ Sia $g : \mathbb{R} \rightarrow \mathbb{R}$ una funzione invertibile, derivabile, con derivata continua e mai nulla $g'(x) \neq 0$.
- ▶ Allora $g(X)$ ammette densità continua e vale

$$p(g(X) = z|I) = p(X = g^{-1}(z)|I) \cdot \frac{1}{|g'(g^{-1}(z))|}$$

Una estensione (utile nei problemi):

- ▶ se X assume con probabilità 1 valori in un intervallo $E \subseteq \mathbb{R}$

Una estensione (utile nei problemi):

- ▶ se X assume con probabilità 1 valori in un intervallo $E \subseteq \mathbb{R}$
- ▶ $g : E \rightarrow \mathbb{R}$ è tale che si può decomporre E in una unione finita di intervalli a due a due disgiunti in cui, all'interno di ciascun intervallo, g sia invertibile, derivabile con derivata continua e mai nulla $g'(x) \neq 0$.

Una estensione (utile nei problemi):

- ▶ se X assume con probabilità 1 valori in un intervallo $E \subseteq \mathbb{R}$
- ▶ $g : E \rightarrow \mathbb{R}$ è tale che si può decomporre E in una unione finita di intervalli a due a due disgiunti in cui, all'interno di ciascun intervallo, g sia invertibile, derivabile con derivata continua e mai nulla $g'(x) \neq 0$.
- ▶ Allora $g(X)$ ammette densità continua

$$p(g(X) = z|I) = \sum_{x \in g^{-1}(z)} p(X = x|I) \cdot \frac{1}{|g'(x)|}.$$

Trasformazione della densità: caso vettoriale

- ▶ Sia X una variabile aleatoria vettoriale, a valori in \mathbb{R}^d , con densità continua $p(X = x|I)$.

Trasformazione della densità: caso vettoriale

- ▶ Sia X una variabile aleatoria vettoriale, a valori in \mathbb{R}^d , con densità continua $p(X = x|I)$.
- ▶ Sia $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ una funzione invertibile, derivabile con derivata continua e invertibile in ogni punto, ossia

$$\det \left(\left(\frac{\partial g_i}{\partial x_j}(x) \right)_{i,j=1,\dots,d} \right) \neq 0, \quad \text{per ogni } x \in \mathbb{R}^d.$$

Trasformazione della densità: caso vettoriale

- ▶ Sia X una variabile aleatoria vettoriale, a valori in \mathbb{R}^d , con densità continua $p(X = x|I)$.
- ▶ Sia $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ una funzione invertibile, derivabile con derivata continua e invertibile in ogni punto, ossia

$$\det \left(\left(\frac{\partial g_i}{\partial x_j}(x) \right)_{i,j=1,\dots,d} \right) \neq 0, \quad \text{per ogni } x \in \mathbb{R}^d.$$

- ▶ Allora $g(X)$ ammette densità continua e vale

$$p(g(X) = z|I) = p(X = g^{-1}(z)|I) \cdot \frac{1}{|\det(Dg)(g^{-1}(z))|}$$

Problema

Sia X una variabile con densità uniforme continua su $(-1, 3)$.

1. Scrivere la densità di X

Problema

Sia X una variabile con densità uniforme continua su $(-1, 3)$.

1. Scrivere la densità di X
2. Scrivere la densità di X^2 .

Problema

Sia X una variabile con densità uniforme continua su $(-1, 3)$.

1. Scrivere la densità di X
2. Scrivere la densità di X^2 .
3. Scrivere la densità di X^4 , verificare che il risultato ottenendo applicando la formula di cambio di variabile direttamente da X è lo stesso che applicando in sequenza due volte la formula per $x \mapsto x^2$, ossia $X^4 = (X^2)^2$.

Problema

Sia X una variabile con densità uniforme continua su $(-1, 3)$.

1. Scrivere la densità di X
2. Scrivere la densità di X^2 .
3. Scrivere la densità di X^4 , verificare che il risultato ottenendo applicando la formula di cambio di variabile direttamente da X è lo stesso che applicando in sequenza due volte la formula per $x \mapsto x^2$, ossia $X^4 = (X^2)^2$.
4. La variabile $\max\{0, X\}$ è continua?

Variabile congiunta

Date due variabili aleatorie $X \in E$, $Y \in F$, come definire la variabile *congiunta* (X, Y) a valori nelle coppie ordinate $(x, y) \in E \times F$?

- *Esempio*: si vuole applicare una funzione $g(x, y)$.

Variabile congiunta

Date due variabili aleatorie $X \in E$, $Y \in F$, come definire la variabile *congiunta* (X, Y) a valori nelle coppie ordinate $(x, y) \in E \times F$?

- ▶ *Esempio*: si vuole applicare una funzione $g(x, y)$.
- ▶ Dai sistemi di alternative $(\{X = x\})_{x \in E}$, $(\{Y = y\})_{y \in F}$, si definisce la variabile aleatoria *congiunta* (X, Y) a valori in $E \times F$ tramite il sistema di alternative

$$\{(X, Y) = (x, y)\} = \{X = x, Y = y\} = \{X = x\} \text{ e } \{Y = y\}.$$

Le variabili X , Y (considerate separatamente) sono le *marginali* della variabile congiunta (X, Y) .

Dalla legge congiunta alle leggi marginali

- ▶ Dalla legge congiunta è sempre possibile ottenere le leggi delle marginali, tramite composizione di opportune funzioni.

Dalla legge congiunta alle leggi marginali

- ▶ Dalla legge congiunta è sempre possibile ottenere le leggi delle marginali, tramite composizione di opportune funzioni.
- ▶ La variabile X è ottenibile dalla (X, Y) tramite la funzione

$$g : E \times F \rightarrow E, \quad (x, y) \mapsto x.$$

Dalla legge congiunta alle leggi marginali

- ▶ Dalla legge congiunta è sempre possibile ottenere le leggi delle marginali, tramite composizione di opportune funzioni.
- ▶ La variabile X è ottenibile dalla (X, Y) tramite la funzione

$$g : E \times F \rightarrow E, \quad (x, y) \mapsto x.$$

- ▶ La variabile Y è ottenibile dalla (X, Y) tramite la funzione

$$g : E \times F \rightarrow F, \quad (x, y) \mapsto y.$$

- Troviamo quindi, nel caso discreto,

$$P(X = x|I) = \sum_{y \in F} P(X = x, Y = y|I) = \sum_{y \in F} P((X, Y) = (x, y)|I),$$

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

Nel caso continuo $E = \mathbb{R}^d$, $F = \mathbb{R}^k$,

$$p(X = x|I) = \int_{\mathbb{R}^k} p((X, Y) = (x, y)|I) dy.$$

$$p(Y = y|I) = \int_{\mathbb{R}^d} p((X, Y) = (x, y)|I) dx.$$

Dalle leggi marginali alla legge congiunta?

- ▶ La conoscenza delle leggi di X , Y (separatamente) **non** determina la legge di (X, Y) .

Estensione al caso di k variabili aleatorie $X_1 \in E_1, \dots, X_k \in E_k$.

- La variabile congiunta $X = (X_1, \dots, X_k)$ a valori nelle k -uple ordinate

$$x = (x_1, \dots, x_k) \in E_1 \times \dots \times E_k$$

è definita tramite il sistema di alternative

$$\begin{aligned}\{X = x\} &= \{X_1 = x_1, X_2 = x_2, \dots, X_k = x_k\} \\ &= \{X_1 = x_1\} \text{ e } \dots \{X_k = x_k\}.\end{aligned}$$

Densità congiunta e regola del prodotto

- ▶ La densità di una variabile congiunta non è esprimibile in termini delle densità marginali

Densità congiunta e regola del prodotto

- ▶ La densità di una variabile congiunta non è esprimibile in termini delle densità marginali
- ▶ Dalla regola del prodotto si ha

$$\begin{aligned}P((X, Y) = (x, y)) &= P(X = x \text{ e } Y = y) \\&= P(X = x)P(Y = y|X = x).\end{aligned}$$

Densità congiunta e regola del prodotto

- ▶ La densità di una variabile congiunta non è esprimibile in termini delle densità marginali
- ▶ Dalla regola del prodotto si ha

$$\begin{aligned}P((X, Y) = (x, y)) &= P(X = x \text{ e } Y = y) \\&= P(X = x)P(Y = y|X = x).\end{aligned}$$

- ▶ Nel caso di variabili discrete, la densità della variabile congiunta è determinata da

Densità congiunta e regola del prodotto

- ▶ La densità di una variabile congiunta non è esprimibile in termini delle densità marginali
- ▶ Dalla regola del prodotto si ha

$$\begin{aligned}P((X, Y) = (x, y)) &= P(X = x \text{ e } Y = y) \\&= P(X = x)P(Y = y|X = x).\end{aligned}$$

- ▶ Nel caso di variabili discrete, la densità della variabile congiunta è determinata da
 1. La densità della marginale X

Densità congiunta e regola del prodotto

- ▶ La densità di una variabile congiunta non è esprimibile in termini delle densità marginali
- ▶ Dalla regola del prodotto si ha

$$\begin{aligned}P((X, Y) = (x, y)) &= P(X = x \text{ e } Y = y) \\&= P(X = x)P(Y = y|X = x).\end{aligned}$$

- ▶ Nel caso di variabili discrete, la densità della variabile congiunta è determinata da
 1. La densità della marginale X
 2. La densità della marginale Y , condizionata all'informazione $X = x$ per ciascun $x \in E$ (condizionata ad X).

Densità congiunta e regola del prodotto

- ▶ La densità di una variabile congiunta non è esprimibile in termini delle densità marginali
- ▶ Dalla regola del prodotto si ha

$$\begin{aligned}P((X, Y) = (x, y)) &= P(X = x \text{ e } Y = y) \\ &= P(X = x)P(Y = y|X = x).\end{aligned}$$

- ▶ Nel caso di variabili discrete, la densità della variabile congiunta è determinata da
 1. La densità della marginale X
 2. La densità della marginale Y , condizionata all'informazione $X = x$ per ciascun $x \in E$ (condizionata ad X).
- ▶ In pratica in molti casi la legge congiunta è proprio definita tramite queste due quantità.

Formula di Bayes

Formula di bayes per variabili: caso discreto

La formula di Bayes si può quindi riscrivere

$$\begin{aligned} P(X = x|Y = y) &= P(X = x) \cdot \frac{P(Y = y|X = x)}{P(Y = y)} \\ &\propto P(X = x)L(X = x; Y = y) \end{aligned}$$

- Il denominatore $P(Y = y)$ si calcola imponendo che la somma su x sia 1:

$$P(Y = y) = \sum_{x \in E} P(X = x)L(X = x; Y = y).$$

Stime MAP e MLE: caso discreto

Spesso interessa solo il valore più probabile della X , avendo osservato $Y = y$.

Si definisce la stima per X :

- ▶ di massimo a posteriori

$$x_{\text{MAP}} \in \arg \max \{P(X = x) \cdot L(X = x; Y = y) : x \in E\}$$

Stime MAP e MLE: caso discreto

Spesso interessa solo il valore più probabile della X , avendo osservato $Y = y$.

Si definisce la stima per X :

- ▶ di massimo a posteriori

$$x_{\text{MAP}} \in \arg \max \{P(X = x) \cdot L(X = x; Y = y) : x \in E\}$$

- ▶ di massima verosimiglianza (a priori uniforme)

$$x_{\text{MLE}} \in \arg \max \{L(X = x; Y = y) : x \in E\}.$$

Stime MAP e MLE: caso discreto

Spesso interessa solo il valore più probabile della X , avendo osservato $Y = y$.

Si definisce la stima per X :

- ▶ di massimo a posteriori

$$x_{\text{MAP}} \in \arg \max \{P(X = x) \cdot L(X = x; Y = y) : x \in E\}$$

- ▶ di massima verosimiglianza (a priori uniforme)

$$x_{\text{MLE}} \in \arg \max \{L(X = x; Y = y) : x \in E\}.$$

- ▶ Anche se E è infinito la stima x_{MLE} ha senso (anche se non c'è densità a priori uniforme)

Esempio

- ▶ un'urna contiene metà palline rosse e metà palline blu, da cui si effettuano un certo numero di estrazioni con rimpiazzo.

Esempio

- ▶ un'urna contiene metà palline rosse e metà palline blu, da cui si effettuano un certo numero di estrazioni con rimpiazzo.
- ▶ Il numero esatto delle estrazioni non è noto, ma solo che il numero di palline rosse estratte è 10.

Esempio

- ▶ un'urna contiene metà palline rosse e metà palline blu, da cui si effettuano un certo numero di estrazioni con rimpiazzo.
- ▶ Il numero esatto delle estrazioni non è noto, ma solo che il numero di palline rosse estratte è 10.
- ▶ Possiamo stimare il numero M di estrazione effettuate?

- ▶ A priori uniforme su $\{0, 1, \dots, \bar{m}\}$:

$$P(M = m|\Omega) = \frac{1}{\bar{m}}$$

- ▶ A priori uniforme su $\{0, 1, \dots, \bar{m}\}$:

$$P(M = m|\Omega) = \frac{1}{\bar{m}}$$

- ▶ Verosimiglianza

$$\begin{aligned} L(M = m; N_R = k) &= P(N_R = k|M = m) = \binom{m}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{m-k} \\ &= \binom{m}{k} \frac{1}{2^m}. \end{aligned}$$

- ▶ A priori uniforme su $\{0, 1, \dots, \bar{m}\}$:

$$P(M = m|\Omega) = \frac{1}{\bar{m}}$$

- ▶ Verosimiglianza

$$\begin{aligned} L(M = m; N_R = k) &= P(N_R = k|M = m) = \binom{m}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{m-k} \\ &= \binom{m}{k} \frac{1}{2^m}. \end{aligned}$$

- ▶ Formula di Bayes (per $m \leq \bar{m}$)

$$P(M = m|N_R = k) \propto \binom{m}{k} \frac{1}{2^m}.$$

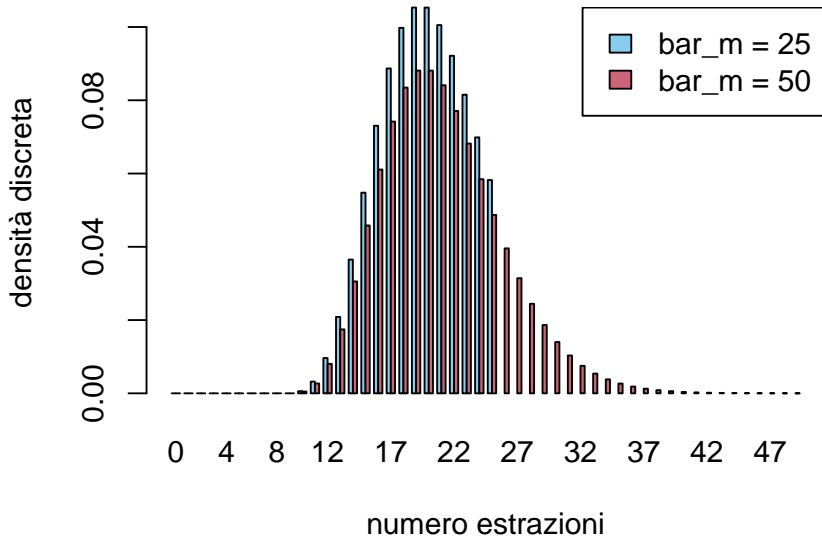


Figure 1: grafico della densità discreta di M sapendo $N_R = 10$

Caso variabile congiunta continua

Si introduce una “regola del prodotto” nel caso in cui vi sia la densità congiunta:

$$\begin{aligned} p((X, Y) = (x, y)) &= p(X = x, Y = y) \\ &= p(X = x)p(Y = y|X = x), \end{aligned}$$

- il termine $p(Y = y|X = x)$ è la *densità condizionale di Y rispetto ad X*

Caso variabile congiunta continua

Si introduce una “regola del prodotto” nel caso in cui vi sia la densità congiunta:

$$\begin{aligned} p((X, Y) = (x, y)) &= p(X = x, Y = y) \\ &= p(X = x)p(Y = y|X = x), \end{aligned}$$

- ▶ il termine $p(Y = y|X = x)$ è la *densità condizionale di Y rispetto ad X*
- ▶ formula di Kolmogorov nel caso continuo:

$$p(Y = y|X = x) = \frac{p(X = x, Y = y)}{p(X = x)}.$$

- formula di Bayes per densità continue:

$$\begin{aligned} p(X = x|Y = y) &= p(X = x) \cdot \frac{p(Y = y|X = x)}{p(Y = y)} \\ &\propto p(X = x)L(X = x; Y = y) \end{aligned}$$

dove

$$p(Y = y) = \int_{\mathbb{R}^d} p(Y = y|X = x)p(X = x)dx.$$

- formula di Bayes per densità continue:

$$\begin{aligned} p(X = x|Y = y) &= p(X = x) \cdot \frac{p(Y = y|X = x)}{p(Y = y)} \\ &\propto p(X = x)L(X = x; Y = y) \end{aligned}$$

dove

$$p(Y = y) = \int_{\mathbb{R}^d} p(Y = y|X = x)p(X = x)dx.$$

- Stima del massimo a posteriori

$$x_{\text{MAP}} \in \arg \max \{p(X = x)L(X = x; Y = y), x \in E\}$$

- ▶ formula di Bayes per densità continue:

$$\begin{aligned} p(X = x|Y = y) &= p(X = x) \cdot \frac{p(Y = y|X = x)}{p(Y = y)} \\ &\propto p(X = x)L(X = x; Y = y) \end{aligned}$$

dove

$$p(Y = y) = \int_{\mathbb{R}^d} p(Y = y|X = x)p(X = x)dx.$$

- ▶ Stima del massimo a posteriori

$$x_{\text{MAP}} \in \arg \max \{p(X = x)L(X = x; Y = y), x \in E\}$$

- ▶ Stima di massima verosimiglianza

$$x_{\text{MLE}} \in \arg \{ \max p(Y = y|X = x) : x \in [a, b] \}.$$

Un esempio

La durata della carica di un dispositivo (ad esempio, uno smartphone, o un drone) è modellizzata da una variabile aleatoria T avente densità continua *esponenziale* di un certo parametro λ , ossia

$$p(T = t) = \lambda e^{-\lambda t} \quad \text{per } t \geq 0,$$

mentre $p(T = t) = 0$ altrimenti.

Per stimare Λ si suppone che sia a priori uniformemente distribuito su $[0, 1]$, ossia, per $\lambda \in [0, 1]$, $p(\Lambda = \lambda) = 1$, mentre la verosimiglianza vale

$$L(\Lambda = \lambda; T = t) = \lambda e^{-\lambda t}.$$

Supponendo di osservare $T = 10$, formula di Bayes implica

$$\begin{aligned} p(\Lambda = \lambda | T = 10) &\propto p(\Lambda = \lambda) L(\Lambda = \lambda; T = 10) \\ &\propto \lambda e^{-\lambda t}. \end{aligned}$$

Imponendo che sia una densità (per $\lambda \in [0, 1]$), troviamo il denominatore

$$p(T = 10) = \int_0^1 \lambda e^{-10\lambda} d\lambda.$$

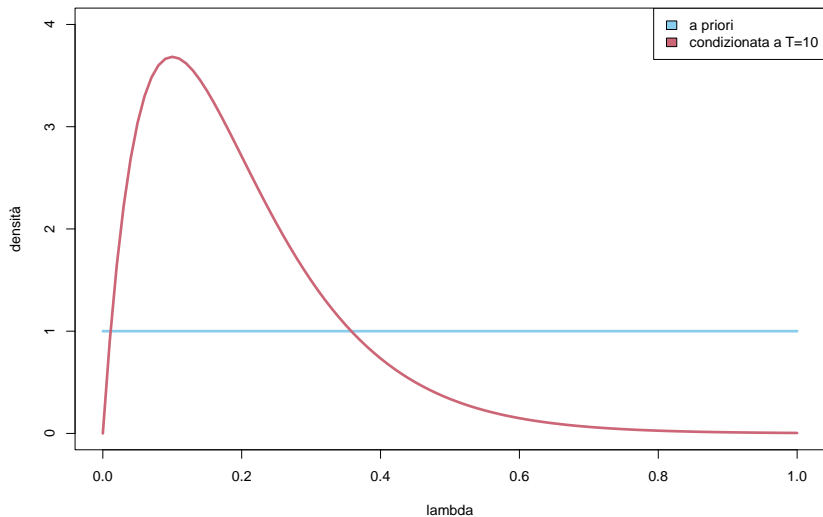


Figure 2: densità per Λ : a priori e avendo osservato $T = 10$

Per la stima di massima verosimiglianza, si deve determinare

$$\lambda_{\text{MLE}} \in \arg \max \left\{ \lambda e^{-10\lambda} : \lambda \in [0, 1] \right\},$$

- Basta derivare e imporre che la derivata sia nulla, trovando

$$e^{-10\lambda} - 10\lambda e^{-10\lambda} = 0 \quad \text{ossia} \quad \lambda = 1/10.$$

Per la stima di massima verosimiglianza, si deve determinare

$$\lambda_{\text{MLE}} \in \arg \max \left\{ \lambda e^{-10\lambda} : \lambda \in [0, 1] \right\},$$

- ▶ Basta derivare e imporre che la derivata sia nulla, trovando

$$e^{-10\lambda} - 10\lambda e^{-10\lambda} = 0 \quad \text{ossia} \quad \lambda = 1/10.$$

- ▶ (andrebbe anche verificato che il massimo non sia raggiunto agli estremi dell'intervallo, ma dal grafico sopra è evidente).

Formula di Bayes: casi misti

Menzioniamo anche il caso misto, in cui $X \in E$ ha densità discreta mentre Y , condizionata ad $\{X = x\}$, ha densità continua continua, per ogni $x \in E$.

- ▶ Il problema è che in generale la variabile congiunta non ha densità né continua né discreta (bisogna usare un approccio più flessibile)

Formula di Bayes: casi misti

Menzioniamo anche il caso misto, in cui $X \in E$ ha densità discreta mentre Y , condizionata ad $\{X = x\}$, ha densità continua continua, per ogni $x \in E$.

- ▶ Il problema è che in generale la variabile congiunta non ha densità né continua né discreta (bisogna usare un approccio più flessibile)
- ▶ La formula di Bayes rimane valida:

$$P(X = x|Y = y) = P(X = x) \cdot \frac{p(Y = y|X = x)}{p(Y = y)} \\ \propto P(X = x)L(X = x; Y = y).$$

dove

$$p(Y = y) = \sum_{x \in E} p(Y = y|X = x)P(X = x).$$

Formula di Bayes: casi misti

Menzioniamo anche il caso misto, in cui $X \in E$ ha densità discreta mentre Y , condizionata ad $\{X = x\}$, ha densità continua continua, per ogni $x \in E$.

- ▶ Il problema è che in generale la variabile congiunta non ha densità né continua né discreta (bisogna usare un approccio più flessibile)
- ▶ La formula di Bayes rimane valida:

$$P(X = x|Y = y) = P(X = x) \cdot \frac{p(Y = y|X = x)}{p(Y = y)} \\ \propto P(X = x)L(X = x; Y = y).$$

dove

$$p(Y = y) = \sum_{x \in E} p(Y = y|X = x)P(X = x).$$

- ▶ Nel caso simmetrico, ossia X continua e Y discreta:

$$p(X = x|Y = y) \propto p(X = x)L(X = x; Y = y)$$

Un esempio

Un'urna contiene palline rosse oppure blu, ma non è noto né il numero totale né la frazione di palline rosse sul totale.

- ▶ Avendo effettuato 10 estrazioni con rimpiazzo, sono state osservate 3 palline rosse.

Un esempio

Un'urna contiene palline rosse oppure blu, ma non è noto né il numero totale né la frazione di palline rosse sul totale.

- ▶ Avendo effettuato 10 estrazioni con rimpiazzo, sono state osservate 3 palline rosse.
- ▶ Come stimare la frazione di palline rosse sul totale?

Un esempio

Un'urna contiene palline rosse oppure blu, ma non è noto né il numero totale né la frazione di palline rosse sul totale.

- ▶ Avendo effettuato 10 estrazioni con rimpiazzo, sono state osservate 3 palline rosse.
- ▶ Come stimare la frazione di palline rosse sul totale?
- ▶ Variabile aleatoria $F_R \in [0, 1]$ per indicare la frazione di palline rosse, con densità a priori uniforme.

Un esempio

Un'urna contiene palline rosse oppure blu, ma non è noto né il numero totale né la frazione di palline rosse sul totale.

- ▶ Avendo effettuato 10 estrazioni con rimpiazzo, sono state osservate 3 palline rosse.
- ▶ Come stimare la frazione di palline rosse sul totale?
- ▶ Variabile aleatoria $F_R \in [0, 1]$ per indicare la frazione di palline rosse, con densità a priori uniforme.
- ▶ Verosimiglianza (binomiale)

$$L(F_R = r; N_R = k) = P(N_R = k | F_R = r) = \binom{10}{k} r^k (1-r)^{10-k}.$$

La formula di Bayes (nel caso “misto”) implica, per $r \in [0, 1]$,

$$p(F_R = r | N_R = 3) = p(F_R = r | \Omega) \frac{L(F_R = r; N_R = 3)}{P(N_R = 3 | \Omega)} \propto \binom{10}{3} r^3 (1-r)^7,$$

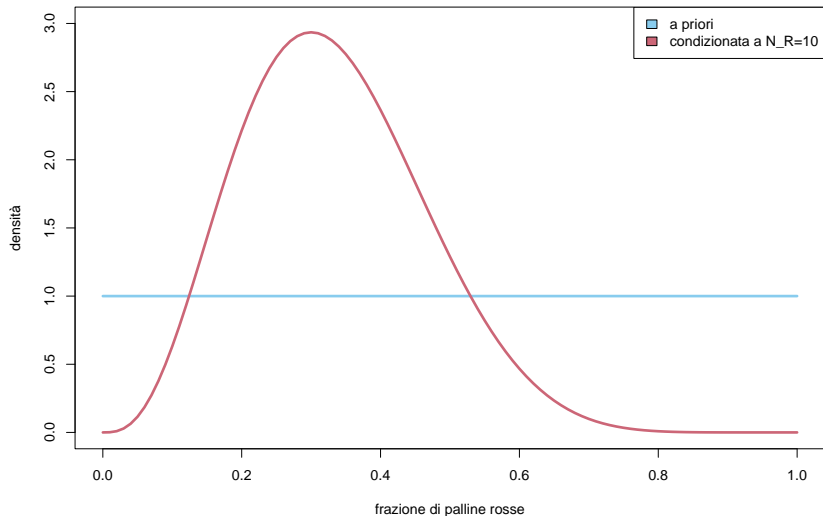


Figure 3: densità continua della frazione di palline rosse a priori e avendo osservato in 10 estrazioni con rimpiazzo 3 palline rosse

Indipendenza

Indipendenza: caso discreto

L'indipendenza tra sistemi di alternative si traduce per variabili discrete:

- Siano $X_1 \in E_1, \dots, X_k \in E_k$ variabili aleatorie con densità discreta (rispetto ad una informazione nota I). Allora esse si dicono indipendenti (condizionatamente ad I) se vale

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k | I) = \prod_{i=1}^k P(X_i = x_i | I),$$

per ogni $x_1 \in E_1, x_2 \in E_2, \dots, x_k \in E_k$, o equivalentemente, per ogni sottoinsieme $J \subseteq \{1, \dots, k\}$,

$$\begin{aligned} &P(X_j = x_j \text{ per ogni } j \in J | I, X_\ell = x_\ell \text{ per ogni } \ell \notin J) \\ &= P(X_j = x_j \text{ per ogni } j \in J | I). \end{aligned}$$

Indipendenza: caso discreto

L'indipendenza tra sistemi di alternative si traduce per variabili discrete:

- ▶ Siano $X_1 \in E_1, \dots, X_k \in E_k$ variabili aleatorie con densità discreta (rispetto ad una informazione nota I). Allora esse si dicono indipendenti (condizionatamente ad I) se vale

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k | I) = \prod_{i=1}^k P(X_i = x_i | I),$$

per ogni $x_1 \in E_1, x_2 \in E_2, \dots, x_k \in E_k$, o equivalentemente, per ogni sottoinsieme $J \subseteq \{1, \dots, k\}$,

$$\begin{aligned} &P(X_j = x_j \text{ per ogni } j \in J | I, X_\ell = x_\ell \text{ per ogni } \ell \notin J) \\ &= P(X_j = x_j \text{ per ogni } j \in J | I). \end{aligned}$$

- ▶ Il membro a sinistra è la densità discreta della variabile congiunta (X_1, \dots, X_k) , mentre a destra abbiamo il prodotto delle densità discrete delle marginali.

Indipendenza: caso continuo

- Siano $X_1 \in \mathbb{R}^{d_1}, \dots, X_k \in \mathbb{R}^{d_k}$ variabili aleatorie con densità continua (rispetto ad una informazione nota I). Allora esse si dicono indipendenti (condizionatamente ad I) se la variabile congiunta $X = (X_1, \dots, X_k)$ ammette densità continua e vale

$$p(X = x|I) = p(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k|I) = \prod_{i=1}^k p(X_i = x_i|I),$$

per ogni $x_1 \in \mathbb{R}^{d_1}, x_2 \in \mathbb{R}^{d_2}, \dots, x_k \in \mathbb{R}^{d_k}$, o
equivalentemente, per ogni sottoinsieme $J \subseteq \{1, \dots, k\}$,

$$\begin{aligned} & p(X_j = x_j \text{ per ogni } j \in J | I, X_\ell = x_\ell \text{ per ogni } \ell \notin J) \\ &= p(X_j = x_j \text{ per ogni } j \in J | I). \end{aligned}$$

Indipendenza: caso generale

- Possiamo immaginare definizioni valide anche per i casi “misti”, in cui alcune variabili sono discrete e altre continue. Ma è possibile dare una definizione generale (che include quelle sopra).

Indipendenza: caso generale

- Possiamo immaginare definizioni valide anche per i casi “misti”, in cui alcune variabili sono discrete e altre continue. Ma è possibile dare una definizione generale (che include quelle sopra).
- Siano $X_1 \in E_1, \dots, X_k \in E_k$ variabili aleatorie (generali). Allora esse si dicono indipendenti (condizionatamente ad una informazione nota I) se vale

$$P(X_1 \in U_1, X_2 \in U_2, \dots, X_k \in U_k | I) = \prod_{i=1}^k P(X_i \in U_i | I),$$

per ogni $U_1 \subseteq E_1, U_2 \subseteq E_2, \dots, U_k \subseteq E_k$, o equivalentemente, per ogni sottoinsieme $J \subseteq \{1, \dots, k\}$,

$$\begin{aligned} &P(X_j \in U_j \text{ per ogni } j \in J | I, X_\ell \in U_\ell \text{ per ogni } \ell \notin J) \\ &= P(X_j \in U_j \text{ per ogni } j \in J | I). \end{aligned}$$

Indipendenza e composizione

Vale il seguente risultato:

- ▶ Siano $X_1 \in E_1, \dots, X_k \in E_k$ variabili aleatorie (generali). Allora esse sono indipendenti (condizionatamente ad una informazione nota I) se e solo se, dato un qualsiasi sottoinsieme $J \subseteq \{1, \dots, k\}$, qualsiasi affermazione A associata alle variabili $\{X_j\}_{j \in J}$ è indipendente (sapendo I) da qualsiasi affermazione B associata alle rimanenti variabili $\{X_\ell\}_{\ell \in \{1, \dots, k\} \setminus J}$.

Indipendenza e composizione

Vale il seguente risultato:

- ▶ Siano $X_1 \in E_1, \dots, X_k \in E_k$ variabili aleatorie (generali). Allora esse sono indipendenti (condizionatamente ad una informazione nota I) se e solo se, dato un qualsiasi sottoinsieme $J \subseteq \{1, \dots, k\}$, qualsiasi affermazione A associata alle variabili $\{X_j\}_{j \in J}$ è indipendente (sapendo I) da qualsiasi affermazione B associata alle rimanenti variabili $\{X_\ell\}_{\ell \in \{1, \dots, k\} \setminus J}$.
- ▶ Una **conseguenza fondamentale**: ogni variabile ottenuta tramite funzione delle sole $(X_j)_{j \in J}$, è indipendente da ogni variabile ottenuta tramite funzione delle sole $(X_\ell)_{\ell \notin J}$.

